# An efficient algorithm to measure arrival times of weak seismic phases

Lei Li [ORCID], Stéphanie Durand, Yanick Ricard [ORCID] and Eric Debayle

*Laboratoire de Géologie de Lyon: Terre, Planètes et Environnement, ENSL, Univ Lyon 1, CNRS, 69100, Villeurbanne, France. E-mail: ll.ynyf@gmail.com*

## SUMMARY

In seismic tomography, traveltime information of seismic body phases is commonly used to invert the seismic velocities of the subsurface structure. At long periods or for later seismic phases, the arrival time of seismic phases lack definitive onset and a direct picking of the absolute arrival time has large uncertainty and reproducibility. A common practice is to estimate the relative delay between the observed and synthetic signals that maximizes the correlation coefficient. For that aim, we must first select appropriate time windows around the candidate signals. To improve the ability to detect and extract weak signals, we develop a new morphological time window selection (MTWS) algorithm that adapts to the shape of signals and has robust performance in automated processing of massive data. The MTWS method consists of two successive steps. First, we detect the major peaks on the waveform envelope using a maximum filter. Secondly, we solve for the beginning and end of the time windows surrounding the peaks straightforwardly from simple geometrical equations. The efficiency and robustness of the MTWS algorithm make it very suitable for automated processing of huge data sets. We demonstrate the implementation of the method with both synthetic and observed long period (20–40 s) SH waves. From ~100 000 traces of transverse-component seismograms recorded by global seismic networks over the course of a year, we obtain ~15 000 $S_{\mathrm{diff}}$, ~7500 $ScS$ and also some $ScS$ multiples. The global map of $S_{\mathrm{diff}}$ correlation time delays shows consistent patterns with the shear wave velocity perturbations on the core–mantle boundary in the recent tomographic models.

**Key words:** Statistical methods; Time-series analysis; Body waves; Seismic tomography.

## 1 INTRODUCTION

Seismic tomography is a primary tool to image the deep, inaccessible structure of the Earth. High-fidelity 3-D seismic imaging, especially details in regions where seismic speeds are anomalous, is critical to better constrain the physical and chemical properties of materials in the deep Earth, and to understand plate tectonics and mantle dynamics (e.g. Debayle *et al.* 2020; Ritsema & Lekić 2020; Chen 2021). The increase in seismological observations since the pioneering models of the 1980s, provides seismologists with the opportunity to improve the details of the imaging of the Earth's interior (e.g. Bozdağ *et al.* 2016; Durand *et al.* 2017; Lu *et al.* 2019; Hosseini *et al.* 2020; Lei *et al.* 2020). Meanwhile, the explosive growth in the data volume (the IRIS-DMC archives increased by two orders of magnitude in 20 yr) poses urgent requirements for the development of automated and efficient computational processing techniques. The manual or hybrid manual/cross-correlation methods to pick phase arrivals become no longer viable with large numbers of stations Jand earthquakes (Houser *et al.* 2008).

The automation of data processing is not a new topic in seismology. The most common cases are the automatic detection of seismic events, identification of seismic phases and picking of arrival times (e.g. Cansi 1995; Withers *et al.* 1998). The automatic phase picking can provide backbones to many important seismic applications, such as real-time event location (e.g. Baer & Kradolfer 1987; Baillard *et al.* 2014), earthquake and tsunami early warning (e.g. Lomax *et al.* 2012; Menager *et al.* 2023) or seismic traveltime tomography (e.g. Ritsema *et al.* 2011; Bozdağ *et al.* 2016; Örsvuran *et al.* 2020). Automatic phase picking is most easily implemented at short periods for first arrivals that have impulsive onsets. The onset time of later phases is usually not as distinct and more difficult to determine accurately. However, the later arrivals that constitute the main portion of the earthquake records are of great interest in that they allow for significant expansion of spatial sampling and impose more constraints on the Earth's structure (Zhao 2019). To overcome the difficulty of picking the absolute onset time of seismic phases, the correlation technique is commonly used to estimate the relative time shift between the observed signal and the synthetic signal generated from a reference model (e.g. Maggi *et al.* 2009; Zaroli

*et al.* 2010). Note, however that manual observations of onset times and their determination by observed-synthetics correlations may be slightly different due to the frequency dependence of arrival times (Hung *et al.* 2000). At long periods (above 10 s), the correlation-based measurement is particularly favoured because signals rarely have abrupt onsets. The correlation-derived time delay is related to the deviation of the Earth's structure from the reference model via well-established mathematical theory, and thus can be used to invert for the true model (Tarantola 2005).

The correlation method requires the appropriate extraction of time windows that encompass the signals and meanwhile avoid the noise-dominated parts. A crude but common practice is to select a window centred on a reference time (e.g. Ritsema *et al.* 2011; Hosseini & Sigloch 2015). The reference time is generally taken from the ray-based traveltime prediction of a specific seismic phase in a reference model, for instance, the theoretical arrival time of *S*-wave travelling in the IASP91 reference earth model (Kennett & Engdahl 1991) as predicted by the TauP program (Crotwell *et al.* 1999). The width of the time window is generally set to a multiple (e.g. triple) of the dominant seismic period. Evidently, using a constant time window is easy to implement but not optimal.

To select optimal time windows for seismic signals, Maggi *et al.* (2009) proposed an automated method and released an open-source Fortran program called FLEXWIN that has been widely adapted for applications in adjoint tomography thereafter (e.g. Tape *et al.* 2009, 2010; Bozdağ *et al.* 2016; Lei *et al.* 2020; Örsvuran *et al.* 2020). The FLEXWIN method isolates wave packets on a function derived from the waveform envelope. From a seismogram, it calculates the ratios between the short-term and long-term averages (STA and LTA) of the waveform envelope, finds the local maxima and minima on the STA:LTA time series, and identifies candidate energy time windows from various combinations of the local maxima and minima. A candidate time window starts and ends at local minima and contains one or more maxima. Only windows meeting conditions defined by a set of empirically tuned parameters are accepted as valid signal windows and passed on to the subsequent processing. Relying on factors like the event magnitude, source radiation pattern, epicentral distance and propagation medium properties, the amplitudes of seismic signals can vary over a very broad range and are difficult to predict accurately. The STA:LTA ratio used by FLEXWIN is dimensionless and has the merit of allowing the application of empirical thresholds independent of the signal amplitudes. However, the recursive moving average operations in the derivation of the STA:LTA curve can lead to a nuisance: weak signals are easily masked by preceding signals with large amplitudes. For instance, weak core reflections are hardly detectable from the derived STA:LTA curves owing to the preceding large-amplitude surface waves (Lei *et al.* 2020).

With the development of computational infrastructure and the advances in machine learning techniques in the past decades, artificial neural networks (ANN) have made remarkable achievements in various scientific and industrial applications (e.g. see Abiodun *et al.* 2018 for a review). The ANN technique consists of a training stage and a production stage. During the training phase, the parameters of the ANN are adjusted so that the correct answers can be predicted from a large number of input examples. The major time cost of ANN occurs during this stage. Once a set of model parameters has been obtained, the application of the ANN model to new data sets is generally efficient. Since about 2019, applications of ANN in seismology have boomed (e.g. Chen *et al.* 2019; Zhu & Beroza 2019; Garcia *et al.* 2021). Existing studies have proven the feasibility of extracting the seismic signal time windows using

the ANN technique (e.g. Chen *et al.* 2017; Jiang *et al.* 2022). It is a promising automated technique. Yet, ANN has not shown a prevailing performance over the traditional methods. In the latest global seismic tomography practices, FLEXWIN remains the most popular method for the automated time window selection (e.g. Lei *et al.* 2020; Örsvuran *et al.* 2020).

In this paper, we propose a simple and efficient morphological algorithm to automate the time window selection. The initial motivation for developing a new algorithm is to improve the capability to extract time windows for weak seismic signals, especially the core reflections. Core reflections are very important for imaging the lowest mantle that plays a critical role in understanding mantle dynamics. Another motivation is to develop an approach well suited to ray and first-order scattering theories, which efficiently handle massive data sets and are generally reasonable assumptions for global tomography at long period. In particular, we aim to develop a method able to extract an independent time window around each seismic phase that can be discerned on the seismogram. This allows us to perform different measurements (e.g. traveltime and potentially attenuation) for the considered phase, which can then be inverted using the appropriate kernel computed around the corresponding ray path. FLEXWIN was originally designed for full waveform imaging (FWI) driven by the parallel spectral element simulations (Maggi *et al.* 2009; Tape *et al.* 2009). The FWI technology requires a 3-D initial model capable of producing synthetic seismograms close enough to actual observations. Model parameters are updated on the basis of sensitivity kernels that are calculated numerically and of the mismatches between synthetics and observations. FWI does not require explicit knowledge on the ray paths, and is not concerned whether or not multiple phases are present in the same time window. Our new algorithm therefore complements FLEXWIN in applications of ray-based tomography.

In the following sections, we first introduce the mathematical background for the new algorithm. Weak signals are often more distinct on the original waveform envelope than on the envelope-derived STA:LTA curve which uses a moving average causing a 'smearing' problem as described previously. The window selection will thus be implemented on the waveform envelope rather than on the STA:LTA curve. We then demonstrate the implementation of the algorithm with both synthetic and real data. The data used here for the demonstration are long-period global earthquake seismograms. However, we mention that the algorithm can be used for the generic purpose of selecting an energy time window from any time series, regardless of the frequency bands or spatial scales. Finally, we apply the method to synthetic and observed seismograms for hundreds of earthquakes that occurred in a year, to investigate the performance of the method on global data set. In future work, the algorithm will be applied to a much longer period of earthquake records to improve our current mantle seismic tomography model SEISGLOB2 (Durand *et al.* 2017).

