Bulletin of the Seismological Society of America

Ⓔ

From Relative to Absolute Teleseismic Travel Times: The Absolute Arrival-Time Recovery Method (AARM)

by Alistair Boyce, Ian D. Bastow, Stéphane Rondenay, and Robert D. Van der Hilst

Abstract   Dense, short-term deployments of seismograph networks are frequently used to study upper-mantle structure. However, recordings of variably emergent teleseismic waveforms are often of lower signal-to-noise ratio (SNR) than those recorded at permanent observatory sites. Therefore, waveform coherency across a network is frequently utilized to calculate relative arrival times between recorded traces, but these measurements cannot easily be combined or reported directly to global absolute arrival-time databases. These datasets are thus a valuable but untapped resource with which to fill spatial gaps in global absolute-wavespeed tomographic models.

We developed an absolute arrival-time recovery method (AARM) to retrieve absolute time picks from relative-arrival-time datasets, working synchronously with filtered and unfiltered data. We also include a relative estimate of uncertainty for potential use in data weighting during subsequent tomographic inversion. Filtered waveforms are first aligned via multichannel cross correlation. These time shifts are applied to unfiltered waveforms to generate a phase-weighted stack. Cross correlation with the primary stack or the SNR of each trace is used to weight a second-higher SNR stack. The first arrival on the final stack is picked manually to recover absolute arrival times for the aligned waveforms.

We test AARM on a recently published dataset from southeast Canada (∼10,000 picks). When compared with the available equivalent earthquake–station pairs on the International Seismological Centre (ISC) database, ∼83% of AARM picks agree to within ±0.5 s. Tests using synthetic P-wave data indicate that AARM produces absolute arrival-time picks to accuracies of better than 0.25 s, akin to uncertainties in ISC bulletins.

Electronic Supplement: Graphical output from testing of the absolute arrival-time recovery method (AARM) on observed dataset and an archive containing a copy of the AARM code, plotting scripts, and user guide.

Introduction

Over the past two decades, the number of temporary regional seismic networks deployed to study upper-mantle structure has grown dramatically (Evans *et al.*, 2015). However, because of suboptimal deployment conditions, teleseismic waveforms of variably emergent nature are generally of lower signal-to-noise ratio (SNR) than those recorded from sparse permanent observatory sites. Fortunately, temporary networks are often small enough in aperture (often <1000 km) to capitalize on waveform coherency (Fig. 1) at teleseismic distances (e.g., VanDecar and Crosson, 1990; Chevrot, 2002; Rawlinson and Kennett, 2004). With the aid of zero-phase filtering, which increases seismogram SNR and preserves the relative timing of the peak amplitudes, relative arrival times can thus be calculated via identification of coherent peaks or troughs of

energy for a given earthquake across a network. Relative-arrival-time inversions (e.g., VanDecar *et al.*, 1995; Rondenay *et al.*, 2000; Rawlinson *et al.*, 2006; Bastow *et al.*, 2008; Frederiksen *et al.*, 2013; Boyce *et al.*, 2016) are thus our primary source of tomographic images of upper-mantle structure beneath many regions of tectonic and geodynamic interest. Despite their high resolution, these models are lacking in one critical respect: they contain no information about the region's average velocity structure, whether it be fast (e.g., the shields) or slow (e.g., hotspots). Within global tomographic models (e.g., Li *et al.*, 2008), these small temporary networks provide a hitherto untapped resource with which to fill gaps in spatial coverage, with important implications for the resolution. The challenge is thus to determine absolute phase-arrival times.

Here, we develop an approach for determining absolute arrival times (the absolute arrival-time recovery method [AARM]) that exploits the efforts of existing relative arrival-time studies. The first arrival can be picked from a phase-weighted stack of previously aligned, yet unfiltered, waveforms, allowing absolute arrival times to be recovered. This is combined with a quantitative pick-quality estimate and is rigorously tested on data containing increasing levels of teleseismic background noise. Temporary network deployments can therefore be used to report accurate absolute arrival-time measurements to global-pick databases, such as the International Seismological Centre (ISC; Di Giacomo *et al.*, 2014; ISC, 2016), and be used in global absolute-wavespeed tomographic inversions.

Relative versus Absolute Arrival Times

Relative arrival times T_{rel} align the first coherent peak or trough (i.e., maximum or minimum) across a network (T_{align}, Fig. 1) by cross correlation (e.g., VanDecar and Crosson, 1990), stacking (e.g., Rawlinson and Kennett, 2004), or otherwise (e.g., Chevrot, 2002). The necessity for waveform coherency during calculation of T_{rel} limits the aperture of a station network to approximately <1500 km. Relative arrival-time residuals RES_{rel} (equation 1) are calculated by comparison to predicted arrivals T_{exp} based, for instance, on travel-time tables (such as IASP91 and ak135: Kennett and Engdahl, 1991; Kennett *et al.*, 1995),

$$RES_{rel} = T_{rel} - (T_{exp} - \bar{T}_{exp}), \qquad (1)$$

in which $\bar{T}_{exp}$ is the mean expected arrival time for an earthquake across the network. During relative-arrival-time analysis, for each earthquake, rays are assumed to follow a similar path until they diverge beneath the regional network.

Resultantly, the contribution from the background mean velocity structure $\bar{T}_{exp}$ is lost because arrival-time variations are assumed to result from local wavespeed structure beneath the network (e.g., Bastow, 2012). However, this background may vary substantially, particularly between cratonic (e.g., Boyce *et al.*, 2016) and active (e.g., Bastow *et al.*, 2008) regions. Thus, relative-arrival-time datasets cannot be combined easily or used to directly recover absolute velocities.

The absolute arrival time (or onset time) T_{abs} marks the start of incoming energy on a recorded waveform (obscured by noise in Fig. 1a–c). Absolute arrival-time residuals RES_{abs} are calculated through direct comparison with the arrival time that is expected for some reference Earth model T_{exp} for a given source–station pair:

$$RES_{abs} = T_{abs} - T_{exp}. \qquad (2)$$

High-SNR waveforms, an accurate earthquake catalog, and often, manual picking by a skilled analyst are required; so, absolute arrival times are mostly limited to sparse permanent observatory sites.
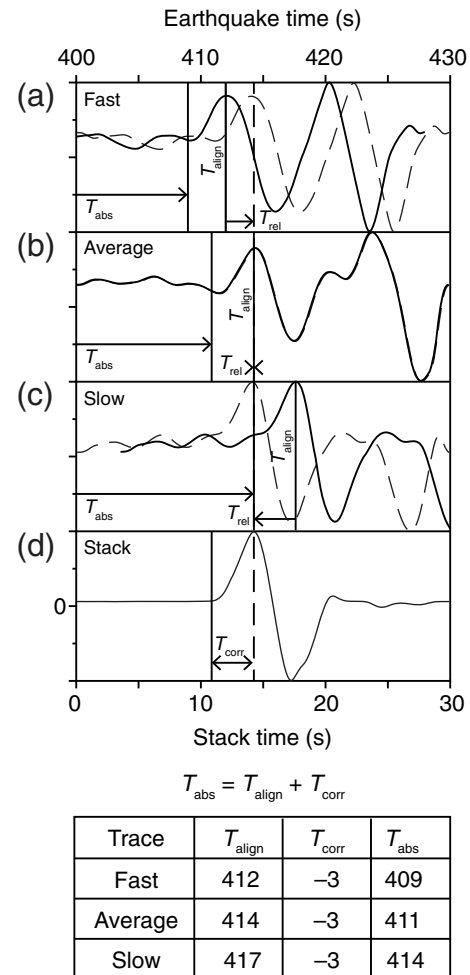

Trace	T_{align}	T_{corr}	T_{abs}
Fast	412	−3	409
Average	414	−3	411
Slow	417	−3	414

Figure 1. Difference between relative (T_{rel}) and absolute (T_{abs}) arrival times. Waveform (a) is fast (2 s) relative to the network average (b) and thus arrives first. The slow waveform (c) arrives later (3 s). Using the first major peak, the waveforms are shifted into alignment (dashed waveforms) with the average trace. This point represents the relative-arrival-time alignment point T_{align}. (d) The first arrival occurs before this (3 s, indicated by the stacked trace), so the correction T_{corr} can be made to find T_{abs}, indicated by the vertical lines.

