Bulletin of the Seismological Society of America, Vol. 105, No. 3, pp. 1628–1641, June 2015, doi: 10.1785/0120140267

Continuous Wavelet Decomposition Algorithms for Automatic Detection of Compressional- and Shear-Wave Arrival Times

by Petros Bogiatzis and Miaki Ishii

Abstract The determination of *P*- and *S*-wave arrival times is important for a variety of seismological applications including earthquake detection and seismic tomography. The method is based on the continuous wavelet transform of the waveforms. Unlike Fourier transform, the basis functions are localized in time and frequency, therefore, wavelet transform is suitable for analysis of nonstationary signals. For detecting the *P*-wave arrival, the wavelet coefficients are calculated using the vertical component of the seismogram. In the case of the *S*-wave arrival, we take advantage of the polarization of the shear waves, and cross examine the wavelet coefficients from two horizontal recordings. In addition, shear-wave splitting, the time delay of polarized *S* waves, can be measured using real and imaginary wavelets. Because these steps can be automated, application of the technique can easily generate a large database of shear-wave splitting measurements for studies of anisotropy.

Introduction

Many applications in seismology such as earthquake hypocenter determination (e.g., Billings *et al.*, 1994; Engdahl *et al.*, 1998; Horiuchi, 2003; Bondár and Storchak, 2011), source mechanism analysis (e.g., Hardebeck and Shearer, 2002), seismic tomography (e.g., Zhao *et al.*, 1992), and hydrocarbon reservoir imaging (Oye and Roth, 2003) require accurate picks of compressional- and shear-wave arrivals. Various seismological organizations have been relying mainly upon human experts to manually pick seismic phase arrivals of interest or review picks that have been automatically determined (e.g., Willemann and Storchak, 2001). However, the deployment of dense arrays such as the USArray (e.g., Levander and Nolet, 2013) that generates large amounts of data make manual picking and reviewing impractical. Furthermore, manually picked phases are neither errorless nor objective. Their accuracy is influenced by factors such as signal-to-noise ratio (SNR), shape of waveform onset, sample rate, filter, and misidentification. For example, Douglas *et al.* (1997) compares precision of *P*-wave picks, and estimates that the errors are 0.1 s for explosions and 0.5 s for teleseismic earthquakes. Diehl *et al.* (2009) argue that phase detection by network analysts include significant numbers of mispicks (i.e., picking of incorrect arrivals) and inconsistencies in error assessment. Zeiler and Velasco (2009) show considerable differences, both random and systematic, that exist in phase picks performed by different institutions. They estimate the average root mean square (rms) error between institutions to be 0.43 s for *P* waves and 5.53 s for *S* waves for events within regional distances (0°–30°). In particular, the average time difference for the *Sg*-wave arrival is 8.33 s. They also note a bias for the manual analyses to tend toward late picks (Zeiler and Valasco, 2009).

Although over the years a significant number of different approaches to automate phase picking have been proposed, the development of robust and accurate automatic algorithms that perform satisfactorily for various situations such as source type, distance, noise level, and instrument, remain a nontrivial, active field of research (e.g., Withers *et al.*, 1998). For *P*-wave arrival, most strategies are based upon the fact that the seismogram changes dramatically with the phase arrival. Allen (1978) has introduced the concept of the characteristic function, that is, a function that is evaluated over segments of seismogram and identifies changes that correspond to the arrival of the seismic phase of interest. Characteristic function is typically calculated after the seismogram is band-pass filtered within narrow windows around the expected signal. Such functions can be the absolute value (Allen, 1982; Baer and Kradolfer, 1987), energy and frequency (Panagiotakis *et al.*, 2008), or the envelope (Earle and Shearer, 1994). One of the first and often-used algorithms utilizes the ratio of short-term average over the long-term average, STA/LTA, as a characteristic function (e.g., Allen, 1978, 1982; Baer and Kradolfer, 1987). In addition to these changes, polarization information and coherency between *SH* and *SV* waves can be used in an algorithm for detection of *S*-wave arrivals (e.g., Amoroso *et al.*, 2012).

More recently, Persson (2003) has argued that seismic waves are non-Gaussian, and as such, higher order statistics should be used to characterize waveforms. A variety of such methods have been developed using attributes such as kurtosis and skewness (e.g., Saragiotis *et al.*, 2002; Galiana-Merino *et al.*, 2008; Küperkoch *et al.*, 2010; Lois *et al.*, 2013; Baillard *et al.*, 2014). Detectors based upon autoregressive-Akaike information criteria are another category of picking algorithms, and they are usually combined with STA/LTA analysis to obtain the initial detection of P or S onset (e.g., Maeda, 1985; Leonard and Kennett, 1999; Sleeman and van Eck, 1999; Leonard, 2000; Zhang *et al.*, 2003). Artificial neural networks are also used (e.g., Dai and MacBeth, 1995, 1997; Wang and Teng, 1995; Mousset *et al.*, 1996) which can be combined with high-order statistics, such as skewness, kurtosis, and their time derivatives (Gentili and Michelini, 2006). Taylor *et al.* (2011) used a novel approach by lifting the time series into a high-dimensional space with the time-delay embedding technique and graph theory to parameterize the resulting phase space.

Wavelet transform is another approach that is suitable for analysis of nonstationary signals. There are two types of wavelet analysis: the discrete wavelet transform (DWT) and the continuous wavelet transform (CWT). The DWT permits a compact representation of a signal, and is useful in signal compression and noise reduction (e.g., Akansu *et al.*, 2010). It has been applied for phase picking, for example, three-component seismogram analysis to detect both P and S phases using amplitude and polarization (Anant and Dowla, 1997). In contrast, the CWT is the redundant counterpart of DWT. The wavelet coefficients are correlated to one another in the CWT (this is not the case in the DWT) resulting in an improved resolution of features localized in frequency or in time. Thus, this redundancy makes the CWT more suitable for applications such as signal detection or feature extraction (e.g., Kumar and Foufoula-Georgiou, 1997). For example, Tibuleac and Herrin (1999) have applied CWT to regional earthquake recordings to detect multireflected postcritical shear waves within the crust. Similarly, Karamzadeh *et al.* (2013) combine the CWT and a characteristic function of stacked envelopes of the signal at different scales for phase detection.

In this work, we take advantage of the continuous wavelet transform properties and define characteristic functions to detect P- and S-wave arrivals. The algorithms provide onset times and estimates of the quality and the robustness of the picks. The approach presented here does not require any filtering or processing of the seismograms. In addition, the algorithms automatically adapt to the characteristics of the arrival to treat different types of signal onset behavior (e.g., sharp or emergent). The output of the algorithms also includes frequency dependent picks which can be used to assess frequency content to the phase arrival and possible body-wave dispersion. Finally, we present a variant of the method that automatically measures shear-wave splitting.

Method

This section provides a brief overview of the background theory of continuous wavelet transform and describes its application to the automatic phase-picking problem.

Continuous Wavelet Transform

Wavelet transform is a popular approach that allows unfolding of a signal into both time and frequency domains. It is used in many applications such as signal analysis, fluid mechanics (e.g., Daubechies, 1992; Farge, 1992; Hudgins *et al.*, 1993; Akansu *et al.*, 2010), image processing, communication systems, and geophysics (e.g., Lilly and Park, 1995; Kumar and Foufoula-Georgiou, 1997; Mallat, 1999; Tibuleac and Herrin, 1999; Grinsted *et al.*, 2004; Walker, 2008; Akansu *et al.*, 2010; Karamzadeh *et al.*, 2013). Similar to the Fourier analysis, it is based on the expansion of functions in terms of basis functions, but the wavelet basis functions are localized in time and frequency, allowing the locality in the signal being analyzed to be retained. For example, if a function is locally smooth, the wavelet coefficients vary smoothly over different scales with similar amplitudes. On the other hand, if the function contains a discontinuity or a singularity, only the coefficients in the vicinity of that feature have large amplitudes. This property enables better representation of nonstationary signals.

Wavelets are zero-mean functions (implemented as the "admissibility condition"), and they are characterized by their localization in time and frequency (Farge, 1992). The wavelets are not necessarily orthogonal, but various wavelet families satisfy the orthogonality condition such as "Daubechies" and "Symlets" (e.g., Daubechies, 1988, 1992; Walker, 2008). The localization of the wavelet in time and scale determines the resolution of the wavelet transform in time and frequency.

The continuous wavelet transform of a function $f(t)$, which depends upon time t, is defined as the inner product between the function and the wavelet family ψ that yields the wavelet coefficients C (Grossmann and Morlet, 1984; Daubechies, 1988, 1992; Heil and Walnut, 1989; Farge, 1992), that is,

$$C(\lambda, \tau) = \langle \psi_{\lambda\tau} | f \rangle = \int_{-\infty}^{\infty} f(t) \psi_{\lambda,\tau}^{*}(t) dt, \qquad (1)$$

in which $\psi_{\lambda,\tau}(t) \equiv \frac{1}{\sqrt{\lambda}} \psi(\frac{t-\tau}{\lambda})$ denotes the wavelet function (also known as mother wavelet), and * indicates the complex conjugate operation. The parameter λ represents the scaling factor and it is a positive real number that controls the dilation ($\lambda > 1$) or contraction ($\lambda < 1$) in time. Parameter τ controls the translation in time. A wavelet function must satisfy the admissibility condition, that is, if ψ is integrable, it has zero mean, or

$$\int_{-\infty}^{\infty} \psi(t) dt = 0. \qquad (2)$$

Equation (1) indicates that the wavelet transform can be described practically as a series of cross-correlation operations between the wavelet at different scales and the signal being analyzed. Consequently, it is desirable to have a wavelet function that reflects the features of the signal or the waveform (e.g., Torrence and Compo, 1998). Sharp changes and steps are better represented with boxcar-like wavelets, such as the Haar wavelet (e.g., Daubechies, 1988, 1992; Walker, 2008), whereas complicated waveforms should be modeled with similarly oscillatory wavelet functions. A useful property of some wavelet families that can be utilized for this purpose is the number of "vanishing moments" that dictates the complexity or oscillatory behavior of a given wavelet (e.g., Daubechies, 1988, 1992; Farge, 1992; Butzer *et al.*, 1994). To examine the behavior of higher order derivatives of a function, the mother wavelet should be insensitive to the lower order variations of the function, that is, it should satisfy the condition (Farge, 1992; Butzer *et al.*, 1994)

$$\int_{-\infty}^{\infty} \psi(t) t^m dt = 0, \tag{3}$$

in which m is the number of vanishing moments, and is the upper limit of the order of the derivate of the signal to be ignored. This implies that the wavelet coefficients are zero for signals that can be described as polynomials of degrees up to $m - 1$. Based on this concept, a wide range of wavelet families can be derived from the mother-wavelet function by changing the number of vanishing moments m. For example, using the Daubechies wavelet family, members with various vanishing moments can be generated (Fig. 1).