## 2 METHOD

In this section, we introduce a morphological time window selection (MTWS) algorithm to automatically detect the presence of any energy arrival and determine the beginning and end of the energy time window. The algorithm is implemented in two successive steps: peak localization and edge determination. The first step consists of applying a maximum filter to the waveform envelope to locate the primary peaks that indicate the plausible presence of seismic signals.
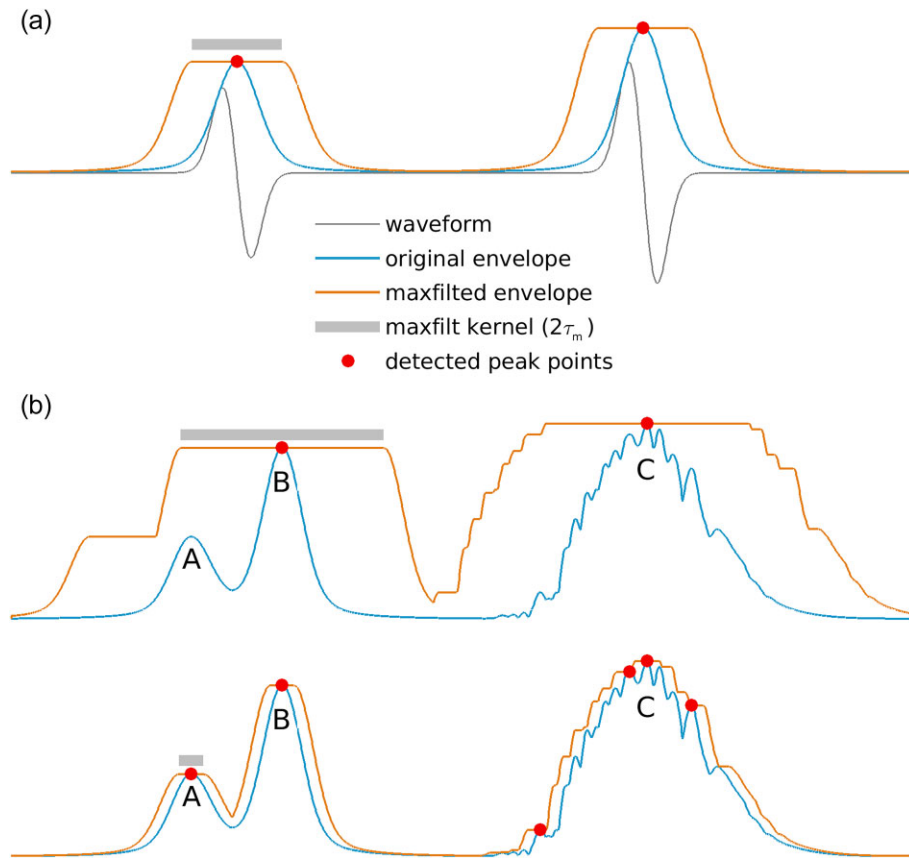
**Figure 1.** Peak localization by the maximum filter method. (a) Peak localization on the envelope of two well separated signals. The envelope of the signal waveform meets the filtered envelope at the peak points. (b) Comparisons of the peaks detected by the maximum filter method with different kernel size ($\tau_m$). An excessively broad kernel (upper) can lead to neglecting a smaller peak adjacent to a larger one. An excessively narrow kernel (lower) can lead to the selection of many secondary peaks.

The second step uses a simple Cartesian geometrical criterion to determine the beginning and end of a time window around a peak.

## 2.1 Peak localization

As illustrated in Fig. 1(a), the signals are characterized as energy wave packets on the envelope $e(t)$ calculated from the Hilbert transformation of the waveform trace. The first step in peak localization is based on the use of maximum filter, also called moving maximum. The maximum filter is analogous to the well-known mean filter, also called moving average, except that the centre sample in the moving window of width $2\tau_m$ is replaced with the maximum rather the mean value of the windowed samples, namely,

$$E(\tau) = \max(e(t), \text{ for } \tau - \tau_m \leq t \leq \tau + \tau_m) \tag{1}$$

Applying the maximum filter to the waveform envelope, one can observe that the original envelope only touches its filtered version at the peak points within the signal windows. Thus, the peaks at time $t_i$, can easily be localized from the condition

$$e(t_i) = E(t_i). \tag{2}$$

The maximum filter is characterized by a unique parameter, the width of the moving window, denoted by $2\tau_m$ (grey shaded bar in Fig. 1), which is also commonly referred to as the filter's kernel in data science. As illustrated in Fig. 1(b), if the half kernel size $\tau_m$ is larger than the distance between two local maxima (Fig. 1b, top), the lower maxima will be neglected (see peaks *A* and *B*). If $\tau_m$ is too

small (Fig. 1b, bottom) many secondary peaks can be selected (see selected local maxima around *C*). The traditional method simply selects the signals higher than their left and right neighbours as peak points—that is equivalent to a special case of the maximum filter method with an extremely narrow moving window where $\tau_m$ is equal to the sampling interval.

Obviously, $\tau_m$ should be large enough to avoid selecting too many secondary peaks. But it should be smaller than the minimum distance between the peaks of two distinguishable adjacent signals, otherwise the smaller one will be undetectable by the maximum filter method. Generally, one can choose $\tau_m$ between $0.5T_c$ and $T_c$, where $T_c$ is the dominant period of the signal.
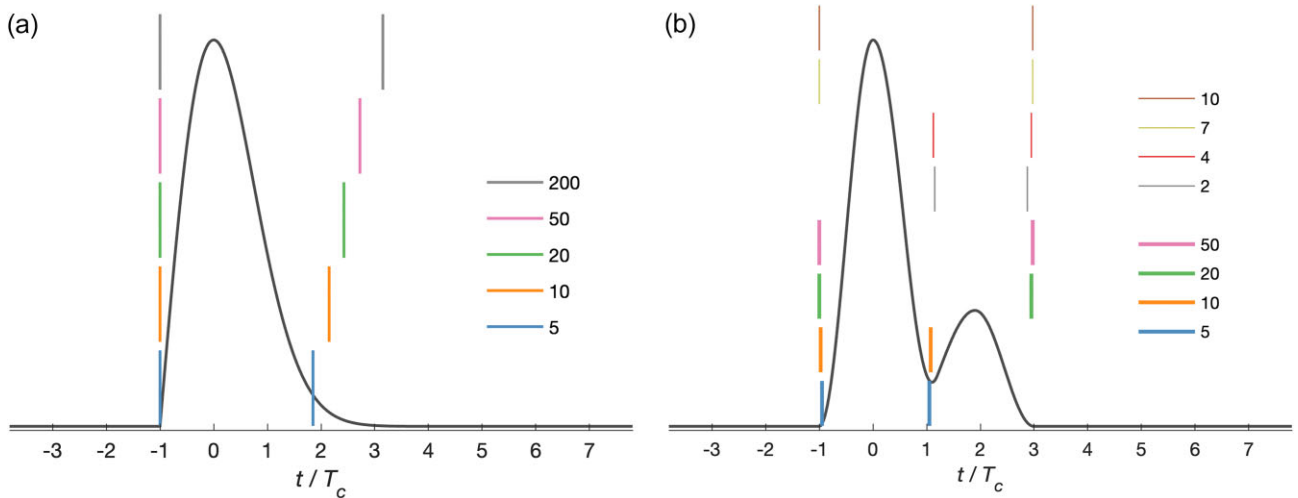
## 2.2 Edge determination

The major peaks located by the maximum filter method described above indicate the presence of energy arrivals. The remaining task is to determine the starting and ending times of the energy packets. In this subsection, we propose a maximum distance criterion to determine the locations of the bound points.

As shown in the sketch of Fig. 2(a), we consider an energy arrival with a peak at point *P* and intend to determine the edge points, *L* and *R*, of the time window that bounds the energy wave packet. By considering two auxiliary lines in opposing directions from *P* to *A* and *B*, we can reasonably identify *L* and *R* as the points farthest from lines *PA* and *PB*, respectively. In this way, the determination of time
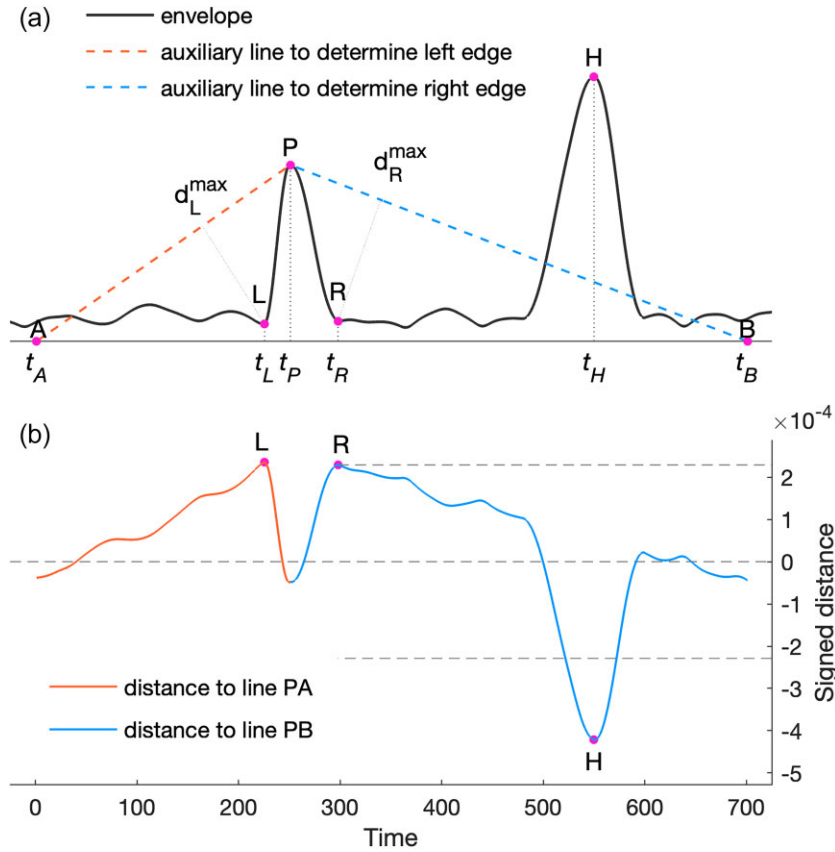
**Figure 2.** (a) Illustration of the concept of the maximum distance criterion for determining the bound points (L and R) of the energy time window peaked at point P. Points A and B lie on the baseline outside the time window. Of the points between A to P on the signal envelope, point L is the furthest from the auxiliary line PA. Similarly, point R is the furthest from line PB. Peak H corresponds to another energy time window characterized by a negative value of $d_H$ and thus will not be selected by the maximum distance criterion. (b) A numerical example for the maximum distance criterion based on eq. (3). Note that to calculate the distances to PA for $t_A < t < t_P$, we replace $(t_1, e_1)$ with $(t_A, 0)$ and $(t_2, e_2)$ with $(t_P, e_P)$ in eq. (3), and to calculate the distance to PB for $t_P < t < t_B$, we replace $(t_1, e_1)$ with $(t_P, e_P)$ and $(t_2, e_2)$ with $(t_B, 0)$.