Conventional Stand-Alone Phase-Picking Methods

Early work in all of seismology was dominated by manual picking. Increasingly large datasets, particularly in controlled-source experiments, motivated the development of automatic-picking routines for static correction calculations (e.g., Cox, 1999). Cross correlations (e.g., Hileman *et al.*, 1968; Taner *et al.*, 1974), increase in signal energy (Coppens, 1985), fractal-based search (e.g., Boschetti *et al.*, 1996), and neural networks (e.g., Dai and MacBeth, 1995) have all been used with varying success. In modern controlled-source seismology applications, these datasets are high frequency with known wavelets and source times, thus typical accuracies are on the order of 10 ms (e.g., Cox, 1999). However, many of these algorithms require training or some manual intervention.
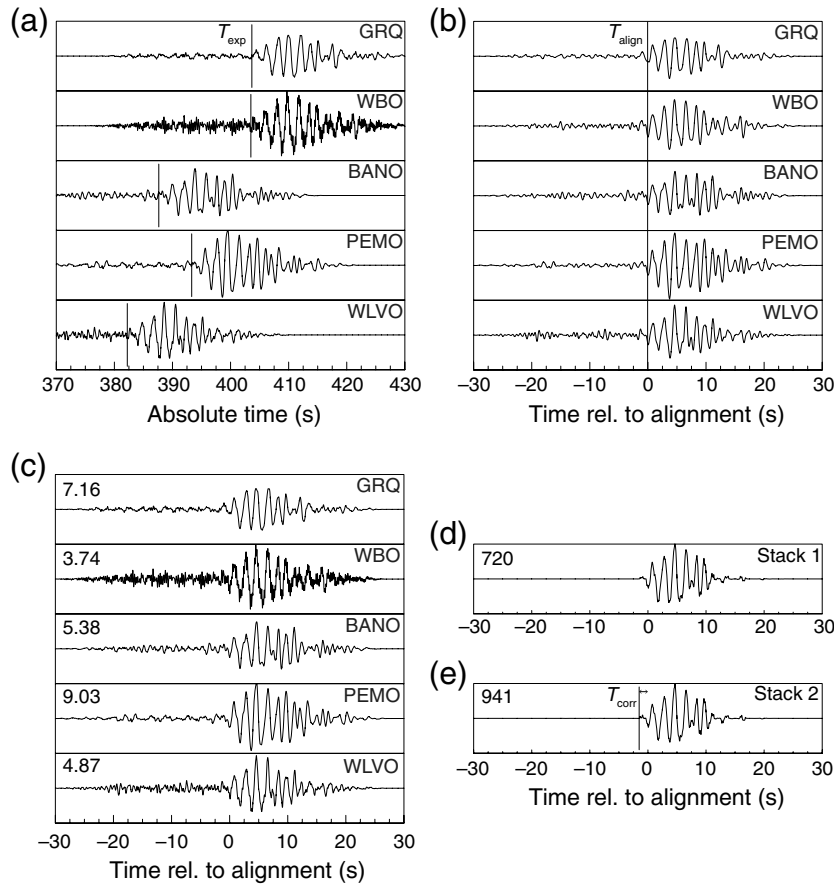
Figure 2. The absolute arrival-time recovery method (AARM) workflow using an example earthquake recorded by a selection of five broadband instruments in southeast (SE) Canada. (a) Unfiltered raw traces with the predicted arrival time T_{exp} for the *P*-wave arrival, shown by the vertical lines. Station names are also shown. (b) Traces are filtered (Butterworth two pole between 0.4 and 2.0 Hz) and aligned on a coherent peak or trough T_{align}, using a relative-arrival-time technique (e.g., VanDecar and Crosson, 1990). (c) Alignments are transferred back to unfiltered velocity seismograms. Numbers within each subpanel refer to the signal-to-noise ratio (SNR) of each trace. (d) An initial phase-weighted stack is calculated; this has an SNR of 720. Weightings for a second stack are calculated using SNR or cross correlation with the initial stack. This results in a higher-SNR second stack (e). Minor trace adjustments can also be made when using the cross-correlation scheme (T_{adj} in equation 7). A manual first arrival picked by the user (vertical line) is then used to calculate the absolute correction T_{corr} to apply to each trace across the network. The absolute arrival times (T_{abs}) are compared with the predicted arrivals (T_{exp}) from ak135 to give absolute arrival-time residuals RES_{abs} (see equation 7).

these methods can lead to large errors in high-noise environments (typical of temporary regional seismic networks), and the algorithms must be carefully trained on a selection of representative reference traces.

During relative-arrival-time analysis, when peaks or troughs of a coherent phase across a network are aligned, an opportunity arises to pick the first arrival from a resulting stack (e.g., Chevrot, 2002; Rawlinson and Kennett, 2004; Pavlis and Vernon, 2010; Lou *et al.*, 2013). However, as far as we have been able to determine, no study to date explicitly lays out a theoretical basis for picking T_{abs} on unfiltered data, nor do they include an associated quantitative estimate of relative uncertainty (for weighting during subsequent tomographic inversion). It is here that we seek improvement.

Absolute Arrival-Time Determination

With the aid of five example seismograms recorded at stations in eastern North America (Fig. 2a), we document our AARM below.

Data Preprocessing

For a temporary network of stations, relative arrival times are routinely calculated (Figs. 1 and 2b; e.g., Frederiksen *et al.*, 2013; Boyce *et al.*, 2016), commonly using the methods of VanDecar and Crosson (1990) or Rawlinson and Kennett (2004). Datasets of filtered waveforms are subject to quality-control (QC) measures that remove cycle skips, timing errors, large outlying residuals, and waveforms of low SNR.