If the function $f(t)$ is discrete, equation (1) becomes a summation, practically a convolution operation of the signal with the scaled and normalized wavelet kernel. The wavelet functions are normalized to have unit energy at each scale, allowing direct comparison between the wavelet transforms and between transforms of different time series (e.g., Torrence and Compo, 1998).

Algorithm for the Detection of Compressional-Wave Arrivals

For the compressional-wave arrival, only the vertical component of the seismogram is considered. The arrival of the first compressional wave produces sudden change in the amplitude, the phase, and the frequency content of the recording, and the continuous wavelet transform allows the detection of these changes by decomposing the signal as a function of time and scale (Fig. 2). At the *P*-wave onset time, the CWT coefficients are expected to increase simultaneously over a range of scales, assuming that the body waves are nondispersive.

The first step in the CWT analysis is to define an analysis time window. A window centered at the theoretical *P* arrival time t_P is selected with a length that is at least the time difference between the theoretical *P* and *S* arrival times, that is, $t_S - t_P$ in which t_S is the expected *S*-wave arrival time. The next step is to determine the wavelet that is appropriate for the waveform under consideration. For example, an abrupt onset is best identified using a sharp wavelet. A rough morphological analysis attempts to identify the sharpness of the arrival using the envelope of the seismogram for two partially overlapping windows. The first window, the signal window, is defined with respect to the time of occurrence t_m of the maximum envelope amplitude. It starts at $t_m - k_1(t_S - t_P)$ and ends at t_m. The parameter k_1 is a value set by the user, so that this window includes maximum amplitude as well as the onset of the arriving wave (Fig. 3). The second window, the noise window, starts earlier at time $t_m - k_0(t_S - t_P)$ and ends at time $t_m - k_2(t_S - t_P)$, in which k_0 and k_2 are parameters set by the user with a condition $k_0 > k_1 > k_2$. This window should mostly contain noise and possibly the very beginning of the wave arrival (Fig. 3). A line is fit to the envelope of each noise and signal window using the least-squares method. For the signal window, the slope calculation includes a weighting scheme based upon the amplitude of each point of the envelope, that is, the most important point is the one that corresponds to the peak amplitude. For the noise window, all points are of equal importance. When the *P*-wave arrival is sharp, the slopes of these two lines differ significantly, but when it is emergent, the difference in slopes is small (Fig. 3). The difference in the slopes is quantified by

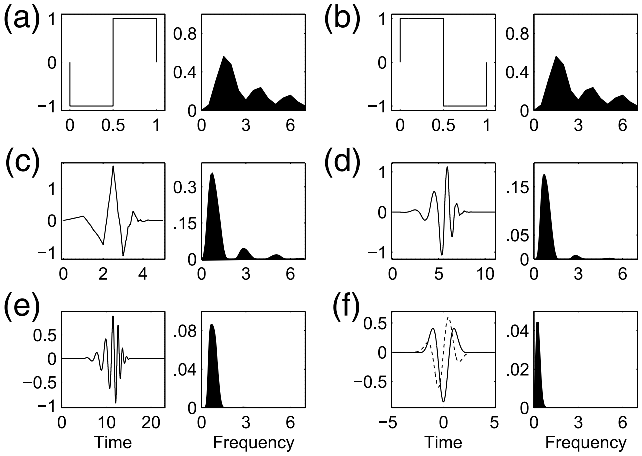

Figure 1. (Left) Mother wavelets used in this study and (Right) the corresponding amplitude spectra. (a) Symlet with one vanishing moment. (b,c,d,e) Daubechies wavelet with 1, 3, 6, and 12 vanishing moments, respectively. (f) Gaussian wavelet of order 2. The solid and dashed curves are the real and imaginary components, respectively.

$$q = (q_1 - q_0)q_1^{-1}, \tag{4}$$

in which q_0 and q_1 are the slopes of the lines of the noise and the signal window, respectively. Parameter q_1 is always positive, whereas q_0 can be negative or positive, depending upon the type of the arrival and the noise characteristics of the seismogram. High q values are associated with abrupt arrivals and low q values indicate emergent arrivals. Based upon the value of q and the SNR, a wavelet (e.g., different vanishing moments) is selected for the analysis of the signal (e.g., Table 1).

After the wavelet for the analysis is selected, the CWT of the waveform is calculated using different scales and time

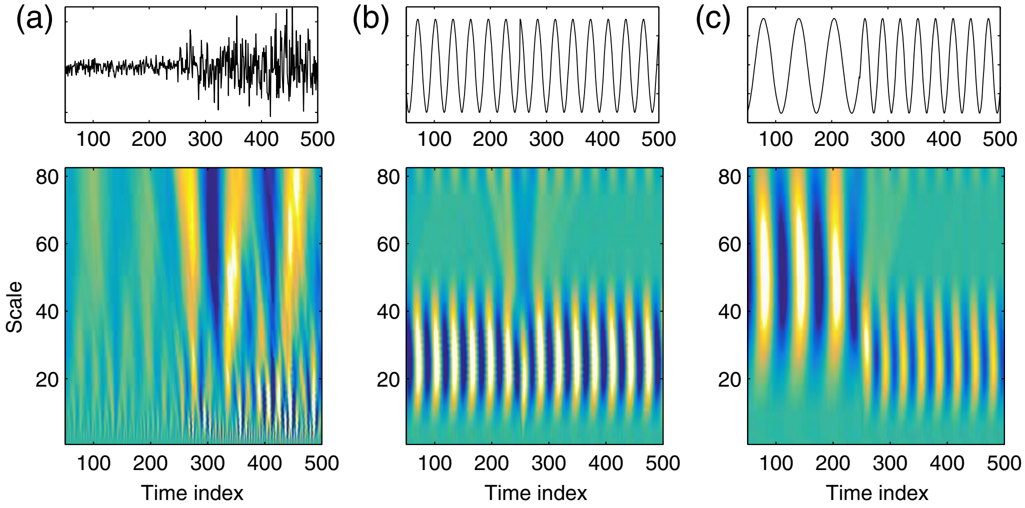

Figure 2. Synthetic examples of the behavior of the continuous wavelet transform to changes in (a) amplitude, (b) phase, and (c) frequency. Top panels show the synthetic data in time domain and bottom panels show wavelet spectrum obtained using Daubechies wavelet with three vanishing moments (Fig. 1c). (a) White Gaussian noise that is amplified by a factor of 4 beyond time index 250. (b) Harmonic oscillation that exhibits a phase discontinuity at time index 250. (c) Harmonic oscillation that changes frequency at time index 250. The color version of this figure is available only in the electronic edition.

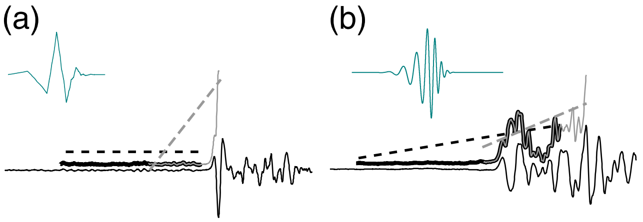

Figure 3. Schematic representation of the analysis of the arrival type that is performed to select the appropriate wavelet. Examples of (a) sharp and (b) gradual arrivals. For each example, raw (thin black curve), its envelope function for noise (thick black curve) and signal (thin gray curve) windows, and corresponding slopes obtained using the noise (black dashed line) and signal (gray dashed line) windows are displayed. Different types of wavelets (upper left corner) are selected based upon the slopes. The color version of this figure is available only in the electronic edition.

offsets. The CWT coefficients are both positive and negative, but the important information, for the purpose of this algorithm, is the abrupt change in the amplitude. To enhance the changes in amplitudes, 2D range filter in scale-time space is defined and applied to the CWT coefficients. At a given time and scale, the filtered value is calculated as the difference between the maximum and minimum CWT coefficients within the range window. The length of the time window is set slightly longer than the expected *P*-wave period; a long window enhances long-period signal and vice versa. This operation implicitly introduces smoothing in time and scale, filtering out potentially isolated features and increasing the statistical significance of the peaks (Torrence and Compo, 1998). Furthermore, to focus on the detection of the onset of a wave arrival, the range filter window is defined asymmetrically in time, that is, the filter uses more coefficients of later times than the prior times (e.g., Table 1). This approach helps to capture the increase in the amplitude of the coefficients that occurs immediately after the onset of the *P*-wave arrival. The coefficients are then normalized with the maximum value at each scale. After the application of the range filter, the following quantity is calculated at each scale

$$M_i = \mathrm{sign}(c_i^+ - c_i^-)(c_i^+ - c_i^-)^2, \tag{5}$$

with

$$c_i^- = \frac{1}{n}\sum_{j=i-n}^{i} C_j \quad \text{and} \quad c_i^+ = \frac{1}{n}\sum_{j=i+1}^{i+n} C_j, \tag{6}$$

in which C_j is the filtered wavelet coefficient at time j, i is the target time index, and n is the predetermined window length. The parameter M_i gives a measure of the difference between average filtered CWT coefficients before (c_i^-) and after (c_i^+) the ith time, and keeps track of the relative amplitudes between c_i^- and c_i^+. Large positive and negative values of M_i indicate abrupt changes in the CWT coefficients with amplitude increase and decrease, respectively, and small values of M_i imply no significant change in the amplitude. Ideally, M_i reaches a maximum value at the *P*-wave onset time. At each scale, therefore, the onset time is chosen as the time when a maximum in M_i occurs for the first time with a value that is above a given threshold. This threshold is set to be a fraction of the peak M_i value for each scale. This fraction is determined using the filtered CWT coefficients, $1 - (\overline{C_s} - \overline{C_n})/\overline{C_s}$, in which $\overline{C_s}$ and $\overline{C_n}$ are the averages of the CWT coefficients over the time window for which M_i values are obtained, and for a noise window prior to M_i calculation, respectively. This parameter controls how close to

Table 1
Summary of Parameters Chosen for *P*-Picking Algorithm

P Wave					
q	Any	> 0.95	(0.8, 0.95]	(0.5, 0.8]	≤0.5
SNR (dB)	> 34	> 10	≤10	≤34	
Wavelet	db1		db3	db6	db12
Range filter window length along scales	1		3		
Range filter window length along time (pts)	3	11		21	1
Detection threshold (upper/lower)	0.01 0.25		0.01 0.1		

The type of wavelet and the range filter applied on the wavelet coefficients are determined based upon the value of q (equation 4) and signal-to-noise ratio (SNR). The wavelets are all Daubechies wavelets with varying number of vanishing moments, for example, db3 indicates Daubechies wavelet with three vanishing moments. The continuous wavelet transform is performed with 40 wavelets ranging in scale between 2 and 128.

the peak M_i time the picked time should be; for a noiseless case, that is, when the fraction is zero, a pick is selected to be the first maximum in M_i value, and for a noisy case, that is, when the fraction is large, a pick is selected close to the peak M_i time. Bounds on the fraction value can be specified by the user to ensure that it is neither 0 nor 1. This procedure is repeated for all scales, and the final *P*-wave onset time is computed as a weighted average of the individual picks from different scales. The weighting value is determined by the product of the SNR and the peak coefficient amplitude around the *P*-wave arrival at each scale. This averaging scheme exploits the multiscale resolution of the CWT domain, and emphasizes the frequencies for which the SNR is high.