**Figure 3.** Sensitivity of edge determination to the offset $\tau_o$. (a) Example of an impulsively starting and gradually fading energy packet. Vertical bars are placed at $t_L$ and $t_R$ determined from eq. (4) with varying $\tau_o$ as labelled in the legend. The time axis is normalized by the dominant seismic period $T_c$, so that the labelled numbers correspond to times of $T_c$. (b) Edge determination for two energy time windows that partially overlap. Thick bars are the results for the larger peak, and thinner bars for the smaller peak. See the main text for more details.
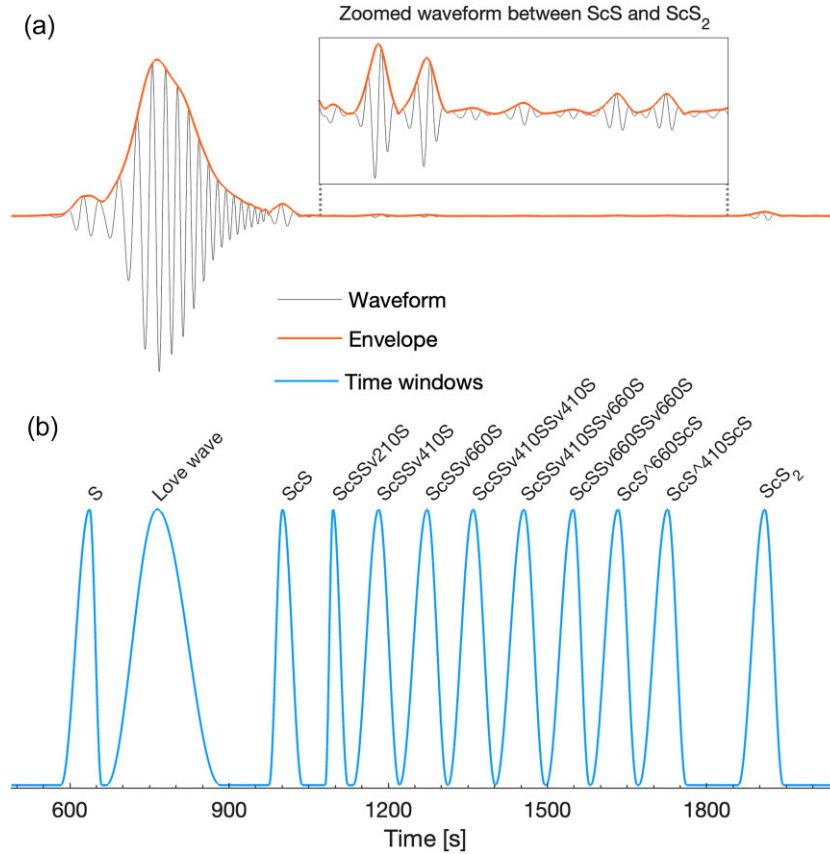
**Figure 4.** Application of the MTWS method to a synthetic transverse seismogram consisting of pure shear body waves and Love waves, excited by a transverse force applied to the surface and recorded at a distance of 27°. (a) Transverse-component seismogram bandpass filtered between 20 and 40 s periods, with an inset showing the magnified portion of the seismogram where tiny signals are detected by the MTWS method. (b) Cosine window trace c(t) calculated from eq. (5) that emphasizes weak signals. The detected tiny signals correspond to upper mantle and transition zone reflections. Phase names are labelled in accordance with the naming convention defined by FDSN. For instances, $S_V 410S$ refers to the *S* reflection from the top side of the 410 km discontinuity, and $S^{\wedge}660S$ refers to the reflection from the bottom side of the 660 km discontinuity. A digit following a phase name indicates a multiple (e.g. ScS₂ for ScSScS).

window edges becomes the solution of a simple planar geometrical problem.

Given two points $X_1$: $(t_1, e_1)$ and $X_2$: $(t_2, e_2)$ with $t_1 < t_2$, the distance from a point $X$: $(t, e(t))$ to line $X_1 X_2$ can be calculated from

$$d(t) = \frac{(t_2 - t_1)(e_1 - e(t)) + (t - t_1)(e_2 - e_1)}{\sqrt{(t_2 - t_1)^2 + (e_2 - e_1)^2}}. \quad (3)$$

The equation corresponds to the definition of a signed distance that is positive only when point $X$ is below the line $X_1 X_2$. The maximum distance criterion based on eq. (3) is illustrated in Fig. 2(b), where $t_L$ and $t_R$ are well identified as the maxima on the distance curve. We also illustrate in Fig. 2(b) a peak $H$ that is further from line $PB$ than point $R$. Since $d(t_H)$ is negative, it will not be selected by the maximum distance criterion.

The constant terms in the numerator and denominator of eq. (3) do not affect the position of the maximum and the point $L$ is determined by the maximum of

$$d_L(t) = e_P t - e(t)(t_P - t_A) \text{ for } t_A \leq t \leq t_P \quad (4a)$$

and point $R$ by the maximum of

$$d_R(t) = e(t)(t_P - t_B) - e_P t \text{ for } t_P \leq t \leq t_B \quad (4b)$$

In Fig. 2, the edges of the energy packet around $P$ are definite. The determination of $t_L$, respectively $t_R$, is thus insensitive to

the offset $t_{AP} = t_P - t_A$, respectively $t_{PB} = t_B - t_P$. However, in some cases, the edge determination may vary with the choice of the offset. This is illustrated in Fig. 3. For simplicity, we will use equal offsets $\tau_o = t_{AP} = t_{PB}$ hereafter. First, we consider, in Fig. 3(a), an energy packet that starts impulsively, peaks after one dominant period $T_c$ and decays smoothly after the peak. We determine $t_L$ and $t_R$ for $\tau_o$ in a wide range from $5T_c$ until $200T_c$. We show that when the arrival (or departure) of the energy packet is impulsive, as on the left-hand side of the signal, the determined $t_L$ is almost invariant. When the departure (or arrival) of the energy packet is smooth, as on the right-hand side of the signal, the determination of $t_R$ varies with $\tau_o$. In the latter case, larger $\tau_o$ leads to a wider window that includes more low amplitudes, although most of the energy packet is still well bounded in the determined windows.

We then show in Fig. 3(b) the case of two energy time windows that partially overlap. The edge determination is implemented for the highest peak and the lowest peak, with $\tau_o$ varying between $5T_c$ and $50T_c$ and between $2T_c$ and $10T_c$, respectively. For the highest peak, $t_L$ is almost uniquely determined regardless of $\tau_o$, whereas $t_R$ depends significantly on $\tau_o$. For smaller $\tau_o$ ($\tau_o = 5 T_c$ or $10 T_c$), each energy time window is well isolated, whereas for larger $\tau_o$ ($\tau_o = 20T_c$ or $50T_c$), the overlapping energy packets are selected as a single time window. Similar phenomena can be observed for the smaller peak. As a rule of thumb, it is thus
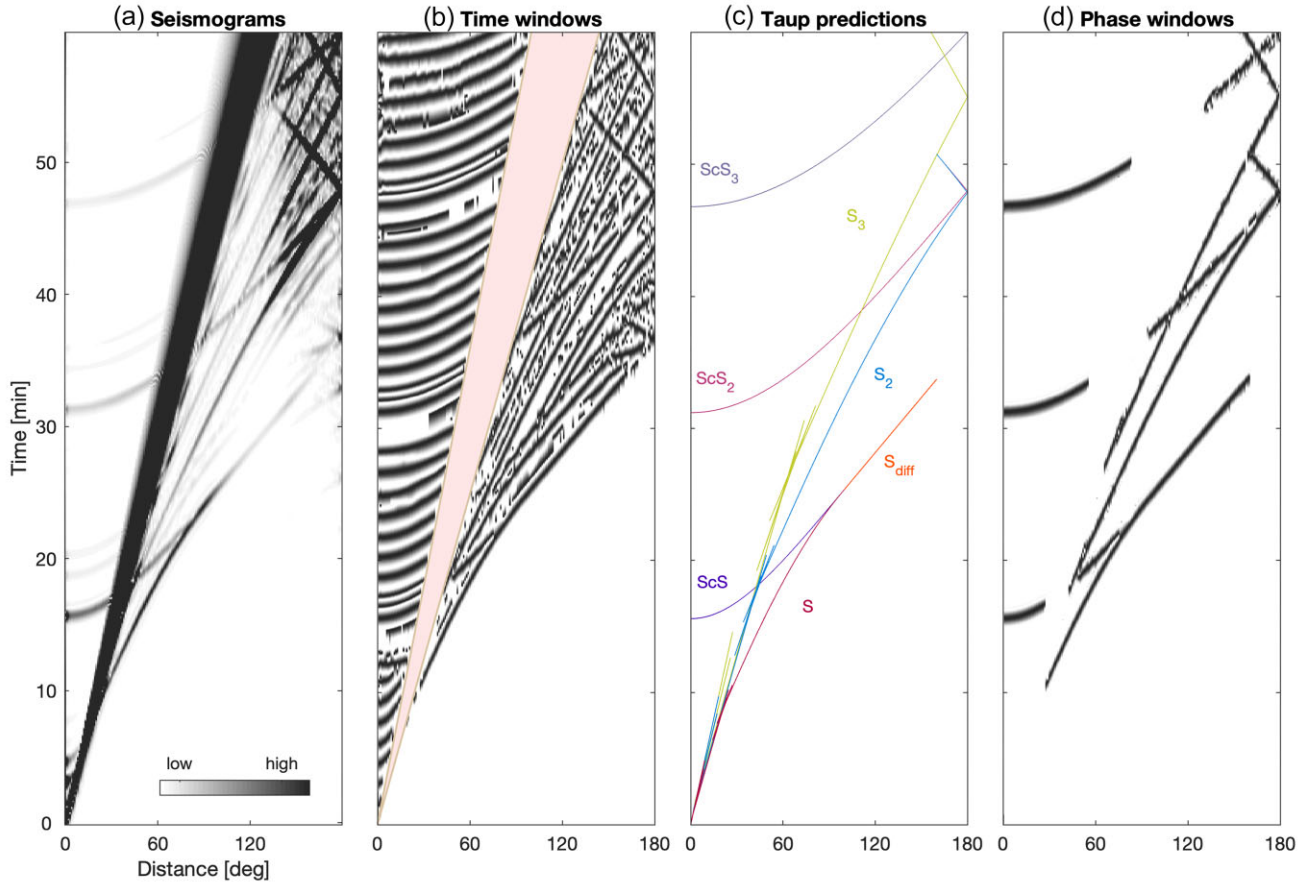
**Figure 5.** Application of the MTWS method to synthetic transverse seismograms consisting of pure shear body waves and Love waves, excited by a transverse force applied to the surface and recorded at distances ranging from 0° to 180°. The period band and the MTWS parameters are the same as in Fig. 4. (a) Envelopes of transverse-component synthetic seismograms. (b) Cosine window traces for the MTWS determined time windows. The pink shaded area corresponds to surface waves with slownesses between 25 and 36 s deg$^{-1}$. (c) Theoretical time–distance curves for some specific shear waves as predicted by the Taup program. (d) Time windows selected from (b) by checking whether a Taup prediction falls within the time window.

recommended to either choose a relatively small $\tau_o$ (e.g. $4T_c$) in order to isolate single energy time windows, which is well suited for ray-based tomography, or to choose a large $\tau_o$ (e.g. $9T_c$) and then move the edge to the internal minimum if a larger peak is present in the determined time window. The latter is illustrated in Fig. 3(b), where the window determined for the lower peak with $\tau_o = 7T_c$ or $10T_c$ encompasses the higher peak on the left, so that we should move the left edge from $t = -T_c$ to the internal minimum at $t = \sim T_c$.

## 3 DATA EXAMPLES

In the following subsections, we demonstrate the performance of the MTWS method for detecting time windows on both synthetic and observed seismograms. For simplicity, we will consider transverse components since they contain only pure SH-type waves. After the time window determination is performed on both synthetic and real seismograms, we will apply a quality control procedure based on cross correlation to select the time windows that correspond to specified seismic phases. The cross correlation also provides time delay estimates for the windowed seismic phases.