To obtain absolute arrival times from these datasets, the alignment times (T_{align} in Fig. 1) should optimally be transferred back into the unfiltered data (Fig. 2c), with a standardized instrument response (here, the broadband velocity response of a standard Blacknest-type seismometer available in Seismic Analysis Code [SAC], Goldstein and Snoke, 2005; Helffrich *et al.*, 2013). Although minimum-phase filters will not lead to a shift in the first arrival, the signal peaks or troughs may be distorted (e.g., Leonard, 2000; Stefano *et al.*, 2006; Küperkoch *et al.*, 2010). Because we rely on alignment of peaks and troughs in the relative-arrival-time step and in our stacking procedure (described in the next section), the best practice for AARM is to avoid filtering during formation of the stack. Indeed Amaru *et al.* (2008) experience a shift (∼0.1 s)
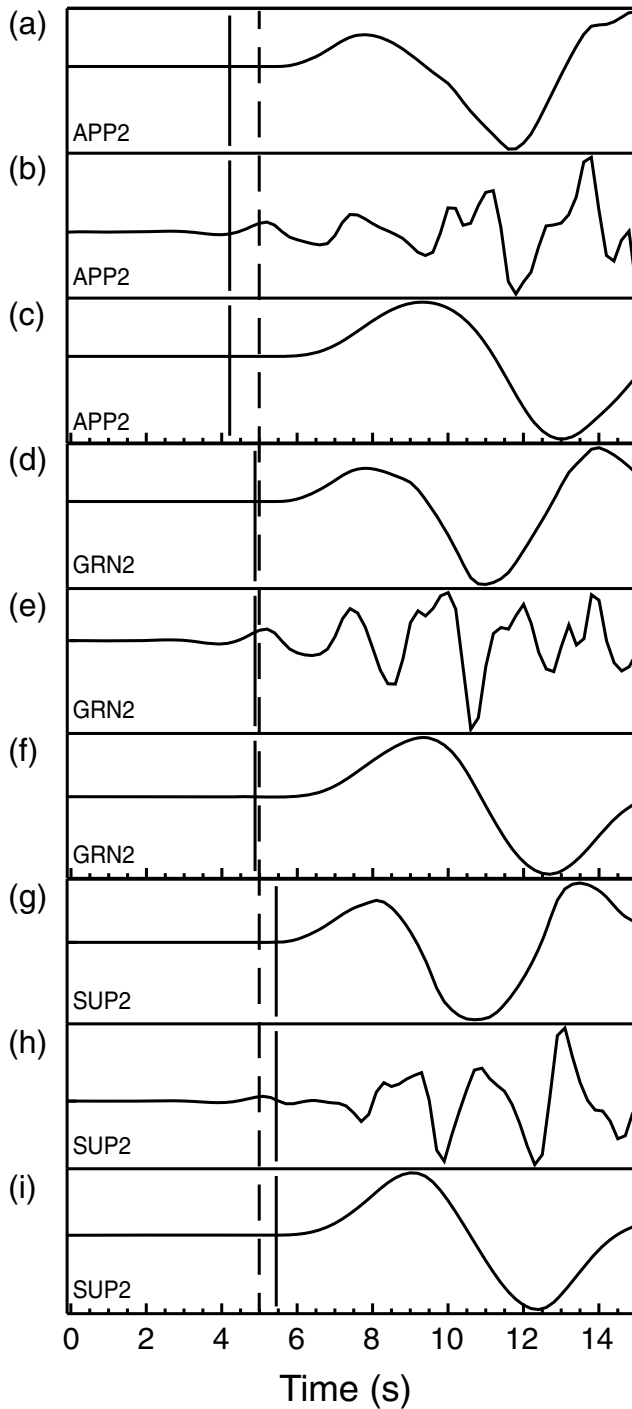
Teleseismic arrival-time determination has a unique set of challenges. Variably emergent waveforms, diverse noise sources (even for one earthquake recorded across a network), lower frequency datasets, and unknown source times and wavelets contribute to a hugely underdetermined problem. Allen (1982), Baer and Kradolfer (1987), and Earle and Shearer (1994) developed the short- and long-term average-ratio method to reliably pick phases trace by trace. Amaru *et al.* (2008) implemented this regime to pick 85,000 absolute arrival times for temporary deployments in Europe, using filtered data. The automated picking algorithm developed by Aldersons (2004) has also been applied to datasets in the Alps (Stefano *et al.*, 2006; Diehl *et al.*, 2009). However,

Figure 3. Aligned, unfiltered, synthetic velocity seismograms (a, d, and g) from three stations (APP2, GRN2, and SUP2). Filtered velocity seismograms (b, e, and h) and unfiltered displacement seismograms (c, f, and i) are also shown. The dashed vertical line refers to the initial trace-alignment point (T_{align}) derived from the multi-channel cross correlation (MCCC) code of VanDecar and Crosson (1990). The solid vertical lines are the predicted absolute-phase-arrival times (T_{exp}) from the 1D velocity model ak135.

toward faster picks, relative to a standard database, due to filtering or varying instrument responses. Figure 3 shows that aligned, filtered velocity seismograms map consistently

into the unfiltered velocity and displacement seismograms, despite not necessarily corresponding to a particular peak or trough.

The unfiltered, aligned data are equalized in sample rate, cut to a predetermined length (60 s for P waves, 120 s for S waves), normalized, and optionally integrated to displacement. We note that, although velocity seismograms are typically more impulsive than displacement seismograms, they often contain more high-frequency noise, but either (velocity or displacement) should be suitable for obtaining absolute arrival times. Because the SNR of unfiltered seismograms can contrast greatly from their filtered counterparts, a low-threshold SNR (< 1) is used to remove particularly noisy traces (e.g., Pavlis and Vernon, 2010). In the following sections, we refer to the SNR frequently, herein defined as

$$\mathrm{SNR} = \frac{A_{\mathrm{signal}}}{A_{\mathrm{noise}}}, \qquad (3)$$

in which A is the root mean squared amplitude defined over a preset window of 25 s (P waves) or 55 s (S waves). The noise and noise-plus-signal windows are separated by a 2 s safety gap akin to Diehl *et al.* (2009) and Stefano *et al.* (2006) around the alignment point (T_{align}). We use P-wave data to describe AARM but note that our testing shows that AARM can be reasonably applied to S-wave data while accounting for their generally longer periods.

Calculation of Arrival Times

Preliminary Stacking

The first step in the methodology is to form an initial stack (Fig. 2d) using the unfiltered aligned waveforms. There are three common stacking methods (e.g., Schimmel and Paulssen, 1997; Rost and Thomas, 2002). A linear stack is defined as

$$f(t)_{\mathrm{lin}} = \frac{1}{N} \sum_{j=1}^{N} s_j(t), \qquad (4)$$

in which $f(t)$ is the stacked trace, N is the number of traces, and $s_j(t)$ is the waveform. An nth-root stack is given by

$$f(t)_{n\mathrm{th}} = \left(\frac{1}{N} \sum_{j=1}^{N} \sqrt[m]{s_j(t)} \right)^m, \qquad (5)$$

in which m is the order of the root. Finally, phase-weight stacking is defined as

$$f(t)_{\mathrm{phase}} = \frac{1}{N} \sum_{j=1}^{N} s_j(t) \times \left(\frac{1}{N} \sum_{j=1}^{N} \exp(i\Phi_k(t)) \right)^v, \qquad (6)$$

in which $\Phi_k(t)$ is the instantaneous phase of the waveform and v defines how readily coherent and incoherent phases are separated, that is, the severity of the phase weighting. A linear stack is retrieved with $v = 0$. The recent dual bootstrap resampling stacking method of Korenaga (2013) is not

considered further because we do not require signal recovery of very-low-SNR (< 1) waveforms.