The statistics of the picks obtained at different scales provide a measure of the robustness of the final pick and associated errors. For example, the minimum and maximum picked times provide the extreme error bounds. Another useful measure is the standard deviation of these picks. In this work, we use the weighted unbiased standard deviation as an estimate of the uncertainty of the arrival time. Nevertheless, higher order statistics can be used to further assess the uncertainty. For example, the skewness of the picks can provide insight into the symmetry of the pick distribution around the final picked time.

Algorithm for the Detection of Shear-Wave Arrivals

Shear-wave arrival detection is, in general, more difficult than the detection of the *P*-wave arrival, because the *S* waves arrive within the coda of *P* and other phases. This problem becomes severe in the case of emerging *S*-wave arrivals in which the wave onset is hidden in the background noise, resulting in picks that are often late. One advantage of the *S*-wave analysis, on the other hand, is that the *S*-wave motion is 2D; hence it allows simultaneous analysis of the two horizontal seismograms to produce robust detection of the *S*-wave arrival. The first stage of the *S*-wave detection algorithm is similar to the *P*-wave counterpart. A window is selected around the *S*-wave theoretical arrival time with the window length that is at least the theoretical time difference between the *P*- and *S*-wave arrivals. For reasons that are discussed later, the adaptive wavelet selection is skipped, and one fixed wavelet is used. The CWT is calculated for both horizontal recordings.

At the *S*-wave onset time, the coefficients of the CWT are expected to increase in absolute values over a wide range of scales assuming that the wave is nondispersive. In contrast to the *P*-wave arrival, this change should occur coherently on the two horizontal seismograms. To identify this localized similarities in time and scale, the wavelet cross spectrum $F(\lambda, \tau)$, is calculated as (Hudgins *et al.*, 1993; Torrence and Compo, 1998)

$$F(\lambda, \tau) = L(C_1^*(\lambda, \tau)C_2(\lambda, \tau)), \tag{7}$$

in which C_1 and C_2 are the CWT of the two signals that depend upon scale λ and time τ, whereas L is a smoothing operator in time and scale. Some of the advantages of the cross examination of waveforms with wavelet cross spectrum are demonstrated in Figure 4. In this synthetic example, Ricker wavelet (i.e., the negative normalized second derivative of a Gaussian function) is used to represent the *S*-wave arrival embedded in white Gaussian noise. The input *S* waves include a phase difference of π radians on two horizontal directions. The wavelet cross spectrum calculated using Symlet-1 wavelet (Daubechies, 1992) is sparse with zero coefficients everywhere except for the time and frequency range of the synthetic *S*-wave arrival that is clearly represented as a region of high anticorrelation value (Fig. 4).

After the cross spectrum is calculated, a range filter is applied in a similar manner as with the *P*-wave algorithm. For the detection of the *S*-wave arrivals, however, the sign of the cross-spectrum coefficients is not important. Therefore, the range filter is designed to bring out regions of high cross-correlation amplitudes (either positive or negative). The filtered coefficients are normalized with the maximum amplitude, and then a characteristic function is examined at different scales. This characteristic function monitors the behavior of the average values of cross spectrum before and after a given time, that is, for a given scale,

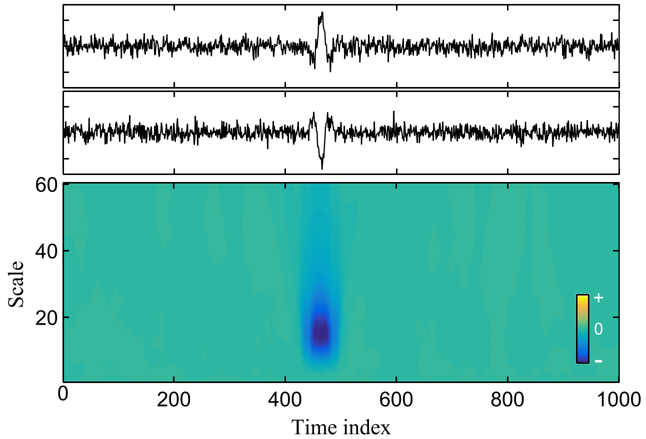

Figure 4. Synthetic example of seismogram wavelet cross spectrum. Ricker wavelet is used to represent the *S*-wave arrival with an average phase difference of π between the two horizontal components (top and middle rows). The wavelet cross spectrum (bottom row) shows a sharp region of negative coefficients, mostly between scales 10 and 20. The color version of this figure is available only in the electronic edition.

$$N_i = f_i^+ - f_i^-, \tag{8}$$

with

$$f_i^- = \frac{1}{n}\sum_{j=i-n}^{i} F_j \quad \text{and} \quad f_i^+ = \frac{1}{n}\sum_{j=i+1}^{i+n} F_j, \tag{9}$$

in which F_j is the filtered cross-spectrum value at time index j, and the index i corresponds to the evaluation time. In the case of *S* waves, the difference in the values between two windows is used instead of the square of the difference considered for the *P*-wave arrival (equation 5). Similar to the *P*-wave algorithm, one pick from each scale is obtained using the first local maximum of N_i that exceeds a given threshold. The threshold is determined using the same average cross-spectrum coefficients as for the *P*-wave case. Assuming the *S* waves to be nondispersive, the final *S*-wave onset time is calculated as a weighted average of these individual estimates. The weights in this case are defined by the difference of the average cross-spectrum coefficient before and during the *S*-wave arrival. The statistics of the picks at various scales can be used to estimate the uncertainty of the final pick, and the weighted unbiased standard deviation is used.

Data, Algorithm Setup, and Applications

We apply the automatic picking algorithm to data recorded by the High Sensitivity Seismograph Network (Hi-net) of Japan (Fig. 5; e.g., Okada *et al.*, 2004). This network consists of more than 750 stations distributed throughout the Japanese islands, and extremely short-period instruments (channel code EH) are installed inside boreholes at depths that are typically more than 100 m below ground. The waveform is digitized at 1000 samples per second, and the data are made available with sampling frequency of 100 Hz and 27 bit resolution (Okada *et al.*, 2004).

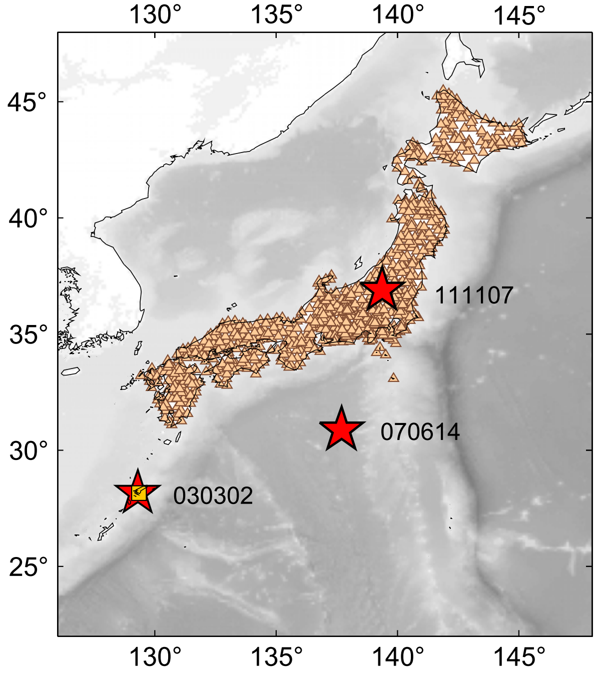

Figure 5. Distribution of 775 High Sensitivity Seismograph Network (Hi-net) stations (triangles) that are available as of 1 July 2014. The two earthquakes used in the arrival-time picking examples are indicated by the two stars labeled 070614 (deep event on 14 June 2007) and 111107 (shallow event on 07 November 2011). The earthquake used in the splitting example is indicated by a star with label 030302 (depth of 127 km on 02 March 2003), and the F-net station AMM is indicated by a rectangle. The color version of this figure is available only in the electronic edition.

We select parameters and apply the automatic algorithms to the Hi-net recordings of a deep, regional earthquake. A deep event is chosen, because surface waves that arrive close to the *S*-wave arrival times at regional distances are not excited, and the arrival times of *P* and *S* waves are sufficiently separated. Furthermore, the Hi-net data exhibit a wide range of disparate behaviors due to strong heterogeneities (e.g., subducting slab) beneath Japan. Even from a single deep event, we observe traces that contain strong high-frequency signals with emergent onset and those that are mostly of low frequency with sharp onset (Figs. 6 and 7). This data set, therefore, is a good example to demonstrate the effectiveness of the algorithms.

Once the target region/stations and types of earthquakes are determined, various parameters must be chosen for the automatic algorithms. Specifically, a user must select a data signal window, fractional parameters k_0, k_1, and k_2 that are used to define windows for the envelope slope calculation, the smoothing operator L for shear-wave detection, the type of wavelets to be used, and how they are selected based upon the data, the 2D range filter, the window for the M_i and N_i calculations, and the thresholds for detecting the onset time.