### 3.1 Synthetics

The synthetic seismograms are obtained from simulations by the spectral element software AxiSEM (Nissen-Meyer *et al.* 2014). We download the Green's function database for the AK135f Earth model (Kennett *et al.* 1995; Montagner & Kennett 1996) that are provided by the IRIS Syngine web service (Hutko *et al.* 2017). Then, using the Python application Instaseis (van Driel *et al.* 2015), we convolve the Green's functions to the same sources and receivers as for the real seismograms. The record duration is one hour and the available seismic periods of the database range from 5 to 100 s (for details, see the official site http://ds.iris.edu/ds/products/syngine/).

Fig. 4(a) shows a synthetic transverse seismogram obtained for a horizontal transverse force located at the surface, so that shear waves are well excited, and a receiver located at 27°. At this distance, the strong Love wave is preceded by the triplicated *S* phases resulting from the 410 and 660 km discontinuities and immediately followed by the *ScS* phase. The seismogram is bandpass filtered between 20 and 40 s (25–50 mHz), with a central period $T_c$ at around 27 s (37.5 mHz). We apply the MTWS method to the waveform envelope, with $2\tau_m = 0.75T_c$ and $\tau_o = 9T_c$. In this study, we prefer a relatively large $\tau_o$, because we intend to investigate the effect of time window duration on the quality of correlation time delay measurements and
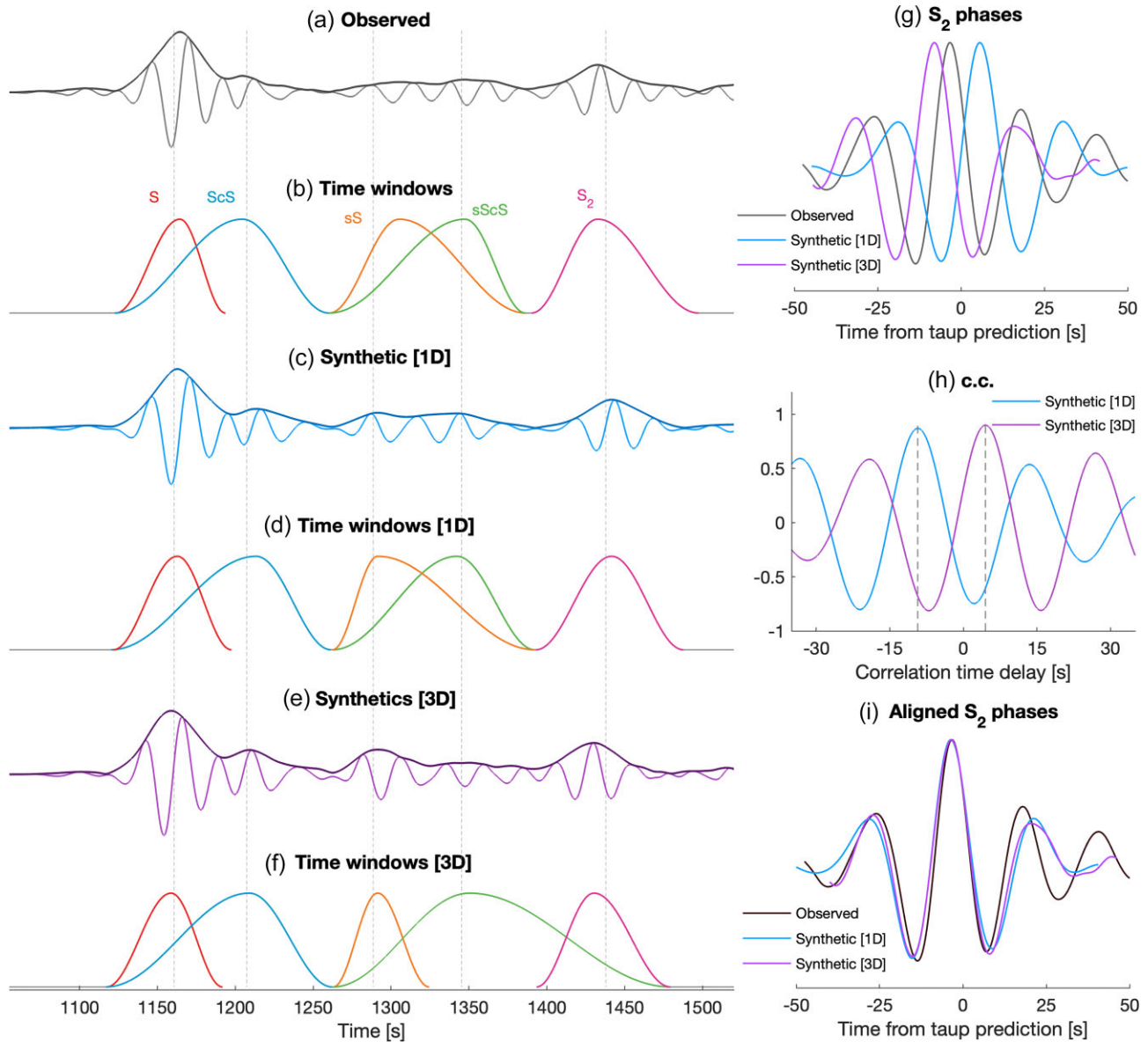
**Figure 6.** Time windows determined by the MTWS method for SH waves from a deep earthquake in Indonesia (2008-02-07 20:58:18.9, $M_w$ 5.7, depth 320 km) recorded in Russia at station IU.PET about 70° away. (a) Real seismogram (grey) and envelope (darker). (b) Cosine window representation of time windows determined from the real seismogram in (a). (c) Synthetic seismogram using the AK135f reference model (blue) and its envelope (darker). (d) Cosine window representation of time windows determined from the synthetic seismogram in (c). (e) Synthetic seismogram in the 3-D tomographic model S40RTS (purple) and envelope (darker). (f) Cosine window representation of time windows determined from the synthetic seismogram in (e). The vertical dash lines mark the theoretical arrival times of seismic phases computed using the Taup program. (g) Synthetic and observed $S_2$ phases to be correlated. (h) Cross-correlation function between the two $S_2$ phases in (g). The dashed line indicates the time delay that maximizes the correlation coefficient (c.c.). (i) Aligned synthetic and observed $S_2$ signals.

a small $\tau_o$ will lead to neglecting wide time windows. The same parameters will be used for all data hereafter.

For a better display of signals of any strength, we compute the asymmetric cosine window function

$$c(t) = \begin{cases} \frac{1}{2}\left(1 - \cos\frac{t-t_L}{t_P-t_L}\pi\right), & t_L \le t \le t_P \\ \frac{1}{2}\left(1 - \cos\frac{t-t_R}{t_P-t_R}\pi\right), & t_P \le t \le t_R \end{cases} \quad (5)$$

for each MTWS detected time window. The notations are the same as in Fig. 2(a). Note that eq. (5) corresponds to a strongly non-linear transform of the raw waveform. It is used in this paper for an equalized display of both weak and strong signals. The derived

asymmetric cosine window trace for the raw trace shown in Fig. 4(a) is displayed in Fig. 4(b). The weak *ScS* phase is well identified and isolated by the MTWS method. In addition, between *ScS* and *ScS₂* waves, many time windows are identified, although the corresponding signals are hardly visible on the raw trace. The enlarged inset in Fig. 4(a) confirms the existence of these very low amplitude signals in the detected time windows. Theoretical traveltimes predicted by the Taup program (Crotwell *et al.* 1999), indicate that the tiny signals correspond to the reflections from upper mantle (210 km) and transition zone (410 and 660 km) interfaces. These results confirm the capability of the MTWS method to extract time windows even for very weak phases.
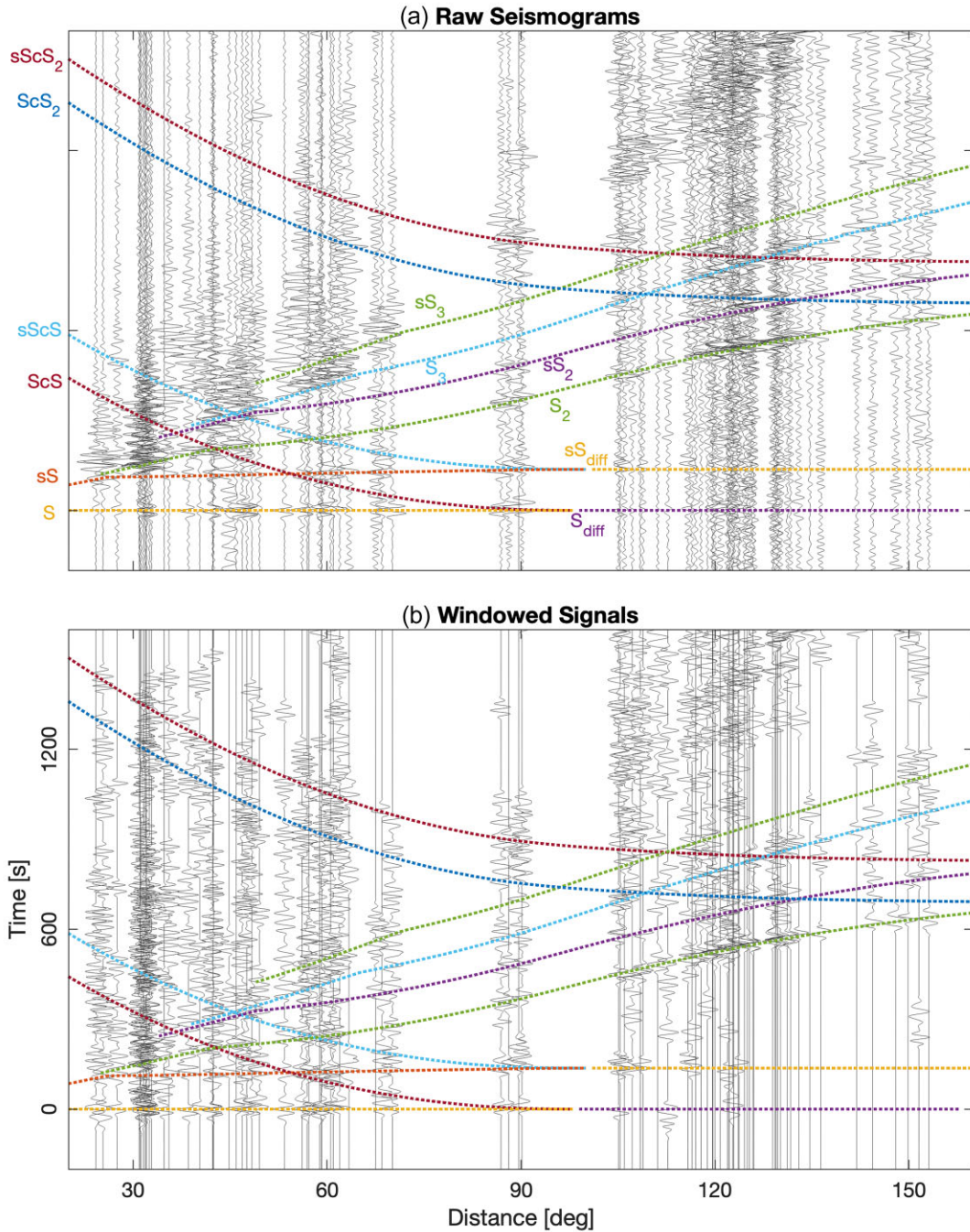
**Figure 7.** (a) Transverse-component seismograms for the same event as in Fig. 6 for a large range of epicentral distance. Traces are bandpass filtered between 20 and 40 s and normalized by their own maxima. (b) Windowed and normalized seismic signals extracted from traces in (a). The time windows are obtained from the MTWS method. The normalization is based on eq. (6). Theoretical travel–distance curves of typical SH waves ($S$, $S_2$, $S_3$, $ScS$, $ScS_2$, $ScS_3$, $S_{diff}$ and their depth phases) are superimposed as coloured dotted curves. Traces are shifted and aligned on the theoretical $S/S_{diff}$ arrival time.