The three stacking schemes (linear, nth root, and phase weight) were tested systematically on observed data, one example of which is shown in Ⓔ Figure S1 (available in the electronic supplement to this article). The main focus is on strong noise suppression without loss of signal in the final stack. Based on this criterion, linear stacking is not considered further, due to the presence of high-frequency prearrival noise masking the true onset time in the stack Ⓔ Fig. S1a). The remaining stack types, nth root and phase weight (Ⓔ Fig. S1b,c), both achieve a high-SNR-stacked trace. Schimmel and Paulssen (1997) and Rost and Thomas (2002) show that nth-root stacking tends to produce more impulsive peaks that may lead to lower pick errors (Douglas *et al.*, 1997). This is consistent with our test dataset. However, during testing, phase-weight stacking was marginally more effective in suppressing pre-arrival noise (Ⓔ Fig. S1c), leading to higher-SNR stacks. Therefore, aligned traces are stacked (Fig. 2d) using phase weighting ($v = 4$ in equation 6) across the entire window (60 s).

Weighting and Adjustment of Traces for Final Stacking

An SNR or cross-correlation-derived weighting scheme (e.g., Pavlis and Vernon, 2010) can be used to form a higher-SNR second stack (Fig. 2e) on which the first arrival is picked by the analyst. For each trace, an SNR approximation (equation 3) is made (Fig. 2c). For each earthquake, these are normalized between 0 and 1 to give a set of SNR-derived weights for the second stack. Alternatively, cross correlation (XC) of each normalized trace with the preliminary stack can also yield a set of XC-derived weights for the second stack, again normalized between 0 and 1. This method also allows small adjustments to the initial alignment points if the cross-correlation function maximum (between the trace and the preliminary stack) is offset from zero. In theory, this should result in lower picking uncertainty.

In cases in which traces are very similar across the network, the cross-correlation technique will tend to assigning relatively high weights to all traces (Ⓔ Fig. S2c). However, the SNR-weighting scheme may produce a broader spread of weightings, and thus the stack will be more robust (Ⓔ Fig. S3c). In a relatively high-noise environment, the stack may be dominated by one quiet trace when using the SNR weighting, but the weights derived from cross correlation may be more uniformly distributed. In practice, the choice of SNR or XC-derived weightings varies from one application to the next. We did not find a strong pattern of improved stacks when using the XC regime on emergent events and the SNR regime on impulsive events (e.g., Pavlis and Vernon, 2010).

Our testing showed that linearly distributing cross correlation or SNR weightings between 0 and 1 is adequate for our purpose. We note that weighting functions have been used to improve signal stacking in some applications (e.g., Restivo

and Helffrich, 1999), but testing of a variety of weighting functions on both observed and synthetic datasets did not show any significant improvement in absolute arrival-time picks. For the majority of earthquakes, for which the number of high-SNR waveforms is > 10, the second stack may appear to offer little improvement in pick accuracy (e.g., Fig. 2d,e). However, when only a limited number (< 10) of variably noisy waveforms are available, the arrival time on the initial stack can vary by > 0.3 s when compared to the final stack (Ⓔ Fig. S4a,b). We, therefore, choose to compute the second stack in all cases at a minor computational cost.

First Arrival Picking

The linear weighting process leads to an improved second phase-weighted stack (equation 6), on which the first arrival or onset time is picked manually by an analyst (Fig. 2e). We use a consistent window size and axes scaling for the entire dataset while picking the first arrival (e.g., Douglas *et al.*, 1997; Diehl *et al.*, 2009). Manual picking does introduce some human bias into the data (e.g., Leonard, 2000) but is more randomly distributed than automatic-picking methods (e.g., Aldersons, 2004; Amaru *et al.*, 2008). Furthermore, the variably emergent nature of unfiltered, stacked signals, necessary training of algorithms (for pick accuracy), and generally low numbers of earthquakes within a relative arrival-time dataset (often ~200) dictate that automatic-picking methods would provide no logical benefit here.

For each earthquake, enough stations across the network must record at high-enough SNR to form a high-SNR stack (Fig. 2e), in which random pre-arrival noise is sufficiently suppressed. This is required to accurately pick the stack manually.

Computation of Absolute Arrival Times and Residuals

Following manual picking of the final stack (Fig. 2e), similarly to equation (2), absolute arrival-time residuals ($\mathrm{RES}_{\mathrm{abs}}[i]$) can be calculated with the following expression:

$$\mathrm{RES}_{\mathrm{abs}}[i] = (T_{\mathrm{align}}[i] + T_{\mathrm{adj}}[i] + T_{\mathrm{corr}}) - T_{\mathrm{exp}}[i]. \qquad (7)$$

The difference between the final alignment point in the stack and the user-picked onset time (negative when the onset occurs before the final alignment point, see Fig. 1) gives the time correction (T_{corr}) to be applied to each trace. This value is consistent across all traces for a teleseismic earthquake in which the waveform does not change shape over a network (an assumption of prior relative-arrival-time analysis). The absolute arrival times T_{abs} (i.e., the term in parentheses in equation 7) are easily calculated by adding the correction to each of the alignment times ($T_{\mathrm{align}}[i]$). This may also be adjusted using the cross-correlation correction ($T_{\mathrm{adj}}[i]$). These times are then compared with a predicted time for each ray ($T_{\mathrm{exp}}[i]$), ak135 in this case, producing absolute arrival-time residuals $\mathrm{RES}_{\mathrm{abs}}[i]$. Upon processing of each event, visual inspection of residual distribution enables fast identi-

fication of outliers. We also use our pick-quality-assessment parameters discussed below to remove further picking inconsistencies.

Methodological Assessment of Pick Quality

A visual first-order indication of picking error is given by the broadness of the final stacked trace (Rawlinson and Kennett, 2004) controlled by the impulsivity of the unknown source wavelet. Unlike controlled-source experiments in which the wavelet can be tuned for the purpose, earthquakes exhibit variability of signal onsets, from impulsive to emergent (Ⓔ Figs. S2 and S3). In extreme cases, less-impulsive first arrivals from long rise-time earthquakes can give onset timing errors of up to 0.5 s (Douglas *et al.*, 1997).

Chevrot (2002) proposed a quantitative method for evaluating this relative picking error. First, the autocorrelation of the final stack is calculated. Then, each trace is individually cross correlated with the final stack to give a cross-correlation function. As the autocorrelation of the stack decreases from 1 from its center (in a normalized case), it will pass through the absolute maximum value observed on the cross-correlation function of the stack and the individual trace. This point will be offset from zero in time and thus gives an approximate measure of relative picking error. We set an upper bound on this autocorrelation-derived relative-error estimate of 0.25 s, above which we exclude the picks from the output absolute arrival times. This ensures that the individual relative-pick uncertainty is lower than the absolute errors of global pick databases.

The cross correlation of the primary stack with each trace can be used to adjust the alignment of traces within the final stack, if the maximum ($T_{adj}[i]$) is offset from zero (e.g., Amaru *et al.*, 2008; Pavlis and Vernon, 2010). This indicates that the trace is better aligned with the stack by a small time shift. When using the cross-correlation-derived weighting scheme, it is possible to correct for this; therefore individual traces are better aligned for the second stack, giving more-accurate arrival-time picks. However, in some cases after the second stack, the value of the maximum offset can be nonzero (more often using the SNR weighting scheme, which does not make this correction). Because our stacked traces are well aligned initially by the relative-arrival-time analysis, we find this measure to be on the order of the sample interval for $>95\%$ of traces in our test datasets. Thus, we use this as a tool to remove poorly aligned traces from our absolute arrival-time dataset (e.g., Pavlis and Vernon, 2010) rather than to directly assess pick quality.