For the *P*-wave analysis, the fractional values k_0, k_1, and k_2 chosen are 0.09, 0.04, and 0.01, respectively, for the

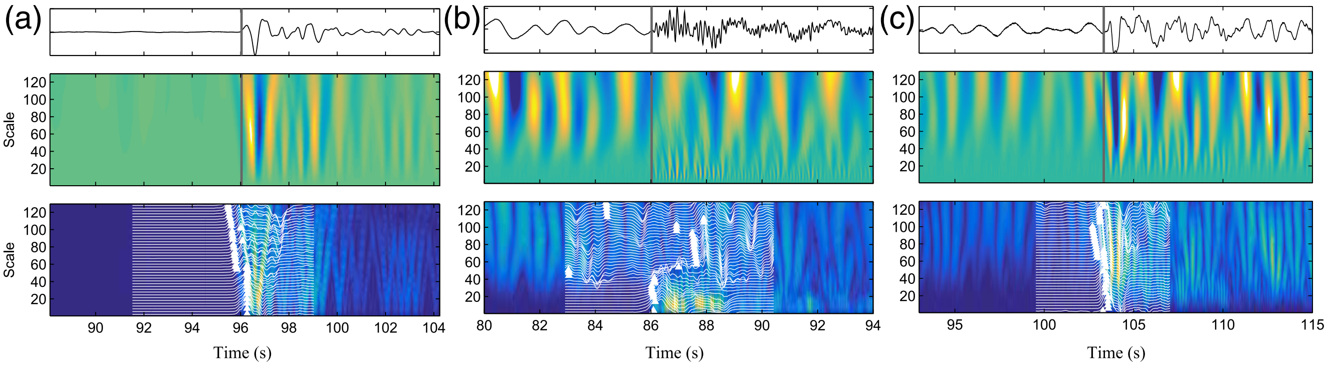

Figure 6. Three examples of P-wave arrival picks at (a) IKNH, (b) KTUH, and (c) SNWH stations of the Hi-net array. The top row shows the vertical component of the seismogram centered at the theoretical P-wave arrival time. The middle row shows the continuous wavelet transform (CWT) coefficients, and the bottom row shows the same coefficients after the application of the range filter. White curves represent the detection function, normalized for each scale. Triangles mark the pick times at each scale. The final P-onset time is shown with the vertical line. The wavelets used are Daubechies with (a) one and (b,c) three vanishing moments. The color version of this figure is available only in the electronic edition.

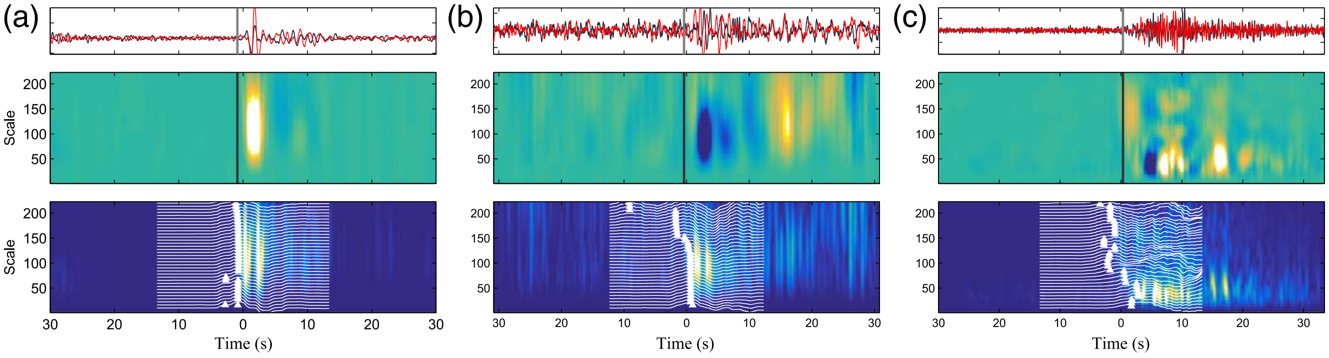

Figure 7. Three examples of S-wave arrival picks at (a) AYKH, (b) HMMH, and (c) ASGH stations of the Hi-net array. The top row shows the two horizontal seismograms centered on the theoretical S-wave arrival time. The middle row shows the corresponding wavelet cross spectrum, and the bottom row shows the filtered version of the cross spectrum. White curves represent the detection function, normalized for each scale. Triangles mark the pick times at each scale. The final S-wave onset time is shown with the vertical line. The color version of this figure is available only in the electronic edition.

determination of the sharpness of the waveform onset. The q value that quantifies the sharpness is used in conjunction with the SNR to choose the number of vanishing moments between 1 and 12 (Table 1). Seismograms with small values of q (emergent arrivals) are modeled with high vanishing moments (more oscillatory wavelet), and those with high values of q (sharp arrivals) are decomposed using wavelets with low vanishing moments (more step-like wavelets). The wavelet family chosen for the P-wave analysis is the Daubechies wavelet family (Fig. 1b–e). The Daubechies are used in a broad range of applications as they are orthogonal, compact, and easily specified with vanishing moments (e.g., Daubechies, 1988, 1992; Mallat, 1999; Vonesch *et al.*, 2007; Walker, 2008). The 2D range filter that is applied to coefficients of CWT is defined with variable lengths in the time and scale domains. The length is set between 1 and 21 samples in time based upon the value of q and the SNR (Table 1). The scale dimension of the range filter is set to three except when the SNR is high. When the ratio is above 34 dB, the scale dimension is ignored, and filtering is performed only over time. The time window for evaluation of M_i (equation 5) starts at the smaller of 3.5 or $0.4(t_S - t_P)$ seconds before the time of interest, and ends at the smaller of 4.0 or $0.4(t_S - t_P)$ seconds after the time of interest. The scale-dependent threshold for obtaining picks at each scale is bound to be no smaller than 0.01 of the peak M_i value for all traces. The upper limit for the fraction is set to be 0.25 for traces with SNR below 34 dB and 0.01 for traces with SNR equal to or greater than 34 dB (Table 1).

For the S-wave analysis, the signal time window is selected to be 50% of the time between S and P arrivals, that is, $0.5(t_S - t_P)$. We find that wavelets that are well localized in time and have a small number of vanishing moments yield sharp changes in the wavelet cross spectra, and that wavelet type does not alter the detection ability, that is, a single wavelet is adequate for most cases. Therefore, the Symlet-1 wavelet (Fig. 1a; Daubechies, 1992) is chosen for the S-wave algorithm. It is poorly localized in the frequency domain; hence, the resolution in scale is poor. Consequently, the scale

Table 2
Summary of Parameters Chosen for *S*-Picking Algorithm

	S Wave
Smoothing window length along scales	1
Smoothing window length along time (s)	3% of $t_S - t_P$
Range filter window length along scales	1
Range filter window length along time (s)	3% of $t_S - t_P$
Wavelet	sym1

The Symlet wavelet with one vanishing moment (sym1) is used for the *S*-wave analysis. The continuous wavelet transform is performed with 36 wavelets ranging in scale between 4 and 220.

dimension is not considered in the smoothing operators on the cross-wavelet coefficients (equation 7) or the range filter. The extent of the smoothing operator (equation 7) in time is set to $0.03(t_S - t_P)$ seconds (Table 2). The window for the N_i calculation (equation 8) is centered on the *S*-wave theoretical arrival time and spans 40% of the $t_S - t_P$ time. The detection threshold is bound to be between 0.1 and 0.99 of the peak N_i amplitude at each scale.

Examples

As an example, we choose a magnitude 5.1 earthquake that occurred on 14 June 2007 at 15:03:14 UTC. According to the Japan Meteorological Agency (JMA), this event is located at latitude 30.876° N, longitude 137.687° E, and a depth of 507.2 km (see Data and Resources). It is recorded by most of the stations of the Hi-net array. When there is a clear arrival with high SNR such as at the station IKNH (Fig. 6a), the picks show minimum amount of scatter at different scales. For the example at station KTUH (Fig. 6b), the higher frequency *P* wave is visible at scales smaller than 40, but scales above 60 are dominated by noise. Such low-frequency noise that is of comparable amplitude to the high-frequency *P*-wave signal is clearly separated by the multiresolution characteristics of the wavelet domain. Individual picks up to a scale of about 40 are consistent, associated with the *P*-wave arrival. At larger scales, the energy of the *P* arrival is small, and the picks show scatter, reflecting the background noise. These picks associated with noise do not contribute significantly to the final onset-time estimate, because the SNRs at these scales are low, and the individual picks are weighted down in the final average calculation. Lastly, when a waveform arrives with low SNR, it becomes difficult to manually pick the onset (Fig. 6c). However, for scales with sufficient SNR, an appropriate final pick can be obtained. The three examples shown in Fig. 6a–c give *P*-wave arrival-time estimates of 96.04 ± 0.28, 86.03 ± 0.77, and 103.33 ± 0.24 s, respectively.

The application of the *S*-wave detection algorithm shows similar results (Fig. 7). When there is a clear *S*-wave arrival, a region of high positive cross-spectrum coefficients is apparent (Fig. 7a). The picks at various scales are relatively consistent with one another, and the final pick corresponds well to visually observable onset of the *S*-wave arrival. Even when the SNR is poor, cross-spectrum coefficients remain small prior to the *S*-wave arrival and increases sharply when the wave arrives (Fig. 7b). The individual picks at different scales are also consistent, with the exceptions at a few larger scales that are noise-related. Because both the amplitude of the coefficients and the SNR are small at these large scales, small weights are assigned to these picks, and they do not have much effect on the final arrival-time estimate. Lastly, when an arrival of the *S* wave is emerging, the onset time is difficult to identify from visual examination of the seismograms (Fig. 7c). However, in the wavelet cross-spectrum domain, the *S*-wave arrival produces a sharp initiation of a region of nonzero coefficients. The sign of the cross-correlation amplitude varies along scales, and individual pick times show scatter, but the onset time can be estimated objectively and automatically. The *S*-wave arrival times selected for stations AYKH (Fig. 7a), HMMH (Fig. 7b), and ASGH (Fig. 7c) are 150.58 ± 0.43, 139.05 ± 2.18 and 151.46 ± 2.10 s, respectively. The uncertainty in AYKH station (Fig. 7a) is particularly small, reflecting the sharp and clear arrival of the *S* phase. At the other two stations, the estimated uncertainties are considerably larger. At the HMMH station (Fig. 7b), the high uncertainty arises mainly from two picks at large scales that are significantly earlier than the rest. Although these outliers have small influence on the final pick, they strongly affect the uncertainty estimate, because calculation of weighted unbiased standard deviation involves squared distances from the average, that is, large deviations are weighted heavily. At the ASGH station (Fig. 7c), the picks at different scales are scattered, reflecting the ambiguity of the onset time due to an emergent arrival and resulting in a large uncertainty estimate.