In contrast to the well bounded body waves, the Love wave window determined by MTWS in Fig. 4(b) appears narrower than the Love wave packet shown in Fig. 4(a). In order to capture the entire Love wave window, a larger $\tau_o$ would be necessary. However, as it has been demonstrated in Fig. 3(b), very large $\tau_o$ can lead to the detection of very long time window embracing multiple phases. For instance, if $\tau_o = 30T_c$, the detected time window for the direct $S$ arrival will start before $S$ and stop after $ScS$, which is not favourable.

Since our focus in this study is on body waves, we favour smaller $\tau_o$ and mute the surface waves.

In Fig. 5, we present the results for a wide range of epicentral distances. Fig. 5(a) shows the envelopes of transverse components, bandpass filtered between 20 and 40 s. The MTWS method is applied to each trace, with the same parameters used in Fig. 4. Again, for a better display, we plot in Fig. 5(b) the cosine window traces for each distance based on eq. (5). We mute surface waves by removing
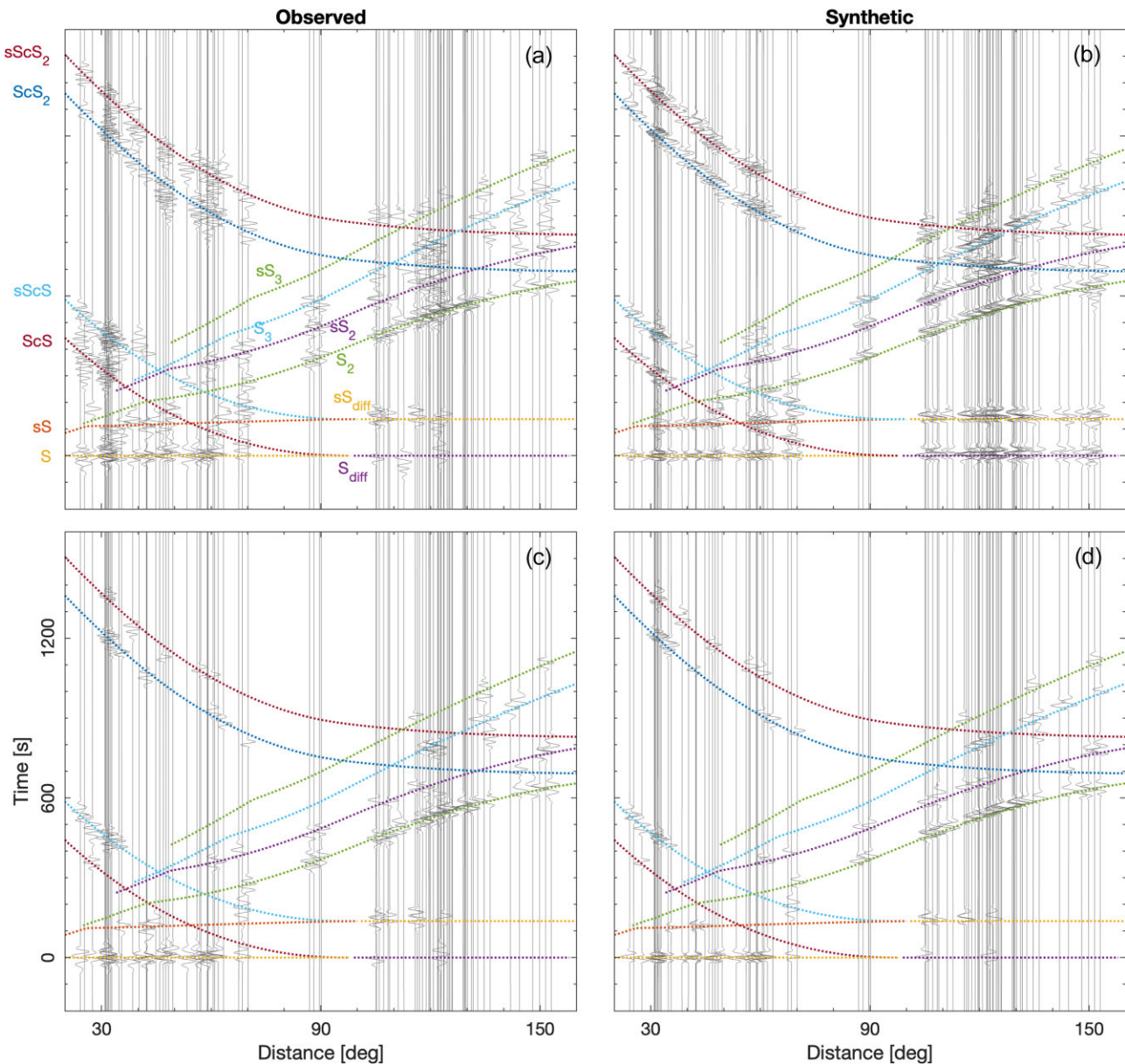
**Figure 8.** Selection of signals based on phase association and observation–synthetic correlation. Candidate time windows for the specified SH phases: (a) observation; (b) synthetics. Highly correlated (c.c. $\geq$0.8) time windows: (c) observation; (d) synthetics. Theoretical time–distance curves are superimposed as coloured dotted curves. Traces are aligned with the theoretical $S/S_{\mathrm{diff}}$ arrival time. All signals are normalized using eq. (5).

## 3.2 Observations

In this subsection, we present examples of applying the MTWS method to real seismograms. We download the LH channel records

any time window whose slowness is between 25 and 36 s deg$^{-1}$ (pink shaded area in Fig. 5b). Comparing Fig. 5(b) to the theoretical traveltimes of commonly observed SH-type phases ($S$, $S_2$, $S_3$, $ScS$, $ScS_2$, $ScS_3$ and $S_{\mathrm{diff}}$) shown in Fig. 5(c), we note the presence of many coherent branches other than $ScS$ and its surface multiples, arriving after the surface waves in Fig. 5(b). As demonstrated in Fig. 4, these branches correspond to the reflections from the upper mantle and transition zone discontinuities. Finally, by only plotting the time windows that include the ray-based traveltime predictions in Fig. 5(c), one can easily select the desired time windows for the specified seismic phases (see Fig. 5d).

of earthquakes from the IRIS data centre, and rotate the horizontal components to yield the transverse component. We use Python program ObspyDMT (Hosseini & Sigloch 2017) for data requests and process data with Python package Obspy (Krischer *et al.* 2015). Synthetic seismograms are also computed for comparison. The synthetics for a 1-D reference model are extracted from the IRIS Syngine database using the Instaseis program (van Driel *et al.* 2015). The synthetics for the 3-D model S40RTS (Ritsema *et al.* 2011) are calculated with the program SPECFEM3D_GLOBE (Komatitsch & Tromp 2002a, b).

Fig. 6(a) shows an example of a real seismogram containing SH waves arising from a 320 km deep earthquake in Indonesia recorded in Russia, about 70° away. The corresponding synthetics using the 1-D and 3-D models are plotted in Figs 6(c) and \(e). We apply the MTWS method to the three signals and their cosine time windows (see eq. 5) are shown in panels (b), (d) and (f). In this example,
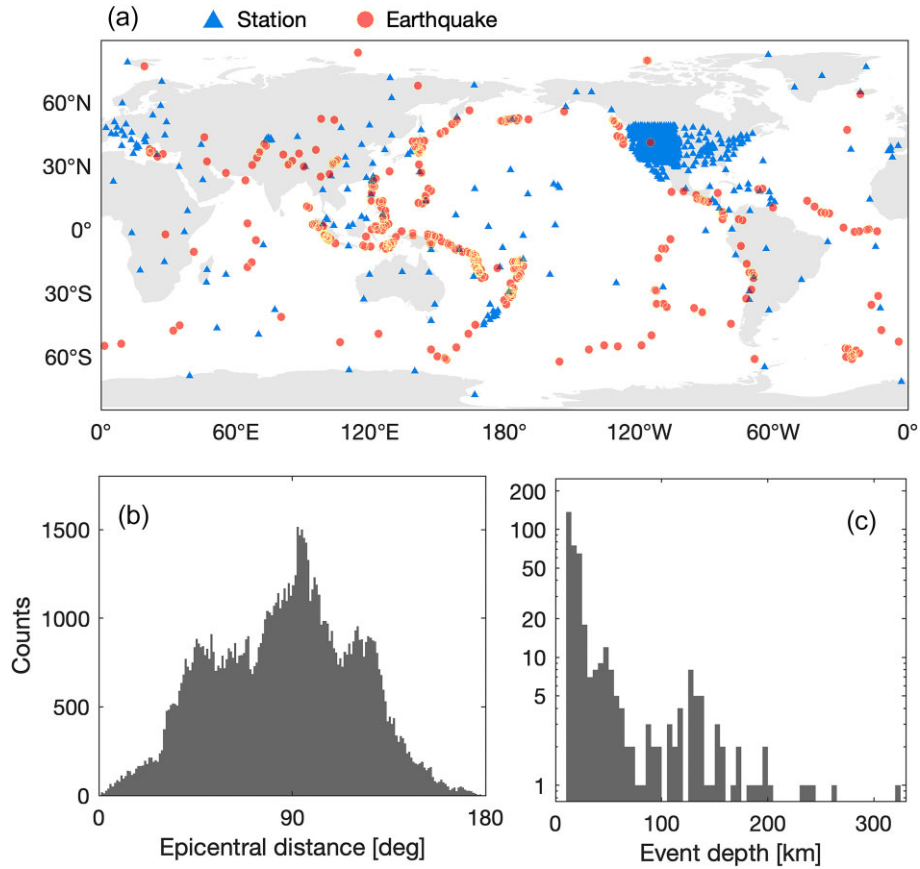
**Figure 9.** (a) Geographical distribution of ∼400 earthquakes (red dots) and ∼950 seismic stations (blue triangles) in 2008. (b) Histogram of epicentral distance. (c) Histogram of event depth.
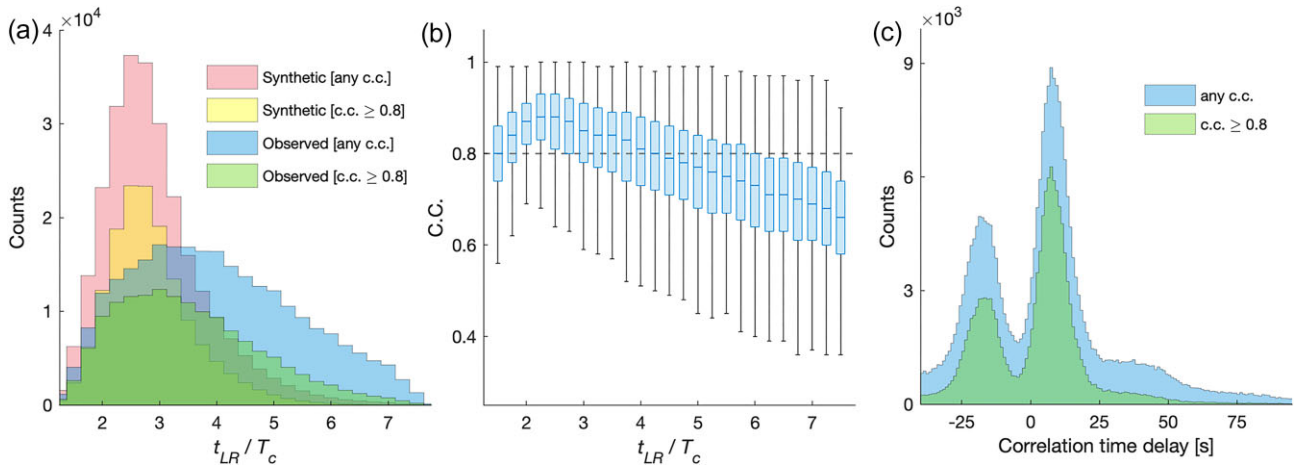