As described above, the cross-correlation function maximum between the primary stack and each trace can be used to weight a second stack. The distribution of these weights is a useful indicator of sources of error, although this is not a direct measure in the time domain. For earthquakes in which a high number of traces have weights > 0.6, the stack will be a reliable approximation to the ideal trace for the network (e.g., Ⓔ Figs. S2 and S3) or array beam. Conversely, in cases where $<10\%$ of traces have weights >0.6, the majority of traces are down-weighted heavily in the final phase-weighted stack. The final stack is not necessarily a reliable estimate for the array beam, and thus the first arrival on the stack may not be well constrained. This does provide a strategy for weighting of individual earthquakes within a subsequent tomographic inversion, however. We note that traces with low weightings in the second stack do not contribute significantly, but arrival-time picks can still be produced or removed as necessary.

Testing of Observed and Synthetic Datasets

We test AARM using both synthetic and observed datasets. All synthetic and observed input data are normalized to the broadband velocity response of a standard Blacknest-type seismometer (available in SAC, Goldstein and Snoke, 2005; Helffrich *et al.*, 2013). Deconvolution with the individual seismometer response directly to displacement is ill-advised, due to the imposed filtering and therefore possible distortion of the onset time (e.g., Leonard, 2000; Stefano *et al.*, 2006; Küperkoch *et al.*, 2010). We have also tested the picking regime on both displacement and velocity seismograms. We find that high-frequency noise suppression during stacking is strong enough to use velocity seismograms to pick absolute arrival times for our test datasets. Despite the lack of direct absolute-error estimates derived from manual picking, we nevertheless explore quantitatively how well absolute arrival times are recovered by AARM.

Absolute Arrival-Time Recovery for Observed Datasets

AARM is first tested on a published, observed, relative arrival-time dataset (Fig. 4) of ∼10,000 picks from southeast (SE) Canada (Boyce *et al.*, 2016) and compared with the absolute arrival-time database of the ISC (Di Giacomo *et al.*, 2014; ISC, 2016). Because of strict QC measures imposed (e.g., SNR <1 and relative-pick accuracy <0.25 s), we recover 9053 ($>90\%$) absolute arrival-time measurements (Fig. 4a) for subsequent use in a global tomographic inversion (such as Li *et al.*, 2008). Almost all (99.7%) residuals are distributed between ± 3 s, with a mean of -0.44 s. This is expected for a region of generally elevated wavespeed relative to the global average (e.g., Li *et al.*, 2008).

The ISC (Di Giacomo *et al.*, 2014; ISC, 2016) collates travel-time data from a number of sources and contributing agencies, in which picks can be manual or automatic. However, it does not routinely provide associated pick errors, let alone any method of calculating them. The oldest records available on the ISC (1964 onward) are thought to be accurate to only ± 2 s (Kennett and Engdahl, 1991), whereas the present average-origin-time uncertainties are ∼1 s for the ISC catalog (Kagan, 2003). Random picking errors are also thought to be high (∼0.5 s; e.g., Gudmundsson *et al.*, 1990). The regional hypocenter determination work of Husen *et al.*
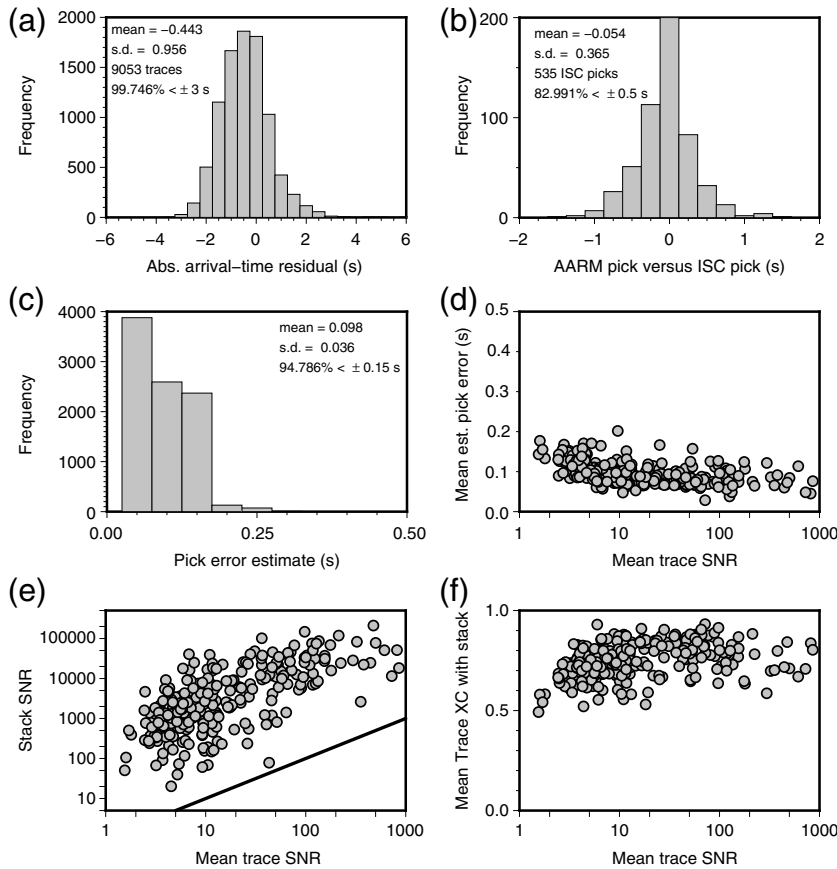
Figure 4. Dataset overview for the methodology (AARM) applied to the relative arrival-time picks of Boyce *et al.* (2016). (a) The distribution of absolute arrival-time residuals for the dataset. (b) The difference between the calculated absolute arrival times (from AARM) and the International Seismological Centre (ISC) picks, where available. (c) The autocorrelation pick-error estimate for the entire dataset. (d) The variation of mean autocorrelation pick errors with event-averaged SNR. (e) The variation in SNR of the final stacked trace against event-averaged SNR. The black line shows the line $y = x$. (f) The mean cross-correlation coefficient of each trace with the stack as a function of event-averaged SNR.

(1999) indicates that modern techniques are accurate to ~2 km in source depth. This translates to a timing uncertainty of ~0.3 s for a typical *P*-wave arrival and ~0.55 s for *S* waves.

Within our observed dataset, 535 equivalent earthquake–station pairs are available on the ISC-pick database (Fig. 4b). AARM recovers 83% of absolute arrival-time picks within ±0.5 s of the ISC picks (approximately normally distributed). ISC picks are, on average, slower by 0.05 s in agreement with Amaru *et al.* (2008). The autocorrelation-derived measure of relative error (Fig. 4c) indicates that picks are accurate to 0.1 s, on average, with 94.8% of picks ≤0.15 s. As expected, as the mean trace SNR increases, mean relative-pick error decreases (consistently low above a mean SNR of 2, Fig. 4d), and the SNR of the stack increases (Fig. 4e). However, high-SNR stacks can result from generally lower SNR data. This is likely due to a small proportion of noisy traces that lower the average SNR but are down-weighted heavily in the final stack. Additionally,

higher SNR results in marginally improved cross-correlation coefficients with the final stack (Fig. 4f).

The comparison to the ISC database and relative assessments of error (that provide a weighting for arrivals during tomographic inversion) show that AARM works reliably for our test dataset. However, using waveforms of high SNR, it is possible to go one step further and compare manual reference picks with the picks derived from AARM (e.g., Stefano *et al.*, 2006; Amaru *et al.*, 2008; Küperkoch *et al.*, 2010) to obtain an estimate for absolute error. The temporary network deployments, on which AARM is designed to be used, will not allow for manual reference picking in most cases, due to high background noise levels. Instead, we are able to directly compare picks produced by AARM and manual picks on the noise-free synthetic seismograms described below.