To assess the effectiveness of the automatic algorithms, we compare the picks with those that are made manually. Of the 757 stations that are available for this deep earthquake, some noisy records and arrivals are difficult to identify, reducing the manually picked times to 750 and 679 for *P*- and *S*-wave arrivals, respectively. Seventy percent of the automatic *P*-wave picks are within 0.28 s of the manual picks, and 85% are within 0.54 s (Fig. 8a). Assuming Gaussian distribution of the differences, the mean is −0.06 s and the standard deviation is 0.53 s (Fig. 9a). The data are, however, fit poorly with this Gaussian distribution. In particular, the peak of distribution around 0 s is significantly underpredicted by the Gaussian model. A better fit is obtained using the Student's *t* distribution that reproduces a narrow peak with broad tails (Fig. 9a). The parameters of the Student's *t* distribution are mean of 0.00 s, standard deviation of 0.19 s, and 1.45 degrees of freedom. For *S* waves, 70% of the automatic picks are within 1.02 s of handpicked times, and 85% are within 1.66 s (Fig. 8b). The mean and standard deviation of the data, assuming Gaussian distribution, are 0.01 and 1.71 s, respectively. However, as with the *P*-wave data, the distribution is better modeled with the Student's *t* probability density function with mean of 0.12 s, standard deviation of 0.12 s, and 1.62 degrees of freedom. The Student's model still underestimates the peak amplitude of the *S*-wave pick distribution,

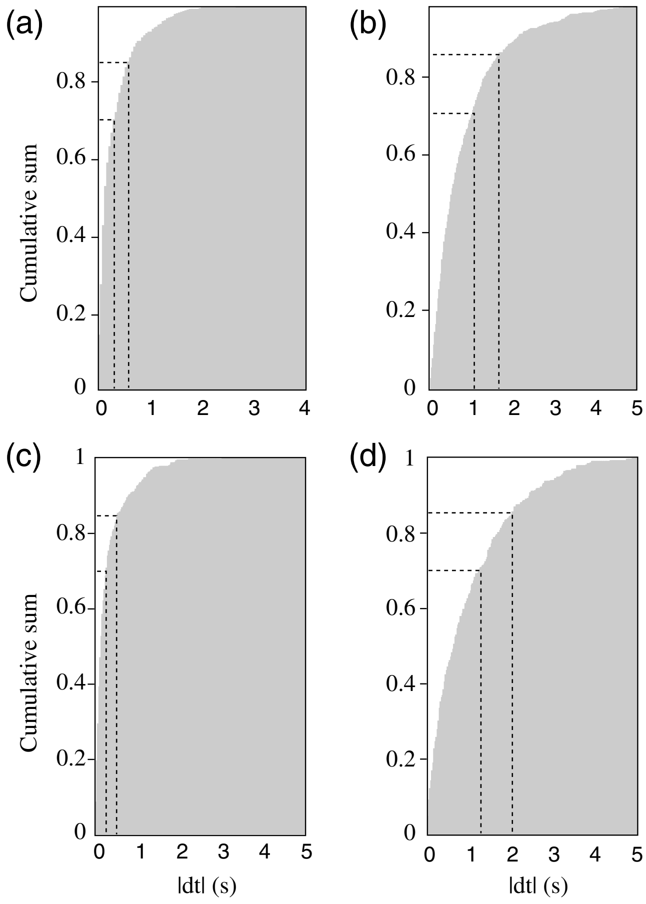

Figure 8. Normalized cumulative number of the stations as a function of absolute difference between automatic and manual picks ($|dt|$) for (a) *P*-wave and (b) *S*-wave arrivals from the analysis of the deep earthquake of 14 June 2007. (c) and (d) are the same as (a) and (b), respectively, except for the analysis of the shallow earthquake of 07 November 2011. The dashed curves show the level of residuals at 70% and 85%.

but characterizes the distribution much better than the Gaussian (Fig. 9b). These statistics show good agreement between manual and automatic picks. The rms differences are 0.53 and 1.71 s for *P* and *S* picks, respectively. The rms differences in picks between institutions are 0.43 s for *P*-wave picks and 5.53 s for *S*-wave picks (Zeiler and Valasco, 2009). Hence, our approach produces *S*-wave picks that show substantially better agreement with manual picks than between institutions. On the other hand, the differences observed for the *P*-wave picks, is slightly larger than interinstitutional difference.

To test if the parameters of the algorithms set for a deep earthquake can be used for other sources, we apply the scheme to a shallow earthquake without changing any of the preset values (Table 1). The event chosen is a magnitude 4.5 earthquake that occurred on 07 November 2011 at 14:41:59 UTC. The JMA catalog lists its location as latitude 36.900° N, longitude 139.381° E, and a depth of 3.4 km. Data for this event are available from 769 stations of the Hi-net array, and the automatic algorithms are applied and compared with manual picks. Seventy percent of the automatic *P*-wave picks are within 0.25 s of hand-selected times, and 85% are within 0.53 s (Fig. 8c). Gaussian fit to the distribution of differences between manual and automatic picks provides mean and standard deviation of −0.13 and 0.49 s, respectively (Fig. 9c). As with the previous example, the distribution is fit better by the Student's *t* distribution with mean of 0.01 s, standard deviation of 0.16 s, and 1.31 degrees of freedom (Fig. 9c). These statistical results indicate that the performance is comparable to the previous example using a deep source. For *S*-wave picks, 70% of the automatic picks are within 1.23 s, and 85% are within 2.01 s from manual picks (Fig. 8d). Gaussian and Student's *t* distribution give the same mean of −0.36 s, but standard deviations are 1.36 and 0.83 s, respectively. The degrees of freedom for the Student's *t* distribution are 2.63 (Fig. 9d). These results show slightly larger discrepancies between automatic and manual picks compared to the results using data from the deep earthquake. This can be attributed to simultaneous arrival of different phases, in particular, surface waves that produce coherent signal on the two horizontal components of the seismogram. Similarly, the incidence angles of the compressional waves and their coda generated by such a shallow event are not vertical, resulting in coherent signal on two horizontal components, further

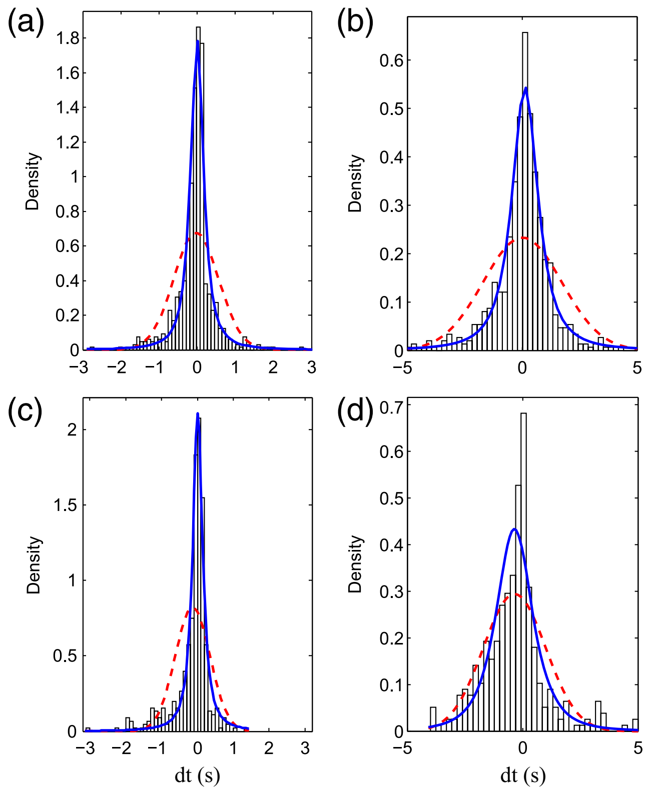

Figure 9. Probability density histograms of the difference between automatic and manual picks and corresponding fits using normal (dashed curve) and Student's *t* (solid curve) distributions for the (a) *P*- and (b) *S*-wave analyses of the deep earthquake. (c) and (d) are the same as (a) and (b), respectively, except for the analyses of the shallow earthquake. The color version of this figure is available only in the electronic edition.

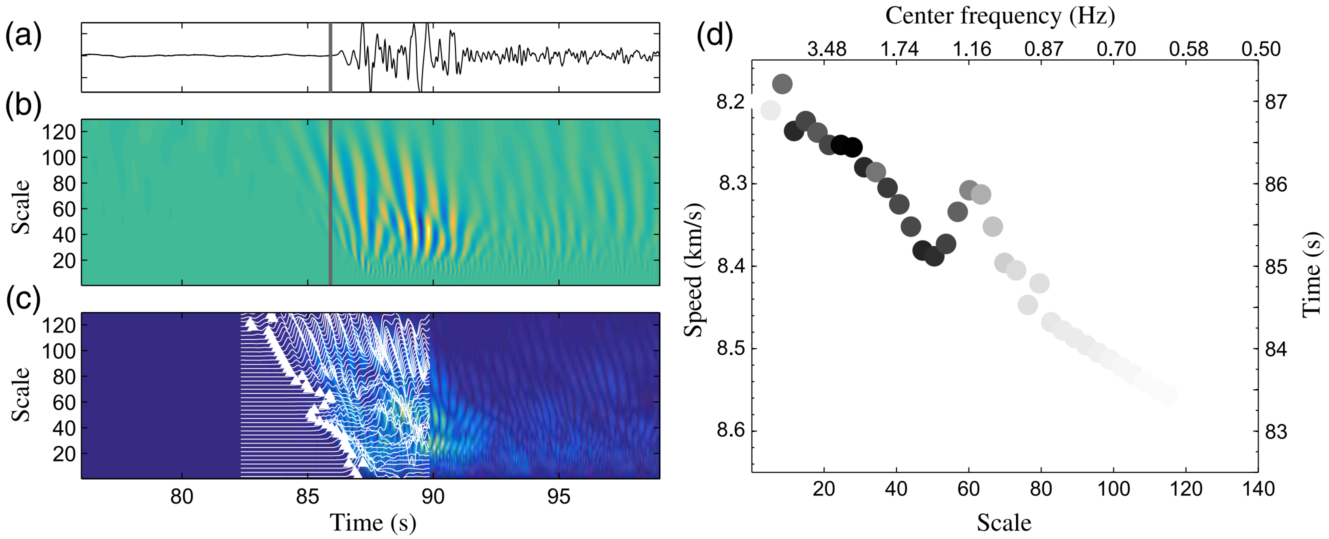

Figure 10. Well-dispersed *P*-wave signal and its dispersion characteristics. (a) Vertical component recorded at YMKH station of the Hi-net array from the deep earthquake of 14 June 2007. The time window is selected around the expected *P*-wave arrival time. (b) CWT of the seismogram. (c) Filtered coefficients. White curves show the variation of the detection function normalized for each scale. White triangles mark the pick at each scale. The *P*-onset time (vertical line) is calculated as the weighted average of the individual picks. (d) Dispersion diagram showing the pick time and the corresponding apparent speed as a function of scale and the associated center frequency. The intensity of the circles shows the importance of each pick that is, dark circles are associated with larger weights used to calculate the arrival time. The color version of this figure is available only in the electronic edition.

interfering with the determination of *S*-wave arrival time. Nonetheless, the discrepancy between manual and automatic picks is small and suggests that the set of parameters chosen for deep sources can be applied successfully to shallow sources.