**Figure 10.** (a) Histograms of time window width ($t_{LR}$) normalized by the dominant seismic period ($T_c$). The pink and blue shaded areas are the statistics on all correlation pairs. The yellow and green shaded areas are the statistics on the highly correlated pairs. (b) Boxplot for the maximum correlation coefficients between the observed and synthetic signals, as a function of $t_{LR}/T_c$ for the observed seismograms. Each blue box around the median goes from the first to the third quartile, and the wiskers from each quartile to the pseudo minimum and maximum. (c) Histograms of correlation time delay for all correlation pairs (blue) and highly correlated pairs (green).

the $S$ and $S_2$ phases are prominent and their time windows are well determined. The $ScS$ phase is less energetic and interferes with the $S$ phase. The weak waveforms of $sS$ and $sScS$ overlap. Our code selects the $S$ and $S_2$ time windows and rejects the $ScS$, $sS$ and $sScS$ phases. Once a pair of time windows for a seismic phase is determined from both observed and synthetic seismograms (e.g. the $S_2$ phases

in Fig. 6g), we can compute the cross-correlation function between the two signals (Fig. 6h). The time lag between observation and synthetic is thus given by the correlation delay that maximizes the correlation coefficients (*c.c.*, Fig. 6h) and allows the optimal alignment of the signals (Fig. 6i). The 3-D model predicts an $S_2$ phase in better agreement with the observations than the simple 1-D model.
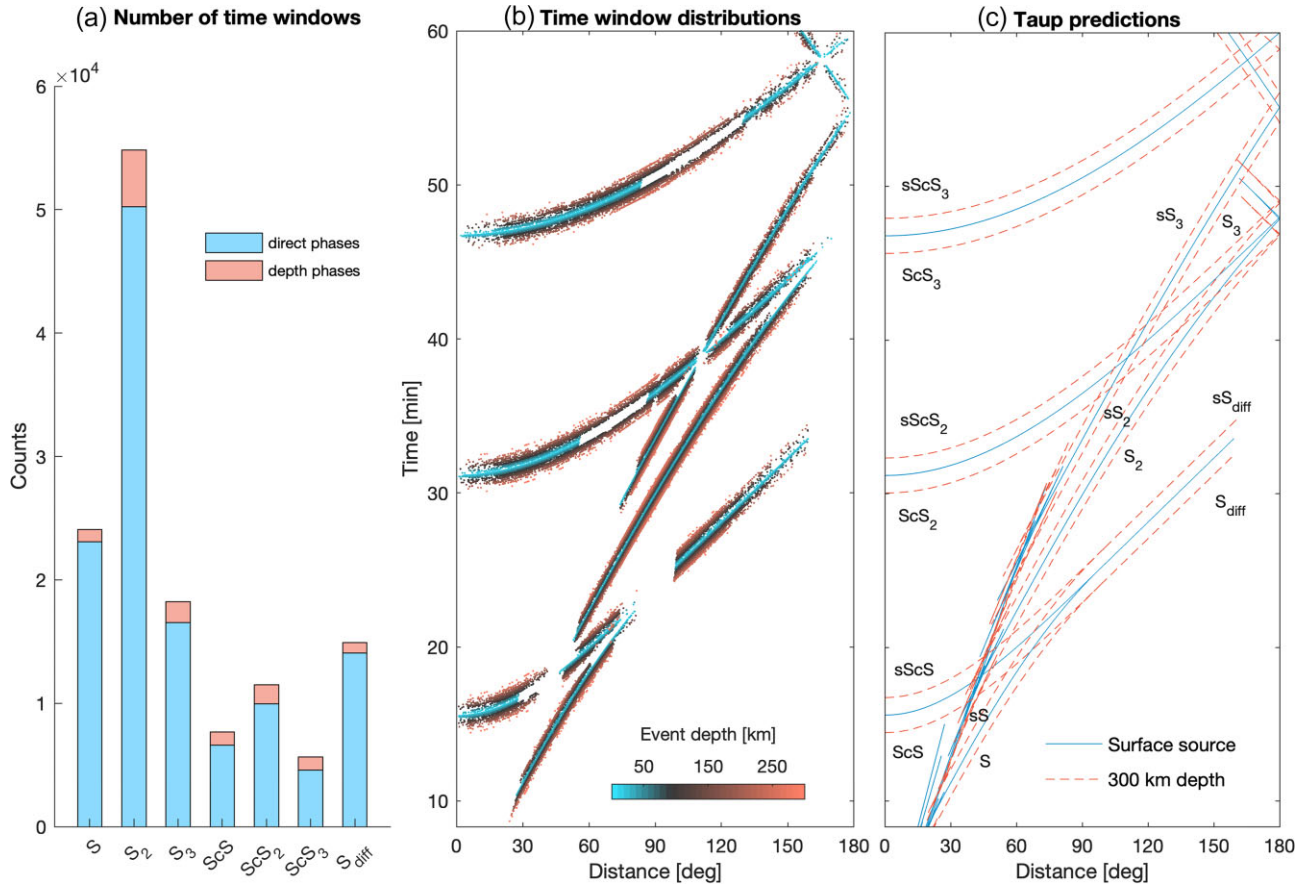
**Figure 11.** (a) Number of estimates for each SH phase. The statistics is based on ~140 000 estimates of correlation time delays with c.c. ≥ 0.8. (b) Time–distance distribution of the SH phases. Each signal is represented by a coloured dot based on the event depth. (c) Theoretical time–distance curves for the phases as shown in (b). Solid lines for a surface source and dashed lines for a 300 km deep source.

In Fig. 7(a), we show the transverse component seismograms for the same event as in Fig. 6 but for a wide range of epicentral distances. We apply the MTWS method to all traces with the same parameters as before and select the traces for which one or more seismic phases are detected. The seismic phases are labelled on the panel, with the theoretical time–distance curves are plotted in dotted lines. For better display of weak signals, we normalize the original waveform trace $s(t)$ with the ratio of the cosine window trace $c(t)$ to the envelope $e(t)$, namely,

$$\bar{s}(t) = s(t) \frac{c(t)}{e(t)}. \tag{6}$$

After this non-linear transform, weak signals become prominent (Fig. 7b). In contrast to the synthetic test shown in Fig. 5, one can observe from Fig. 7(b) that, for real seismograms, many time windows have been detected where no theoretical seismic phases are present. Accounting for the interference of noise and the inherent complexity of real seismic wavefields originating from the 3-D heterogeneities, these observations are not surprising. Thus, the next step is to select the time windows corresponding to real arrivals.

### 3.3 Time window selection

Before any complex processing, a simple pre-analysis retention criterion is applied. We estimate the noise variance before the first

arrival ($S$ or $S_{\text{diff}}$), and the signal variance afterwards. Only recordings for which the signal-to-noise variance ratio is larger than 1.5 are considered.