Absolute Arrival-Time Recovery for Synthetic Datasets

For our synthetic testing, we model a linear network of nine stations evenly spaced between epicentral distances of 3400 and 4200 km (Fig. 5), simulating a teleseismic earthquake from the western United States recorded in eastern Canada (e.g., Boyce *et al.*, 2016). The velocity model (Fig. 5) is perturbed systematically by ±2% to replicate a fast-wavespeed Archean Superior Province (SUP), average-velocity Grenville Province (GRN), and slow-wavespeed Appalachians (APP) (see Boyce *et al.*, 2016) through which synthetic *P*-wave seismograms (dip-slip source time function: Figs. 6a and 7a) are generated using the Computer Programs for Seismology package (Herrmann, 2013). The equivalent procedure is followed to generate the *S*-wave synthetics in Ⓔ Figures S5a and S6a, although a strike-slip source-time function is used to increase testing heterogeneity. These waveforms are initially aligned (Ⓔ Fig. S7) using the relative-arrival-time method of VanDecar and Crosson (1990). We note that the large maximum inter-station distance results in seismograms on the limit of necessary waveform coherence (Fig. 7a and Ⓔ Fig. S6a) for a relative-arrival-time study; thus, in practice, errors are likely to be lower than those observed for our synthetic datasets. Using the teleseismic noise spectra from the study of Peterson (1993), the synthetic seismograms are subject to increasing levels of random noise distributed between the new low-noise model and new high-noise model power spectra (colored curves of Ⓔ Fig. S8). Example outputs for a synthetic station
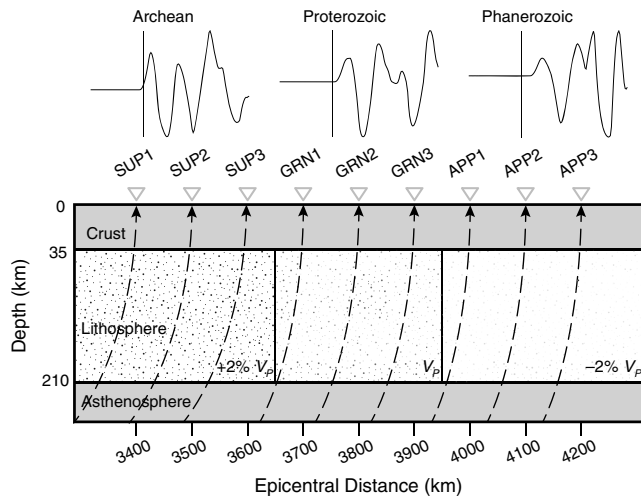
Figure 5. Perturbed 2D ak135 velocity model (+2%, 0%, and −2%dV_P), between 35 and 210 km depth, used to produce symmetrically distributed synthetic waveforms, modeling early, average, and late arrivals (example delayed waveforms are shown above) at increasing epicentral distance (3400–4200 km). This corresponds to the Archean Superior Province (SUP, dense fill), Proterozoic Grenville Province (GRN, average fill), and Phanerozoic Appalachians (APP, light fill) in southeastern Canada, akin to the study of Boyce *et al.* (2016). Ray tracing is used to produce Green's functions that are convolved with a parabolic source-time function representative of a 90° dip-slip earthquake placed at 210 km depth (Herrmann, 2013). Output seismograms are convolved with a standardized broadband velocity response of a standard Blacknest-type seismometer (available in Seismic Analysis Code [SAC], Goldstein and Snoke, 2005; Helffrich *et al.*, 2013).

(SUP2) are given in Figure 6a and Ⓔ Figure S5a for *P* and *S* waves, respectively. The increasing uncorrelated teleseismic noise spectra (0.0–1.0; Ⓔ Fig. S8) give rise to decreased SNR of each trace (Fig. 6a and Ⓔ Fig. S5a).

AARM picks produced for all nine zero-noise synthetic stations are within ±0.25 s of the manual reference pick (Fig. 7f). Thus, we consider AARM to be able to produce *P*-wave picks to accuracies of <0.25 s. This is consistent with the ISC database and other accurate databases worldwide (Leonard, 2000) in which 90% of picks made by an experienced analyst are typically accurate to within ±0.2 s. AARM picks preferentially occur slightly after the manual reference picks (on average less than the sampling interval), in agreement with Amaru *et al.* (2008) and Küperkoch *et al.* (2010).

When compared with the zero-noise case, synthetic seismograms (with average SNR > 8) produce very similar results to the zero-noise case in terms of absolute arrival-time residuals (Fig. 7c,d), estimated errors (Fig. 7b,e), and the manual pick time on the stack (Fig. 6b). Figure 6a shows that, between a trace SNR of 8 and 2, the resulting stack undergoes mild distortion of the first arriving peak (Fig. 6b), and thus manual pick-time errors result. Below this average-trace-SNR level, the stack completely breaks down, containing significant high-frequency noise, and does not show any resemblance to the zero-noise stack (Fig. 6b). However, the previously described QC steps ensure that such low-SNR stacks, with little representation of the average trace for the network, rarely occur.

For the *S*-wave synthetic dataset, the change in source-time function does not affect the accuracy of the picks. The stack breaks down at a lower noise level (Ⓔ Fig. S5b) because the noise spectrum has a greater influence on *S*-wave frequencies than that of the *P*-wave data, but this is to be expected. The absolute errors, compared to individual manual picks, are in agreement to within ±0.5 s (Ⓔ Fig. S6f), again a slight increase over the *P*-wave tests.

Conclusions

We developed a strategy (AARM) to determine absolute arrival times using routinely processed, teleseismic, relative arrival-time datasets from regional networks. A manual pick of the first arrival of stacked prealigned unfiltered traces allows absolute arrival times to be calculated for each station record.

A database of ~10,000 picks from networks in SE Canada is used to test AARM. When checked against the ISC archive, the difference in absolute arrival times shows that 83% are in agreement to within ±0.5 s. The tests also indicate that AARM is effective for a low-average-trace SNR (>2). Further testing of synthetic data indicates that AARM is accurate to better than 0.25 s in an absolute sense, on par with modern worldwide databases. AARM can thus be used to incorporate data from dense
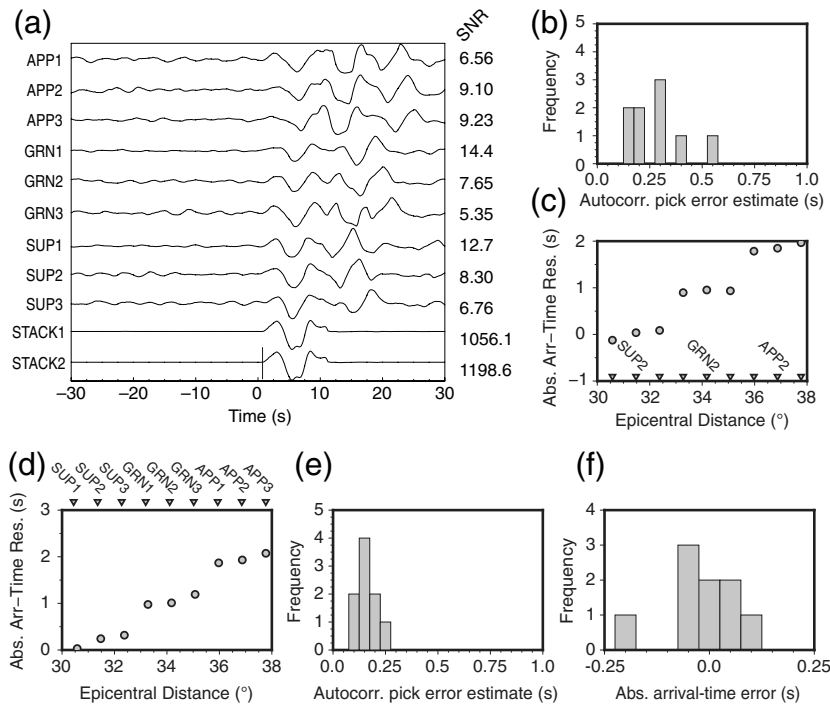