Thus far, we have assumed *P* and *S* waves are nondispersive to obtain a single arrival time. Although body waves are generally not dispersive, there have been observations of considerable dispersion (e.g., Correig, 1991; Devilee *et al.*, 2003; Furumura and Kennett, 2005). The algorithms presented in this manuscript are suitable for retrieving dispersion information, or frequency dependence of arrival times, by keeping track of individual picks along different scales (Fig. 10). However, it should be noted that the wavelet transform of the signal unavoidably introduces artifacts at the edges of the wavelet spectrum due to the finite signal length and incomplete localization of the wavelet in time. The regime of significant artifacts is estimated using the "cone of influence" (Torrence and Compo, 1998; Grinsted *et al.*, 2004) that shows the decorrelation time for a single spike in the signal at each scale. Any scale-dependent signal that is beyond this cone of influence can be interpreted as dispersion effect, and the artifact can be taken into account when evaluating the statistical significance of the measured dispersion. This artifact can also be suppressed using longer time window around the arrival of the signal of interest. For measurements of dispersion, selection of the wavelet for wavelet transform also becomes important. For example, the Symlet-1 wavelet used for *S*-wave detection is poorly localized in frequency and thus has significant leakage over scales, making dispersion measurement more challenging.

Measurement of Shear-Wave Splitting

Shear-wave splitting can be measured using well-established techniques such as rotation correlation method (e.g., Ando *et al.*, 1983; Fukao, 1984; Bowman and Ando, 1987; Levin *et al.*, 1999). However it is also possible, by selecting a complex wavelet, to obtain a direct measurement of the phase difference between the two horizontal components from the complex argument of the wavelet cross spectrum.

The algorithm to measure shear-wave splitting is practically identical to that presented in the previous section, except that the cross-wavelet spectrum (equation 7) becomes complex when the mother wavelet chosen for the analysis is complex. In this case, the cross-wavelet power is given by $|F(\lambda, t)|$, and the relative phase difference of the signals in time and scale is obtained using the complex argument $\arg(F(\lambda, t))$. The amplitude information of the cross-wavelet power spectrum can be used for automatic selection of the region of interest that is, time and scale at which the *S*-wave arrives. The average phase of the wavelet cross spectrum within this region provides an estimate of the phase difference between the two horizontal recordings. Conversion of phase difference to time delay, a quantity conventionally used as a measure of shear-wave splitting, requires an estimate of the period of the *S* wave which can be readily obtained by the type of basis wavelet and the scale of the peak amplitude in the power of the wavelet cross spectrum. If T is the estimated period of the wave, the delay time is calculated as $\delta t = \delta\varphi \times T/(2\pi)$, in which δt is the delay time and $\delta\varphi$ is the measured phase difference. We demonstrate this approach in the following two examples. For the basis complex

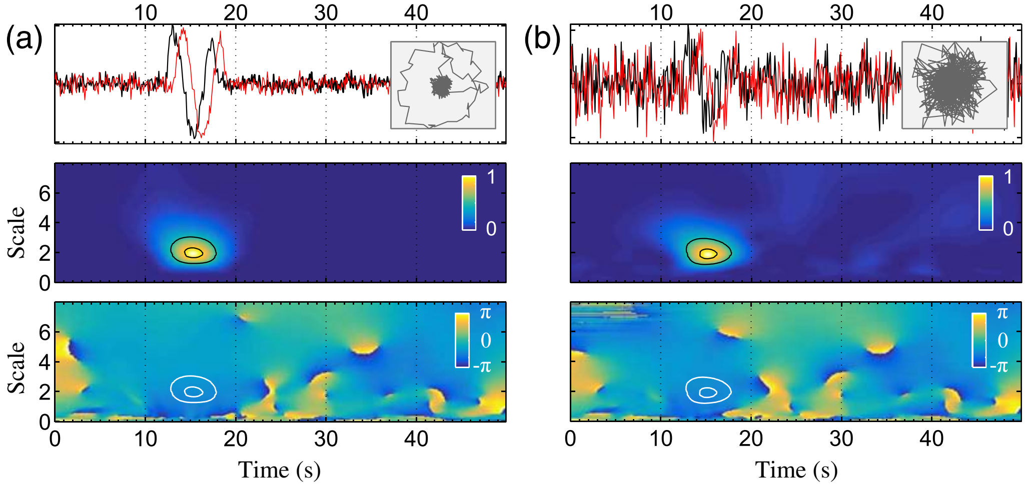

Figure 11. Synthetic examples for determining the time delay between two horizontal seismograms. The top row shows the synthetic waveforms and the corresponding particle motion (inset). The middle row shows the amplitude spectrum normalized by the maximum value, and the bottom row displays the complex argument of the wavelet cross-spectrum coefficients. The contours correspond to 50% and 90% of the peak amplitude of the coefficients. Two different noise levels are considered, (a) 7.3 dB and (b) 0.6 dB. The color version of this figure is available only in the electronic edition.

wavelet, Gaussian wavelet of order 2, that is, the second derivative of the complex Gaussian function, is used (Fig. 1f).

As a first example, we generate two synthetic seismograms using a tapered harmonic function of period $T = 4.6$ s with 1 s time delay, that is, one trace is offset from another by one second (Fig. 11). White Gaussian noise is added to both signals with SNRs of 7.3 dB (low-noise case) and 0.6 dB (high-noise case; Fig. 11a,b). In both cases, the arrival of the *S* wave is clearly identified in the amplitude spectrum, and a region of interest can be defined based upon 50% and 90% of the peak amplitude. In the example where the noise level is low (Fig. 11a), the phase difference is calculated by taking the average within these regions, and based upon the 50% and 90% of the peak amplitude, they are −62.4° and −63.9°, respectively. Using the scale where the peak amplitude in the amplitude spectrum occurs, the period is estimated to be 4.785 s. If we use the 90% contour, the time delay is 1.01 s that reduces to 0.97 s when the contour level is lowered to 50%. Even when the noise level is increased so that the target signal is difficult to discern by eye (Fig. 11b), the wavelet analysis performs well, and recovers delay times of 1.07 and 1.05 s using 90% and 50% contour levels, respectively. These analyses show that the delay time is successfully recovered by the automatic algorithm, and when noise is not correlated on the two components, shear-wave splitting can be obtained even using noisy recordings.

As a second example, we apply the method to real data recorded at the AMM station that is part of the Full Range Seismograph Network of Japan (Fig. 5; Okada *et al.*, 2004). We choose a magnitude 4.2 earthquake that occurred on 02 March 2003 at 20:01:00 UTC with the hypocenter at 27.695° N, 127.374° E, and 127.0 km depth (JMA catalog, see Data and Resources). The station AMM in the Ryuku region is within 1° of the earthquake, and the *S*-wave incident angle is nearly vertical, maximizing energy on the two horizontal components. Furthermore, previous studies have analyzed anisotropic signals observed at this station (e.g., Long and van der Hilst, 2005), providing reference values to which the results of the automatic algorithm can be compared.

The two horizontal seismograms recorded at the AMM station exhibit clear *S*-wave arrival with time delay (Fig. 12a). The region of interest can be easily identified in the amplitude spectrum, and delay time is determined using the phase information. However, shear-wave splitting measurements are typically characterized by the fast-splitting direction and the delay time, rather than delay time alone (e.g., Long and Silver, 2009). To obtain an estimate of the fast-splitting direction, we implement an additional step. Assuming a uniform anisotropy, the waveforms recorded on two horizontal components become most similar when they are rotated to the fast-splitting direction, that is, at this angle, the two waveforms have highest correlation with the delay-time offset. Therefore, we perform a grid search over rotation angles in the horizontal plane. For each rotation angle, the delay time is first determined, and the waveforms are shifted by this delay time to determine maximum value of the amplitude cross spectrum. The rotation angle that gives the largest amplitude, corresponding to most similar waveforms in two horizontal directions, can be interpreted as the fast-splitting direction.