For the noise-free synthetic data (Fig. 5d), there are little doubts in the association of a detected time window with a specific seismic phase. However, for real seismograms that contain noise and waves seeded by undetermined 3-D heterogeneities, ambiguity arises in associating a time window to a specific seismic phase. It is well known that the time shift between the synthetic phase and the observed phase can reach up to dozens of seconds (e.g. Zaroli *et al.* 2010; Ritsema *et al.* 2011; Hosseini & Sigloch 2015; Lei *et al.* 2020; Örsvuran *et al.* 2020). Within the possible range of time offset from the theoretical arrival time calculated in a reference model, there may be multiple candidate time windows detected on the real seismogram for a specific seismic phase. Thus, the comparison with the theoretical arrival time cannot definitely determine the appropriateness of the time window. For a specific phase, we consider any time window with a peak within ±45 s around the theoretical traveltime to be candidate for the phase. The range of ±45 s is wider than typical synthetic-observation correlation time delay, in order to avoid missing time windows in the phase association. We then discard any time window that contains multiple seismic phases, unless it contains only a phase and its depth phase (e.g. $S$ and $sS$) that have similar traveltimes and ray paths in the deep mantle. To confirm the phase association, we correlate the observed signals in the candidate time windows with the corresponding synthetic signals, and
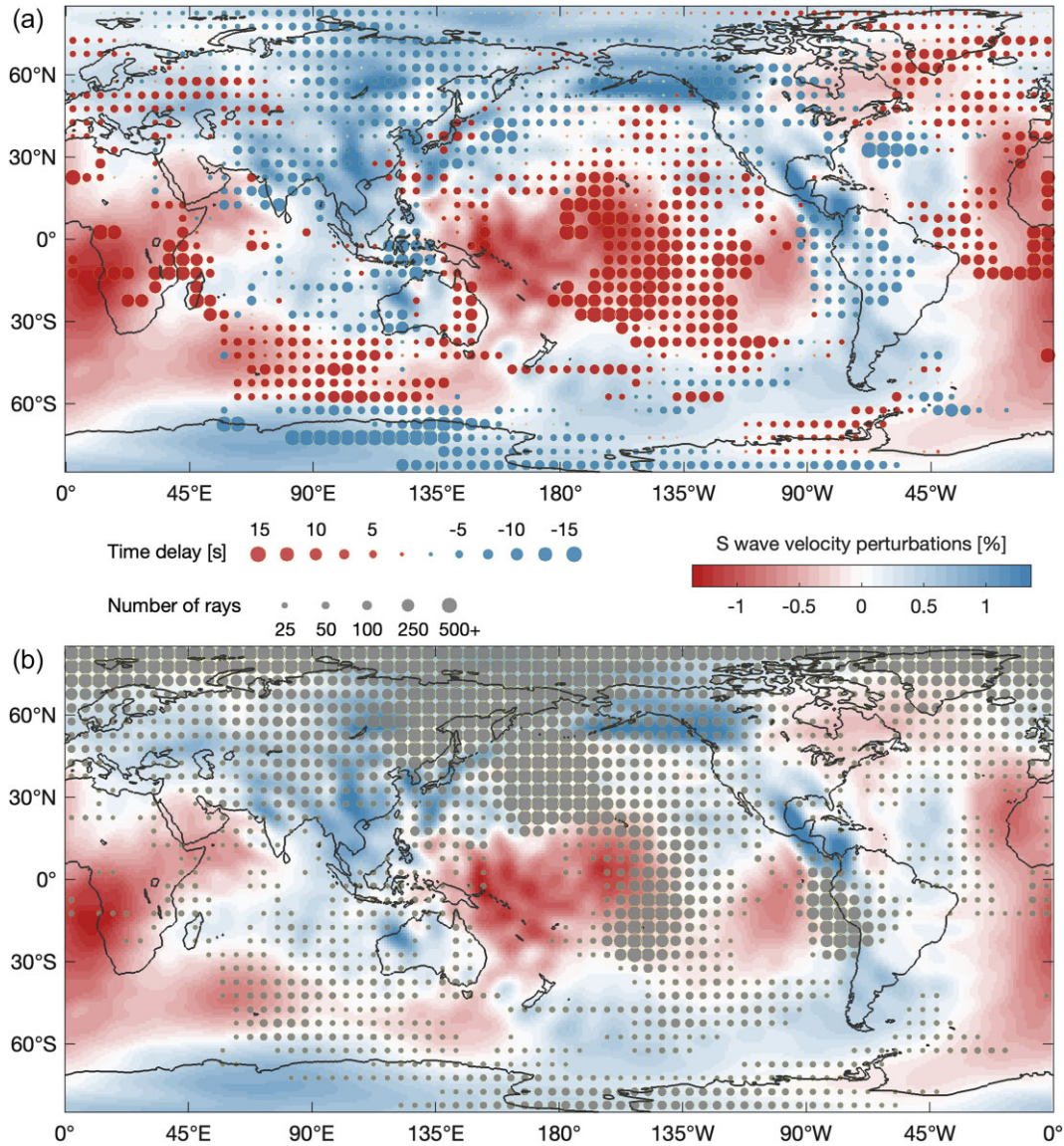
**Figure 12.** (a) Map of correlation delay times for the $S_{\mathrm{diff}}$ phases averaged over equal areas of 500 km × 500 km surface. An overall mean value of 3.8 s is removed. The size of dots varies with the value of the time delay. The background image displays the CMB shear velocity perturbations in the tomographic model SEISGLB2 (Durand *et al.* 2017). Red colour represents slower seismic velocity/positive time delay. Blue colour represents faster seismic velocity/negative time delay. (b) The size of dots represents the number of $S_{\mathrm{diff}}$ ray paths that sample each 500 km × 500 km area.

accept only the windows where the time-shifted signals are highly correlated.

Let us illustrate the time window selection using Fig. 6. The $S + ScS$ phases partially overlap and these two phases have distinct ray paths. The same applies to the $sS + sScS$ phases. The $S_2$ phase, on the other hand, is isolated and our code chooses the corresponding time window straightforwardly.

As was illustrated in Fig. 3, our code successfully identifies the energetic $S$ phase from the $ScS$ phase and selects a reasonable time window (Figs 6b, d and e). Finally, our code selects the time windows for $S$ and $S_2$, discarding the $ScS$, $sS$ and $sScS$ (two arrivals in the same interval). We then extract from the raw trace the waveforms delimited within the detected time windows and correlate the observed and synthetic signals. No tapering is applied before correlation. The observation-synthetic delay time is derived from the maximum of the correlation function. As illustrated in Figs 6(g)–(i) in the case of $S_2$, the observed and synthetic signals are highly

semblable (*c.c.* ~1), so that we deem the observed $S_2$ phase to be reliable. The $S$ arrival (not shown here) is obtained in the same way.

The procedure is applied to each seismogram in Fig. 7(a). The normalized observed and synthetic signals in the candidate time windows for the specified seismic phases are plotted in Figs 8(a) and (b), respectively. We correlate the corresponding pairs of observed and synthetic signals extracted from raw seismograms. A threshold of 0.8 for *c.c.*, as adopted by many authors (e.g. Zaroli *et al.* 2010; Hosseini & Sigloch 2015), is applied for discarding weakly correlated signals. The remaining observed and synthetic normalized signals are displayed in Figs 8(c) and (d), respectively.

## 4 APPLICATIONS AND STATISTICS

Following the same procedure and with the same parameters as described above, we apply the MTWS method to a global data set of real and synthetic transverse-component seismograms generated

by ∼400 earthquakes with magnitudes ranging from 5.5 to 7.5 and recorded at ∼950 seismic stations with horizontal LH channels available in 2008 (Fig. 9a). The histograms of epicentral distances and event depths are shown in Figs 9(b) and (c), respectively. The epicentral distances are mainly concentrated between 30° and 140°. About 72 per cent of the events are shallower than 30 km. Only 48 events (12 per cent) are deeper than 100 km, with the deepest at 320 km. All real seismograms are downloaded from the IRIS data centre and the synthetic seismograms are obtained from the IRIS Syngine database based on the source parameters provided by the GCMT catalog (Ekström *et al.* 2012).

From ∼140 000 traces of seismograms, we are able to obtain ∼270 000 pairs of candidate time windows for the considered SH body phases ($S$, $S_2$, $S_3$, $ScS$, $ScS_2$, $ScS_3$, $S_{\text{diff}}$ and their depth phases), with ∼140 000 pairs with *c.c.* ≥ 0.8. In Fig. 10(a), we show the statistics of the time window width ($t_{\text{LR}}$) for all the correlation and highly correlated estimates. For the synthetic signals, $t_{\text{LR}}$ is mainly concentrated between about $2T_c$ and $3.5T_c$ (∼84 per cent for all *c.c.* and ∼80 per cent for *c.c.* ≥ 0.8). For the observed phases, $t_{\text{LR}}$ distributes in a broader range, mostly between around $2T_c$ and $5T_c$ (∼85 per cent for all *c.c.* and ∼75 per cent for *c.c.* ≥ 0.8) and is systematically wider than the synthetic match. Comparing the histograms of $t_{\text{LR}}$ for the observed phases before and after applying the *c.c.* threshold, one can note a greater reduction in the percentage of wider time windows. For $t_{\text{LR}} \sim 2T_c$, only ∼20 per cent of the time windows with *c.c.* < 0.8 are discarded, whereas for $t_{\text{LR}} \sim 7T_c$, over 80 per cent are discarded. This suggests the feasibility of preconditioning the time window selection prior to correlation. Discarding time windows on real seismograms with $t_{\text{LR}}$ above a threshold (like $6T_c$) before correlation can speed up the computation and will not lead to a notable loss of data.

The boxplot in Fig. 10(b) displays the variation of *c.c.* with respect to $t_{\text{LR}}$ for all the 270 000 observed signals. For $t_{\text{LR}}$ between around $2T_c$ and $4.5T_c$, and especially between $2T_c$ and $3T_c$, most *c.c.* values are above the threshold of 0.8. Contrarily, for most of $t_{\text{LR}} > \sim 5T_c$, the *c.c.* values are below 0.8. Low *c.c.* is not the only common feature for long time windows. They often lead to extremely large positive time delays over 50 s, well beyond the range of pooled delays that lies within ±25 s and occupies ∼77 per cent of all estimates (Fig. 10c). After applying the *c.c.* threshold, the number of estimates with time delays beyond 25 s is largely reduced. Almost all delays longer than 50 s disappear and the percentage of time delays within ±25 s increases to ∼90 per cent. The bimodal pattern of the histograms in Fig. 10(c) originates from the ray coverage. The higher peak with positive time delays can be ascribed to the ray paths crossing the slower mantle beneath Pacific. The lower peak with negative time delays is due to the ray paths crossing the faster mantle beneath Eurasia and America continents.

In Fig. 11(a), we show the statistics for each detected SH phase. The number of phases detected with 1 yr of data is already of the order of what is currently used in global tomographic models (e.g. Ritsema *et al.* 2011; Moulik & Ekström 2014). The majority is the $S$ phase and its surface multiples. We have more $S_2$ phase than other phases, because it is a prominent phase that is observable in a wide distance range. We also have many core reflections and diffractions successfully detected (∼7500 $ScS$; nearly 15 000 $S_{\text{diff}}$). The $S_{\text{diff}}$ time delay estimates are critical to improving the tomographic resolution at the bottom of the mantle. In Fig. 11(b), we project the selected signals onto the time–distance panel. Each dot corresponds to a signal at given epicentral distance and arrival time. The dots are colour-coded and sorted by event depth. The theoretical traveltime curves are shown in Fig. 11(c) for comparisons. It can be seen

that, with increasing event depth, isolation of the depth phase from the direct phase becomes feasible. In summary, Fig. 11 reveals the effectiveness of the MTWS method in capturing the weak core reflections and diffractions, as well as its ability to separate depth phase from direct phase.

In Fig. 12(a), we map the average time delays computed over equal areas of 500 km × 500 km, for $S_{\text{diff}}$ phases with the global average removed. The corresponding density map of $S_{\text{diff}}$ ray paths sampling the CMB is displayed in Fig. 12(b). For comparison, we also show in background the map of shear velocity anomalies at the CMB in the tomographic model SEISGLOB2 (Durand *et al.* 2017). One can see that the average time delays agree well with the seismic velocity variations. Local discrepancies (like in Europe and the Western Atlantic) can generally be ascribed to poor sampling, as shown in Fig. 12(b). The comparison confirms the capability of the MTWS algorithm to detect weak signals and measure delay times.

## 5  CONCLUDING REMARKS

In the era of big data, automated algorithms are indispensable. In this study, we propose a new automated algorithm (morphological time window selection, MTWS) to detect and extract time windows around seismic signals. The MTWS algorithm is a parametric method. It requires only two user-specified parameters, namely, the maximum filter kernel size for peak detection ($2\tau_m$) and the off-peak time-shift ($\tau_o$) for edge determination. The MTWS algorithm is characterized by its high flexibility in the choice of parameters. As a rule of thumb, for extracting long period body phases as demonstrated in this study, we suggest $2\tau_m = 0.75T_c$ and $\tau_o = 4T_c$, though choosing parameters from reasonable ranges (e.g. $0.5T_c \leq 2\tau_m \leq T_c$, $4T_c \leq \tau_o \leq 9T_c$) do not affect the results much.

The method is morphology-based and adapts to the shape of signals. It selects time windows of seismic signals that are more optimal than the constant time windows around reference time in the conventional approach. The MTWS method is superior in its ability to detect weak signals and extract their time windows appropriately. The detection is implemented on the raw seismogram envelope that retains the fidelity of signal energy bursts, rather than on any derived curves that may distort or obscure the signal. The performance of the method is independent of absolute amplitudes. In principle, it can detect any energy wave packets that are distinguishable from background fluctuations. The results of MTWS can be used as the training data to improve the artificial neural networks for automatic time selection.