Figure 6. Plot to show how increasing levels of teleseismic background noise (Peterson, 1993) affect (a) the input velocity synthetic seismograms and (b) the output stacked traces. (a) Increasing levels of teleseismic noise are added to the synthetic velocity seismogram SUP2, resulting in decreased SNR as shown. (b) The increasing levels of noise on the input traces produce velocity seismogram stacks of decreasing SNR and thus inaccurate picks for the onset time. Vertical lines are the manual picks used for absolute arrival-time calculation in AARM.

Figure 7. Results from alignment of noisy synthetic traces compared to key results from a zero-noise case. (a) Nine velocity synthetic seismograms convolved with the 0.1-level power spectra of teleseismic noise (Peterson, 1993). The approximate SNR is shown beside each trace. Traces are aligned and stacked using phase weighting (STACK1), then restacked using cross correlation (used here) or SNR-derived weights to form the higher-SNR final stack (STACK2). The manually picked onset or first arrival is shown as a vertical line. (b) The autocorrelation pick-error estimate for the 0.1-noise-level synthetic traces in (a). (c) Distribution of absolute arrival-time residuals for the 0.1-noise-level synthetic traces in (a). (d) Plot showing absolute arrival-time residuals for the idealized zero-noise case for comparison with 0.1 noise level in (a–c). (e) Autocorrelation-derived pick-error estimates for the zero-noise case. (f) The zero-noise case allows calculated absolute arrival-time residuals (from AARM) to be compared with manually picked arrivals. The plot shows a histogram of the difference between two measurements on which negative values indicate the AARM pick to be after the manual hand pick.

but short-duration temporary seismograph networks into global-pick databases and absolute-wavespeed tomographic inversions.

Data and Resources

A copy of the code (absolute arrival-time recovery method [AARM]) used to produce absolute arrival times from the relative-arrival-time dataset of Boyce *et al.* (2016) is available in the Ⓔ electronic supplement to this article or by contacting the corresponding author. The authors thank J. VanDecar for use of his multichannel cross correlation (MCCC) codes and also Herrmann (2013) for use of the computer programs in seismology package. Seismic Analysis Code (SAC; Goldstein and Snoke, 2005; Helffrich *et al.*, 2013) and Generic Mapping Tool (GMT; Wessel and Smith, 1995) software were also used to process seismic data obtained from the Incorporated Research Institutions for Seismology (IRIS) Data Management Center and from the Canadian National Data Centre (Natural Resources Canada) (http://www.earthquakescanada.nrcan.gc.ca/stndon/

AutoDRM/autodrm_req?en.php, last accessed December 2016).

References

Aldersons, F. (2004). Toward a three-dimensional crustal structure of the Dead Sea region from local earthquake tomography, *Ph.D. Thesis*, Tel Aviv University, Tel Aviv, Israel.

Allen, R. (1982). Automatic phase pickers: Their present use and future prospects, *Bull. Seismol. Soc. Am.* **72,** no. 6B, S225–S242.

Amaru, M., W. Spakman, A. Villaseñor, S. Sandoval, and E. Kissling (2008). A new absolute arrival time data set for Europe, *Geophys. J. Int.* **173,** no. 2, 465–472, doi: 10.1111/j.1365-246X.2008.03704.x.

Baer, M., and U. Kradolfer (1987). An automatic phase picker for local and teleseismic events, *Bull. Seismol. Soc. Am.* **77,** no. 4, 1437–1445.

Bastow, I. (2012). Relative arrival-time upper-mantle tomography and the elusive background mean, *Geophys. J. Int.* **190,** no. 2, 1271–1278, doi: 10.1111/j.1365-246X.2012.05559.x.

Bastow, I., A. Nyblade, G. Stuart, T. Rooney, and M. Benoit (2008). Upper mantle seismic structure beneath the Ethiopian hotspot: Rifting at the edge of the African low velocity anomaly, *Geochem. Geophys. Geosys.* **9,** no. 12, Q12022, doi: 10.1029/2008GC002107.

Boschetti, F., M. D. Dentith, and R. D. List (1996). A fractal-based algorithm for detecting first arrivals on seismic traces, *Geophysics* **61,** no. 4, 1095–1102, doi: 10.1190/1.1444030.

Boyce, A., I. D. Bastow, F. A. Darbyshire, A. G. Ellwood, A. Gilligan, V. Levin, and W. Menke (2016). Subduction beneath Laurentia modified the eastern North American cratonic edge: Evidence from P wave and S wave tomography, *J. Geophys. Res.* **121,** no. 7, 5013–5030, doi: 10.1002/2016JB012838.

Chevrot, S. (2002). Optimal measurement of relative and absolute delay times by simulated annealing, *Geophys. J. Int.* **151,** no. 1, 164–171, doi: 10.1046/j.1365-246X.2002.01755.x.

Coppens, F. (1985). First arrival picking on common offset trace collections for automatic estimation of static corrections, *Geophys. Prospect.* **33,** no. 8, 1212–1231, doi: 10.1111/j.1365-2478.1985.tb01360.x.

Cox, M. (1999). *Static Corrections for Seismic Reflection Surveys*, Society of Exploration Geophysicists, Tulsa, Oklahoma.

Dai, H., and C. MacBeth (1995). Automatic picking of seismic arrivals in local earthquake data using an artificial neural network, *Geophys. J. Int.* **120,** no. 3, 758–774, doi: 10.1111/j.1365-246X.1995.tb01851.x.

Di Giacomo, D., D. A. Storchak, N. Safronova, P. Ozgo, J. Harris, R. Verney, and I. Bondár (2014). A new ISC service: The bibliography of seismic events, *Seismol. Res. Lett.* **85,** no. 2, 354–360, doi: 10.1785/0220130143.

Diehl, T., E. Kissling, S. Husen, and F. Aldersons (2009). Consistent phase picking for regional tomography models: Application to the greater Alpine region, *Geophys. J. Int.* **176,** no. 2, 542–554, doi: 10.1111/j.1365-246X.2008.03985.x.

Douglas, A., D. Bowers, and J. Young (1997). On the onset of *P* seismograms, *Geophys. J. Int.* **129,** no. 3, 681–690, doi: 10.1111/j.1365-246X.1997.tb04503.x.

Earle, P., and P. Shearer (1994). Characterization of global seismograms using an automatic-picking algorithm, *Bull. Seismol. Soc. Am.* **82,** no. 2, 366–376.