This additional step is demonstrated in Figures 12 and 13. At a given angle (e.g., 48°), the wavelet analysis of the two horizontal recordings is performed to obtain the delay time (e.g., 0.68 s; Fig. 12a). The waveforms are then shifted by this delay time, and the wavelet analysis is performed again (Fig. 12b). Because correction has been made for the delay time, the resulting phase difference should be zero, and the quantity of interest is now the maximum value of the amplitude spectrum. This is repeated for different rotation angles

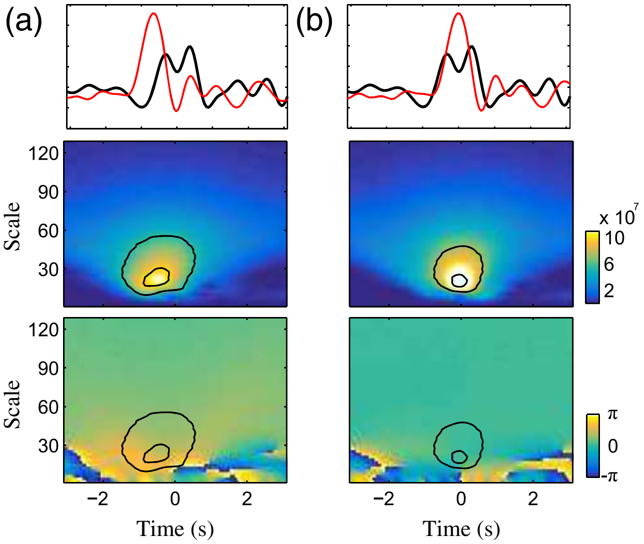

Figure 12. (a, top) Horizontal components recorded at the F-net AMM station rotated by 48° to identify the fast-splitting direction. (a, middle) Amplitude of the wavelet cross-spectrum coefficients. The coherent region shown by 50% and 90% contours constrains the *S*-wave window. (a, bottom) Phase of the complex wavelet cross-spectrum coefficients showing the phase-difference between the two signals. Time zero corresponds to the expected *S*-wave arrival time. (b, top) Horizontal components after the time-delay correction is applied. (b, middle) Wavelet cross-spectrum amplitude for the corrected seismograms. (b, bottom) Wavelet cross-spectrum phase information for the corrected components. The color version of this figure is available only in the electronic edition.

(Fig. 13), and the plot of the peak amplitude spectra shows two local maxima (Fig. 13a). One maximum corresponds to waveforms that are correlated and another corresponds to waveforms that are anticorrelated. These two types are identified by examining the phase difference (Fig. 13b), that is, when the associated phase difference is close to π, the maximum is due to anticorrelation, and is disregarded. Based upon this criterion, the fast-splitting direction is observed to be between 49° and 52° with inferred delay time of 0.64 to 0.68 s (Fig. 13c). Both fast direction and delay time are in good agreement with previous results for this particular station, a fast-splitting azimuth of 46° and a delay time of 0.65 s (Long and van der Hilst, 2005). The above algorithm is objective, and as long as the noise is not coherent on two components being considered, it can automatically determine the splitting parameters even with noisy seismograms.

Discussion and Conclusions

The algorithms presented in this manuscript for automatically picking *P*- and *S*-wave arrivals perform at a level that is comparable to that of human analysts. Working in the wavelet domain allows multiresolution analysis of the waveform, and provides the means to distinguish the phase arrival from random or systematic noise. In particular, the *S*-wave arrival can be identified using the wavelet cross spectrum of the horizontal components even when the noise level is high. Although not implemented for this article, the effectiveness of the algorithm can be improved by rotating three-component seismograms into the QRT coordinates based upon the incidence angle (e.g., Jepsen and Kennett, 1990). It will help isolate *P*- and *S*-wave arrivals on one or two components and reduce coherent noise.

In this work, we attempt to keep the number of parameters that require tuning as low as possible. However, further optimization of parameters can increase the performance of the algorithms by adapting to specific problem of interest. For example, the scale dimension of the range filter should correspond to the frequency content of the signal, and this knowledge can be implemented in the form of wavelet selection to improve detection. Similarly, the time window around the theoretical arrival time should be modified based upon the type of application and the expected difference between observed and theoretical arrival times. Choosing a larger window increases the ability to capture arrivals that are unusually late or early, but increases the chance of misidentifying some later or earlier phase as the target phase. For the onset-time detection, threshold at each scale can be fine-tuned according to the specific noise characteristics of the application. The source properties can also be taken into consideration. For example, recordings of volcanic tremor show highly oscillatory arrival with emerging onset (e.g., Schick, 1981). For such data, using wavelets with high number of van-

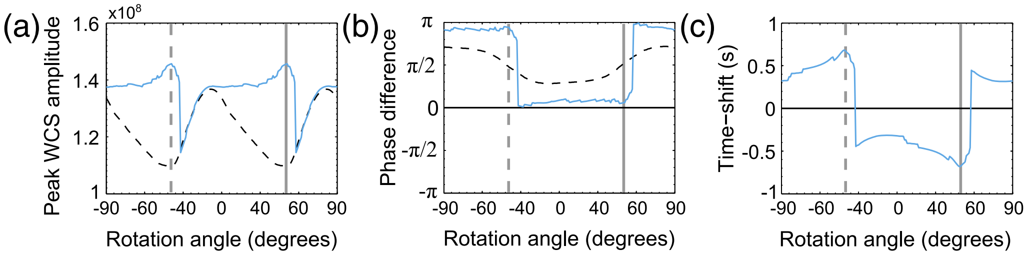

Figure 13. Determination of the fast-splitting direction and delay time of *S* waves recorded at the AMM station. The two vertical lines indicate the maximum coefficient values that correspond to phase difference of 0 (solid line) and π (dashed line), respectively. (a) Peak amplitude before (dashed curve) and after (solid curve) the delay-time correction as a function of rotation angles of the horizontal components. (b) The phase difference at the peak amplitude before (dashed curve) and after (solid curve) the delay-time correction. (c) The delay-time correction inferred from the phase difference at various rotation angles. The color version of this figure is available only in the electronic edition.

ishing moments as basis functions would be preferable. In the case of arrivals of teleseismic signals at a dense array, it is worth investigating the improvement in performance by replacing ordinary mother wavelets with customized wavelets, such as stacked waveforms, to generate the basis functions that are similar to the observed signal. This practice may also be fruitful in exploration applications where the morphology of the first arrival does not vary significantly along different receivers.

In addition to the onset time and the associated uncertainties, the analysis in the wavelet domain can be used to retrieve information such as the dominant frequency of the arrival and strength of dispersion. For example, frequency-dependent pick times can be utilized to derive frequency-dependent sensitivity kernels required in seismic tomography by determining the first Fresnel volume around the ray path. Because the algorithms described in this article are applied to raw seismograms, the frequency information is unbiased (except for the instrument response) and is data driven, as opposed to picking algorithms that require *a priori* narrow band-pass filtering of the waveform.

Finally, we demonstrate that the continuous wavelet transform approach is also suitable for developing an automatic algorithm for measuring the shear-wave splitting. To obtain phase information, the basis wavelet must be complex, and amplitude and phase coefficients of complex wave spectra are analyzed to generate information such as the fast-splitting direction and the delay time. The results obtained for the station AMM are consistent with a previous study, and the automatic implementation of the method will lead to compilation of a large database of shear-wave splitting measurements.

Data and Resources

Seismograms used in this study have been collected and made available by the National Research Institute for Earth Science and Disaster Prevention in Japan. The locations of the earthquakes used in this work are obtained from the Japan Meteorological Agency (JMA; http://www.jma.go.jp; last accessed August 2014).

Acknowledgments

We kindly acknowledge the Japanese National Research Institute for Earth Science and Disaster Prevention for providing the data used in this study. We also thank Lisa Tang for manually picking some of the seismic phases. We also gratefully acknowledge the Editor-in-Chief Diane I. Doser, the Associate Editor Hiroshi Kawase, Yoshihiro Ito, and one anonymous reviewer for their useful reviews and constructive comments that helped to improve the manuscript.

References

Akansu, A. N., W. A. Serdijn, and I. W. Selesnick (2010). Emerging applications of wavelets: A review, *Phys. Comm.* **3,** no. 1, 1–18, doi: 10.1016/j.phycom.2009.07.001.

Allen, R. V. (1978). Automatic earthquake recognition and timing from single traces, *Bull. Seismol. Soc. Am.* **68,** no. 5, 1521–1532.

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

Amoroso, O., N. Maercklin, and A. Zollo (2012). *S*-wave identification by polarization filtering and waveform coherence analyses, *Bull. Seismol. Soc. Am.* **102,** no. 2, 854–861, doi: 10.1785/0120110140.

Anant, K. S., and F. U. Dowla (1997). Wavelet transform methods for phase identification in three-component seismograms, *Bull. Seismol. Soc. Am.* **87,** no. 6, 1598–1612.

Ando, M., Y. Ishikawa, and F. Yamazaki (1983). Shear wave polarization anisotropy in the upper mantle beneath Honshu, Japan, *J. Geophys. Res.* **88,** 5850–5864.

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

Baillard, C., W. C. Crawford, V. Ballu, C. Hibert, and A. Mangeney (2014). An automatic kurtosis-based *P*- and *S*-phase picker designed for local seismic networks, *Bull. Seismol. Soc. Am.* **104,** no. 1, 394–409, doi: 10.1785/0120120347.

Billings, S. D., M. S. Sambridge, and B. L. Kennett (1994). Errors in hypocenter location: Picking, model, and magnitude dependence, *Bull. Seismol. Soc. Am.* **84,** no. 6, 1978–1990.

Bondár, I., and D. Storchak (2011). Improved location procedures at the International Seismological Centre, *Geophys. J. Int.* **186,** no. 3, 1220–1244, doi: 10.1111/j.1365-246X.2011.05107.x.

Bowman, J. R., and M. Ando (1987). Shear-wave splitting in the upper-mantle wedge above the Tonga subduction zone, *Geophys. J. Roy. Astron. Soc.* **88,** 25–41, doi: 10.1111/j.1365-246X.1987.tb01367.x.

Butzer, P., A. Fischer, and K. Rückforth (1994). Scaling functions and wavelets with vanishing moments, *Comput. Math. Appl.* **27,** no. 3, 33–39, doi: 10.1016/0898-1221(94)90044-2.

Correig, A. (1991). Body-wave dispersion: Measurement and interpretation, *Pure Appl. Geophys.* **136,** no. 4, 561–576, doi: 10.1007/BF00878587.

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.

Dai, H., and C. MacBeth (1997). The application of back-propagation neural network to automatic picking seismic arrivals from single-component recordings, *J. Geophys. Res.* **102,** no. B7, 15105–15113, doi: 10.1029/97JB00625.

Daubechies, I. (1988). Orthonormal bases of compactly supported wavelets, *Comm. Pure Appl. Math.* **41,** no. 7, 909–996, doi: 10.1002/cpa.3160410705.

Daubechies, I. (1992). Ten lectures on wavelets, *SIAM, Philadelphia, PA* **61,** 357 pp., doi: 10.1137/1.9781611970104.

Devilee, R. J., J. Trampert, and H. Paulssen (2003). Dispersion measurements of *P* waves and their implications for mantle Q_P, *Pure Appl. Geophys.* **160,** no. 12, 2223–2238, doi: 10.1007/s00024-003-2399-8.

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. B. Young (1997). On the onset of *P* seismograms, *Geophys. J. Int.* **129,** 681–690, doi: 10.1111/j.1365-246X.1997.tb04503.x.