The MTWS algorithm is easy to implement and computationally efficient. We have demonstrated the efficiency and robustness of the method with both synthetic and observed seismograms, making the algorithm very suitable for the automated process of massive data sets. We applied the method to ∼140 000 transverse-component seismograms recorded in a year. Based on the statistical relations between the time window width and the quality of time delay measurement, long-time windows (e.g. $>5T_c$) generally lead to poor similarity between observed and synthetic signals or to unreasonable time delays (e.g. >50 s). Thus, we propose to discard long time windows, or apply additional constraints to long time windows for further quality control. Besides the prominent $S$ phases and surface multiples, we have also obtained a considerable amount of weak phases including $S_{\text{diff}}$, $ScS$ and its surface multiples. The CMB-related phases are essential in improving the seismic tomographic resolution in the lowest mantle. With the demonstration data and an uneven distribution of events and stations, we have already been

able to build a map of $S_{\text{diff}}$ delay times that agrees well with existing tomographic model and a delay time database of the order of what is currently used for global tomography. In future work, we will apply the MTWS algorithm to the seismic records accumulated over the past decades that are available for global tomography. With more time delay measurements (more phases, especially the core phases) and better spatial sampling (more events and stations), we will build improved seismic imaging for the global lower mantle.

## DATA AND CODE AVAILABILITY

## ACKNOWLEDGMENTS

## AUTHOR CONTRIBUTION STATEMENT

ED designed and supervised the project. LL designed the method, developed the code and processed the data. All authors participated the analysis of results. LL produced the figures with aids from SD. LL wrote the original draft and all authors participated in the revision of the manuscript.

## CONFLICT OF INTEREST

The authors declare no conflicts of interest with respect to the research, authorship and publication of this paper.

## REFERENCES

Abiodun, O.I., Jantan, A., Omolara, A.E., Dada, K.V., Mohamed, N.A.E. & Arshad, H., 2018. State-of-the-art in artificial neural network applications: a survey, *Heliyon,* **4,** e00938, doi:10.1016/j.heliyon.2018.e00938.

Baer, M. & Kradolfer, U., 1987. An automatic phase picker for local and teleseismic events, *Bull. seism. Soc. Am.,* **77**(4), 1437–1445.

Baillard, C., Crawford, W.C., Ballu, V., Hibert, C. & Mangeney, A., 2014. An automatic kurtosis-based P- and S-phase Picker designed for local seismic networks, *Bull. seism. Soc. Am.,* **104**(1), 394–409.

Bozdağ, E., Peter, D., Lefebvre, M., Komatitsch, D., Tromp, J., Hill, J., Podhorszki, N. & Pugmire, D., 2016. Global adjoint tomography: first-generation model, *Geophys. J. Int.,* **207,** 1739–1766.

Cansi, Y., 1995. An automatic seismic event processing for detection and location: the P.M.C.C. method, *Geophys. Res. Lett.,* **22,** 1021–1024.

Chen, J., 2021. Tracking the origin of ultralow velocity zones at the base of Earth's mantle, *Natl. Sci. Rev.,* **8,** 308.

Chen, Y., Hill, J.C., Lei, W., Lefebvre, M., Bozdag, E., Komatitsch, D. & Tromp, J., 2017. Automated time-window selection based on machine learning for full-waveform inversion, in *SEG Technical Program Expanded Abstracts 2017,* pp. 1604–1609, Society of Exploration Geophysicists.

Chen, Y., Zhang, G., Bai, M., Zu, S., Guan, Z. & Zhang, M., 2019. Automatic waveform classification and arrival picking based on convolutional neural network, *Earth Space Sci.,* **6,** 1244–1261.

Crotwell, H.P., Owens, T.J. & Ritsema, J., 1999. The TauP Toolkit: flexible seismic traveltime and ray-path utilities, *Seismol. Res. Lett.,* **70,** 154–160.

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

Debayle, E., Bodin, T., Durand, S. & Ricard, Y., 2020. Seismic evidence for partial melt below tectonic plates, *Nature,* **586,** 555–559.

Driel, M.v., Krischer, L., Stähler, S.C., Hosseini, K. & Nissen-Meyer, T., 2015. Instaseis: instant global seismograms based on a broadband waveform database, *Solid Earth,* **6,** 701–717.

Durand, S., Debayle, E., Ricard, Y., Zaroli, C. & Lambotte, S., 2017. Confirmation of a change in the global shear velocity pattern at around 1000 km depth, *Geophys. J. Int.,* **211,** 1628–1639.

Ekström, G., Nettles, M. & Dziewoński, A.M., 2012. The global CMT project 2004-2010: centroid-moment tensors for 13,017 earthquakes, *Phys. Earth planet. Inter.,* **200-201,** 1–9.

Garcia, J.A., Waszek, L., Tauzin, B. & Schmerr, N., 2021. Automatic identification of mantle seismic phases using a convolutional neural network, *Geophys. Res. Lett.,* **48,** e2020GL091658, doi:10.1029/2020GL091658.

Hosseini, K. & Sigloch, K., 2015. Multifrequency measurements of core-diffracted P waves (Pdiff) for global waveform tomography, *Geophys. J. Int.,* **203,** 506–521.

Hosseini, K. & Sigloch, K., 2017. ObspyDMT: A Python toolbox for retrieving and processing large seismological data sets, *Solid Earth,* **8**(5), 1047–1070.

Hosseini, K., Sigloch, K., Tsekhmistrenko, M., Zaheri, A., Nissen-Meyer, T. & Igel, H., 2020. Global mantle structure from multifrequency tomography using P, PP and P-diffracted waves, *Geophys. J. Int.,* **220,** 96–141.

Houser, C., Masters, G., Shearer, P. & Laske, G., 2008. Shear and compressional velocity models of the mantle from cluster analysis of long-period waveforms, *Geophys. J. Int.,* **174**(1), 195–212.

Hung, S.H., Dahlen, F.A. & Nolet, G., 2000. Frechet kernels for finite-frequency traveltimes—II. Examples, *Geophys. J. Int.,* **141,** 175–203.

Hutko, A.R., Bahavar, M., Trabant, C., Weekly, R.T., Fossen, M.v. & Ahern, T., 2017. Data products at the IRIS-DMC: growth and usage, *Seismol. Res. Lett.,* **88,** 892–903.

Jiang, W., Xi, C., Wang, W. & Ruan, Y., 2022. Time window selection of seismic signals for waveform inversion based on deep learning, *IEEE Trans. Geosci. Remote Sens.,* **60,** 1–10,

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

Komatitsch, D. & Tromp, J., 2002a. Spectral-element simulations of global seismic wave propagation—I. Validation, *Geophys. J. Int.,* **149**(2), 390–412. doi:

Komatitsch, D. & Tromp, J., 2002b. Spectral-element simulations of global seismic wave propagation—II. Three-dimensional models, oceans, rotation and self-gravitation, *Geophys. J. Int.,* **150**(1), 303–318.

Krischer, L., Megies, T., Barsch, R., Beyreuther, M., Lecocq, T., Caudron, C. & Wassermann, J., 2015. ObsPy: a bridge for seismology into the scientific Python ecosystem, *Comput. Sci. Discov.,* **8**(1), doi:10.1088/1749-4699/8/1/014003.

Lei, W. *et al.*, 2020. Global adjoint tomography—model GLAD-M25, *Geophys. J. Int.,* **223,** 1–21.

Lomax, A., Satriano, C. & Vassallo, M., 2012. Automatic picker developments and optimization: filterPicker—a robust, broadband Picker for real-time seismic monitoring and earthquake early warning, *Seismol. Res. Lett.,* **83**(3), 531–540.

Lu, C., Grand, S.P., Lai, H. & Garnero, E.J., 2019. TX2019slab: a new P and S tomography model incorporating subducting slabs, *J. geophys. Res.,* **124,** 11 549–11 567.

Maggi, A., Tape, C., Chen, M., Chao, D. & Tromp, J., 2009. An automated time-window selection algorithm for seismic tomography, *Geophys. J. Int.,* **178,** 257–281.

Menager, M., Guilhem Trilla, A. & Delouis, B., 2023. An automated earthquake detection and characterization tool for rapid earthquake and tsunami response in Western Mediterranean, *Geophys. J. Int.,* **234**(2), 998–1011.

Montagner, J.P. & Kennett, B.L.N., 1996. How to reconcile body-wave and normal-mode reference earth models, *Geophys. J. Int.,* **125,** 229–248.

Moulik, P. & Ekström, G., 2014. An anisotropic shear velocity model of the Earth's mantle using normal modes, body waves, surface waves and long-period waveforms, *Geophys. J. Int.,* **199**(3), 1713–1738.

Nissen-Meyer, T., Driel, M.v., Stähler, S.C., Hosseini, K., Hempel, S., Auer, L., Colombi, A. & Fournier, A., 2014. AxiSEM: broadband 3-D seismic wavefields in axisymmetric media, *Solid Earth,* **5,** 425–445.

Örsvuran, R., Bozdağ, E., Modrak, R., Lei, W. & Ruan, Y., 2020. Double-difference measurements in global full-waveform inversions, *Geophys. J. Int.,* **220,** 661–680.

Ritsema, J., Deuss, A., Heijst, H.J.v. & Woodhouse, J.H., 2011. S40RTS: a degree-40 shear-velocity model for the mantle from new Rayleigh wave dispersion, teleseismic traveltime and normal-mode splitting function measurements, *Geophys. J. Int.,* **184,** 1223–1236.

Ritsema, J. & Lekić, V., 2020. Heterogeneity of seismic wave velocity in Earth's mantle, *Annu. Rev. Earth planet. Sci.,* **48,** 377–401.

Tape, C., Liu, Q., Maggi, A. & Tromp, J., 2009. Adjoint tomography of the southern California crust, *Science,* **325,** 988–992.

Tape, C., Liu, Q., Maggi, A. & Tromp, J., 2010. Seismic tomography of the southern California crust based on spectral-element and adjoint methods, *Geophys. J. Int.,* **180,** 433–462.

Tarantola, A., 2005. *Inverse Problem Theory and Methods for Model Parameter Estimation. Inverse Problem Theory and Methods for Model Parameter Estimation,* Society for Industrial and Applied Mathematics.

Withers, M., Aster, R., Young, C., Beiriger, J., Moore, S. & Trujillo, J., 1998. A comparison of select trigger algorithms for automated global seismic phase and event detection, *Bull. seism. Soc. Am.,* **88,** 95–106.

Zaroli, C., Debayle, E. & Sambridge, M., 2010. Frequency-dependent effects on global S-wave traveltimes: wavefront-healing, scattering and attenuation, *Geophys. J. Int.,* **182,** 1025–1042.

Zhao, D., 2019. Importance of later phases in seismic tomography, *Phys. Earth planet. Inter.,* **296,** doi:10.1016/j.pepi.2019.106314.

Zhu, W. & Beroza, G.C., 2019. PhaseNet: a deep-neural-network-based seismic arrival-time picking method, *Geophys. J. Int.,* **216,** 261–273.