Evans, P., A. Strollo, A. Clark, T. Ahern, R. Newman, J. Clinton, H. Pedersen, and C. Pequegnat (2015). Why seismic networks need digital object identifiers, *Eos Trans. AGU* **96,** doi: 10.1029/2015EO036971.

Frederiksen, A., T. Bollmann, F. Darbyshire, and S. van der Lee (2013). Modification of continental lithosphere by tectonic processes: A tomographic image of central North America, *J. Geophys. Res.* **118,** no. 3, 1051–1066, doi: 10.1002/jgrb.50060.

Goldstein, P., and A. Snoke (2005). SAC availability for the IRIS community, *DMS Electr. Newsl.* **7,** no. 1.

Gudmundsson, O., J. Davies, and R. Clayton (1990). Stochastic analysis of global traveltime data: Mantle heterogeneity and random errors in the ISC data, *Geophys. J. Int.* **102,** no. 1, 25–43, doi: 10.1111/j.1365-246X.1990.tb00528.x.

Helffrich, G., J. Wookey, and I. Bastow (2013). *The Seismic Analysis Code: A Primer and User's Guide*, First Ed., Cambridge University Press, New York, New York.

Herrmann, R. B. (2013). Computer programs in seismology: An evolving tool for instruction and research, *Seismol. Res. Lett.* **84,** no. 6, 1081–1088, doi: 10.1785/0220110096.

Hileman, J., P. Embree, and J. Pflueger (1968). Automated static corrections, *Geophys. Prospect.* **16,** 326–358.

Husen, S., E. Kissling, E. Flueh, and G. Asch (1999). Accurate hypocentre determination in the seismogenic zone of the subducting Nazca plate in northern Chile using a combined on-/offshore network, *Geophys. J. Int.* **138,** no. 3, 687–701, doi: 10.1046/j.1365-246x.1999.00893.x.

International Seismological Centre (ISC) (2016). *On-line Bulletin*, International Seismological Centre, Thatcham, United Kingdom, available at http://www.isc.ac.uk (last accessed August 2016).

Kagan, Y. Y. (2003). Accuracy of modern global earthquake catalogs, *Phys. Earth Planet. In.* **135,** nos. 2/3, 173–209, doi: 10.1016/S0031-9201(02)00214-5.

Kennett, B., and E. Engdahl (1991). Traveltimes from global earthquake location and phase identification, *Geophys. J. Int.* **105,** no. 2, 429–465, doi: 10.1111/j.1365-246X.1991.tb06724.x.

Kennett, B. L. N., E. R. Engdahl, and R. Buland (1995). Constraints on seismic velocities in the Earth from traveltimes, *Geophys. J. Int.* **122,** no. 1, 108–124, doi: 10.1111/j.1365-246X.1995.tb03540.x.

Korenaga, J. (2013). Stacking with dual bootstrap resampling, *Geophys. J. Int.* **195,** no. 3, 2023–2036, doi: 10.1093/gji/ggt373.

Küperkoch, L., T. Meier, J. Lee, and W. Friederich (2010). Automated determination of *P*-phase arrival times at regional and local distances using higher order statistics, *Geophys. J. Int.* **181,** no. 2, 1159–1170, doi: 10.1111/j.1365-246X.2010.04570.x.

Leonard, M. (2000). Comparison of manual and automatic onset time picking, *Bull. Seismol. Soc. Am.* **90,** no. 6, 1384–1390, doi: 10.1785/0120000026.

Li, C., R. Van der Hilst, R. Engdahl, and S. Burdick (2008). A new global model for *P* wave speed variations in Earth's mantle, *Geochem. Geophys. Geosys.* **9,** no. 5, Q05018, doi: 10.1029/2007GC001806.

Lou, X., S. van der Lee, and S. Lloyd (2013). AIMBAT: A Python/Matplotlib tool for measuring teleseismic arrival times, *Seismol. Res. Lett.* **84,** no. 1, 85–93, doi: 10.1785/0220120033.

Pavlis, G. L., and F. L. Vernon (2010). Array processing of teleseismic body waves with the USArray, *Comput. Geosci.* **36,** no. 7, 910–920, doi: 10.1016/j.cageo.2009.10.008.

Peterson, J. (1993). Observations and modeling of seismic background noise, *U.S. Geol. Surv. Open-File Rept. 93-322.*

Rawlinson, N., and B. Kennett (2004). Rapid estimation of relative and absolute delay times across a network by adaptive stacking, *Geophys. J. Int.* **157,** no. 1, 332–340, doi: 10.1111/j.1365-246X.2004.02188.x.

Rawlinson, N., A. M. Reading, and B. L. N. Kennett (2006). Lithospheric structure of Tasmania from a novel form of teleseismic tomography, *J. Geophys. Res.* **111,** no. B2, 1–21, doi: 10.1029/2005JB003803.

Restivo, A., and G. Helffrich (1999). Teleseismic shear wave splitting measurements in noisy environments, *Geophys. J. Int.* **137,** no. 3, 821–830, doi: 10.1046/j.1365-246x.1999.00845.x.

Rondenay, S., M. G. Bostock, T. M. Hearn, D. J. White, and R. M. Ellis (2000). Lithospheric assembly and modification of the SE Canadian shield: Abitibi-Grenville teleseismic experiment, *J. Geophys. Res.* **105,** no. B6, 13,735–13,754, doi: 10.1029/2000JB900022.

Rost, S., and C. Thomas (2002). Array seismology: Methods and applications, *Rev. Geophys.* **40,** no. 3, 2-1–2-27, doi: 10.1029/2000RG000100.

Schimmel, M., and H. Paulssen (1997). Noise reduction and detection of weak, coherent signals through phase-weighted stacks, *Geophys. J. Int.* **130,** no. 2, 497–505, doi: 10.1111/j.1365-246X.1997.tb05664.x.

Stefano, D. R., F. Aldersons, and E. Kissling (2006). Automatic seismic phase picking and consistent observation error assessment: Application to the Italian seismicity, *Geophys. J. Int.* **165,** 121–134, doi: 10.1111/j.1365-246X.2005.02799.x.

Taner, M., F. Koehler, and K. Alhilali (1974). Estimation and correction of near-surface time anomalies, *Geophysics* **39,** no. 4, 441–463, doi: 10.1190/1.1440441.

VanDecar, J., and R. Crosson (1990). Determination of teleseismic relative phase arrival times using multi-channel cross-correlation and least squares, *Bull. Seismol. Soc. Am.* **80,** no. 1, 150–169.

VanDecar, J., D. James, and M. Assumpção (1995). Seismic evidence for a fossil mantle plume beneath South America and implications for plate driving forces, *Nature* **378,** no. 6552, 25–31, doi: 10.1038/378025a0.

Wessel, P., and W. Smith (1995). New version of the Generic Mapping Tools released, *Eos Trans. AGU* **76,** no. 33, 329, doi: 10.1029/95EO00198.

Department of Earth Science and Engineering
Royal School of Mines
Imperial College London
Prince Consort Road
London SW7 2BP, United Kingdom
alistair.boyce10@imperial.ac.uk
i.bastow@imperial.ac.uk
    (A.B., I.D.B.)

Department of Earth Science
University of Bergen
Allegaten 41
5007 Bergen, Norway
rondenay@geo.uib.no
    (S.R.)

Department of Earth, Atmospheric and Planetary Sciences
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139
hilst@mit.edu
    (R.D.V.)