Earle, P. S., and P. M. Shearer (1994). Characterization of global seismograms using an automatic-picking algorithm, *Bull. Seismol. Soc. Am.* **84,** no. 2, 366–376.

Engdahl, E. R., R. van der Hilst, and R. Buland (1998). Global teleseismic earthquake relocation with improved travel times and procedures for depth determination, *Bull. Seismol. Soc. Am.* **88,** no. 3, 722–743.

Farge, M. (1992). Wavelet transforms and their applications to turbulence, *Ann. Rev. Fluid Mech.* **24,** no. 1, 395–458, doi: 10.1146/annurev.fl.24.010192.002143.

Fukao, Y. (1984). Evidence from core-reflected shear waves for anisotropy in the earth's mantle, *Nature* **309,** 695–698, doi: 10.1038/309695a0.

Furumura, T., and B. L. N. Kennett (2005). Subduction zone guided waves and the heterogeneity structure of the subducted plate: Intensity anomalies in northern Japan, *J. Geophys. Res.* **110,** no. B10302, doi: 10.1029/2004JB003486.

Galiana-Merino, J. J., J. Rosa-Herranz, and S. Parolai (2008). Seismic *P* phase picking using a Kurtosis-based criterion in the stationary wavelet domain, *IEEE Trans. Geosci. Remote Sens.* **46,** no. 11, 3815–3826.

Gentili, S., and A. Michelini (2006). Automatic picking of *P* and *S* phases using a neural tree, *J. Seismol.* **10,** no. 1, 39–63, doi: 10.1007/s10950-006-2296-6.

Grinsted, A., J. C. Moore, and S. Jevrejeva (2004). Application of the cross wavelet transform and wavelet coherence to geophysical time series, *Nonlinear Process. Geophys.* **11,** nos. 5/6, 561–566, doi: 10.5194/npg-11-561-2004.

Grossmann, A., and J. Morlet (1984). Decomposition of Hardy functions into square integrable wavelets of constant shape, *SIAM J. Math. Anal.* **15,** no. 4, 723–736, doi: 10.1137/0515056.

Hardebeck, J. L., and P. M. Shearer (2002). A new method for determining first-motion focal mechanisms, *Bull. Seismol. Soc. Am.* **92,** no. 6, 2264–2276, doi: 10.1785/0120010200.

Heil, C. E., and D. F. Walnut (1989). Continuous and discrete wavelet transforms, *SIAM Rev.* **31,** no. 4, 628–666.

Horiuchi, S. (2003). Automatic hypocenter location at times of extremely high seismic activity, in *Methods and Applications of Signal Processing in Seismic Network Operations*, Springer, Berlin Heidelberg, 98, 209–219, doi: 10.1007/BFb0117704.

Hudgins, L., C. A. Friehe, and M. E. Mayer (1993). Wavelet transforms and atmospheric turbulence, *Phys. Rev. Lett.* **71,** 3279–3282.

Jepsen, D. C., and B. L. Kennett (1990). Three-component analysis of regional seismograms, *Bull. Seismol. Soc. Am.* **80,** no. 6B, 2032–2052.

Karamzadeh, N., G. Javan Doloei, and A. Reza (2013). Automatic earthquake signal onset picking based on the continuous wavelet transform, *IEEE Trans. Geosci. Remote. Sens.* **51,** no. 5, 2666–2674, doi: 10.1109/TGRS.2012.2213824.

Kumar, P., and E. Foufoula-Georgiou (1997). Wavelet analysis for geophysical applications, *Rev. Geophys.* **35,** no. 4, 385–412, doi: 10.1029/97RG00427.

Küperkoch, L., T. Meier, J. Lee, W. Friederich, and Working Group EGELADOS (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.

Leonard, M., and B. Kennett (1999). Multi-component autoregressive techniques for the analysis of seismograms, *Phys. Earth Planet. In.* **113,** nos. 1/4, 247–263, doi: 10.1016/S0031-9201(99)00054-0.

Levander, A., and G. Nolet (2013). Perspectives on array seismology and USArray, in *Seismic Earth: Array Analysis of Broadband Seismograms*, American Geophysical Union, Washington, D. C., 1–6, doi: 10.1029/157GM01.

Levin, V., W. Menke, and J. Park (1999). Shear wave splitting in Appalachians and Urals: A case for multilayered anisotropy, *J. Geophys. Res.* **104,** 17,975–17,994.

Lilly, M. J., and J. Park (1995). Multiwavelet spectral and polarization analyses of seismic records, *Geophys. J. Int.* **122,** no. 3, 1001–1021 doi: 10.1111/j.1365-246X.1995.tb06852.x.

Lois, A., E. Sokos, N. Martakis, P. Paraskevopoulos, and G.-A. Tselentis (2013). A new automatic *S*-onset detection technique: Application in local earthquake data, *Geophys.* **78,** no. 1, KS1–KS11, doi: 10.1190/geo2012-0050.1.

Long, M. D., and P. G. Silver (2009). Shear wave splitting and mantle anisotropy: Measurements, interpretations, and new directions, *Surv. Geophys.* **30,** nos. 4/5, 407–461, doi: 10.1007/s10712-009-9075-1.

Long, M. D., and R. D. van der Hilst (2005). Upper mantle anisotropy beneath Japan from shear wave splitting, *Phys. Earth Planet. In.* **151,** nos. 3/4, 206–222, doi: 10.1016/j.pepi.2005.03.003.

Maeda, N. (1985). A method for reading and checking phase times in auto-processing system of seismic data, *Zisin* **38,** 365–379.

Mallat, S. (1999). A Wavelet Tour of Signal Processing, in *Wavelet Analysis & Its Applications*, Second Ed., Academic Press, San Diego, California, 620 pp.

Mousset, E., Y. Cansi, R. Crusem, and Y. Souchet (1996). A connectionist approach for automatic labeling of regional seismic phases using a single vertical component seismogram, *Geophys. Res. Lett.* **23,** no. 6, 681–684, doi: 10.1029/95GL03811.

Okada, Y., K. Kasahara, S. Hori, K. Obara, S. Sekiguchi, H. Fujiwara, and A. Yamamoto (2004). Recent progress of seismic observation networks in Japan - Hi-net, F-net, K-NET and KiK-net, *Earth Planets Space* **56,** no. 8, 15–28.

Oye, V., and M. Roth (2003). Automated seismic event location for hydrocarbon reservoirs, *Comput. Geosci.* **29,** no. 7, 851–863, doi: 10.1016/S0098-3004(03)00088-8.

Panagiotakis, C., E. Kokinou, and F. Vallianatos (2008). Automatic *P*-phase picking based on local-maxima distribution, *IEEE Trans. Geosci. Remote Sens.* **46,** no. 8, 2280–2287, doi: 10.1109/TGRS.2008.917272.

Persson, L. (2003). Statistical tests for regional seismic phase characterizations, *J. Seismol.* **7,** no. 1, 19–33, doi: 10.1023/A:1021216313892.

Saragiotis, C., L, Hadjileontiadis, and S. Panas (2002). PAI-S/K: A robust automatic seismic *P* phase arrival identification scheme, *IEEE Trans. Geosci. Remote Sens.* **40,** no. 6, 1395–1404, doi: 10.1109/TGRS.2002.800438.

Schick, R. (1981). Source mechanism of volcanic earthquakes, *Bull. Volcanol.* **44,** no. 3, 491–497, doi: 10.1007/BF02600579.

Sleeman, R., and T. van Eck (1999). Robust automatic *P*-phase picking: An on-line implementation in the analysis of broadband seismogram recordings, *Phys. Earth Planet. In.* **113,** nos. 1/4, 265–275, doi: 10.1016/S0031-9201(99)00007-2.

Taylor, K. M., M. J. Procopio, C. J. Young, and F. G. Meyer (2011). Estimation of arrival times from seismic waves: A manifold-based approach, *Geophys. J. Int.* **185,** 435–452, doi: 10.1111/j.1365-246X.2011.04947.x.

Tibuleac, I. M., and E. T. Herrin (1999). An automatic method for determination of *Lg* arrival times using wavelet methods, *Seismol. Res. Lett.* **70,** 577–595.

Torrence, C., and G. P. Compo (1998). A practical guide to wavelet analysis, *Bull. Am. Meteorol. Soc.* **79,** no. 1, 61–78, doi: 10.1175/1520-0477(1998)079<0061:APGTWA>2.0.CO;2.

Vonesch, C., T. Blu, and M. Unser (2007). Generalized Daubechies wavelet families, *IEEE Trans. Signal Process.* **55,** no. 9, 4415–4429, doi: 10.1109/TSP.2007.896255.

Walker, J. S. (2008). *A Primer on Wavelets and Their Scientific Applications*, Chapman and Hall/CRC, Boca Raton, Florida, 320 pp.

Wang, J., and T.-L. Teng (1995). Artificial neural network-based seismic detector, *Bull. Seismol. Soc. Am.* **85,** no. 1, 308–319.

Willemann, R. J., and D. A. Storchak (2001). Data collection at the International Seismological Centre, *Seismol. Res. Lett.* **72,** no. 4, 440–453, doi: 10.1785/gssrl.72.4.440.

Withers, M., R. Aster, C. Young, J. Beiriger, M. Harris, S. Moore, and J. Trujillo (1998). A comparison of select trigger algorithms for automated global seismic phase and event detection, *Bull. Seismol. Soc. Am.* **88,** no. 1, 95–106.

Zeiler, C., and A. A. Velasco (2009). Seismogram Picking Error from Analyst Review (SPEAR): Single-analyst and institution analysis, *Bull. Seismol. Soc. Am.* **99,** no. 5, 2759–2770, doi: 10.1785/0120080131.

Zhang, H., C. Thurber, and C. Rowe (2003). Automatic P-wave arrival detection and picking with multiscale wavelet analysis for single-component recordings, *Bull. Seismol. Soc. Am.* **93,** 1904–1912, doi: 10.1785/0120020241.

Zhao, D., A. Hasegawa, and S. Horiuchi (1992). Tomographic imaging of P and S wave velocity structure beneath northeastern Japan, *J. Geophys. Res.* **97,** no. B13, 19909–19928, doi: 10.1029/92JB00603.

Department of Earth and Planetary Sciences
Harvard University Cambridge
Massachusetts 02138

Manuscript received 4 September 2014;