

An Automatic Phase Picker for Local Earthquakes with Predetermined Locations: Combining a Signal-to-Noise Ratio Detector with 1D Velocity Model Inversion

by Zefeng Li and Zhigang Peng

ABSTRACT

We present a simple method to pick additional P and S phases for local earthquakes with predetermined locations. Different from conventional characteristic function techniques, our method incorporates 1D velocity inversion into phase picking. It first predicts phase arrivals using initial velocity models and available event locations and then applies a detector function to search genuine phase arrivals around the initial predictions. Using the newest searched phase arrivals (picks), the velocity models are updated accordingly and then used to predict more accurate arrival times. Such a procedure can be iterated multiple times, during which both phase picking and velocity model are improved. We perform a synthetic test with a 1D velocity model. The resulting phase picks are consistent with true arrivals, and true velocity models can be well recovered. Another test with real data recorded by the Anza Seismic Network in southern California shows that the resulting velocity models agree with the average Southern California Earthquake Center community models in this region. Out of 23,932 event–station pairs, 23,770 P picks and 21,935 S picks are obtained from this method. After four iterations, 90% of our P picks and 80% of our S picks have differences within 0.15 s from the phase picks of the Southern California Earthquake catalog. Given its simplicity and efficiency and ability to produce robust P and S picks and 1D velocity models, the technique is particularly suitable for a wide range of seismological research in which phase picks at additional stations and/or refined 1D velocity models are needed.

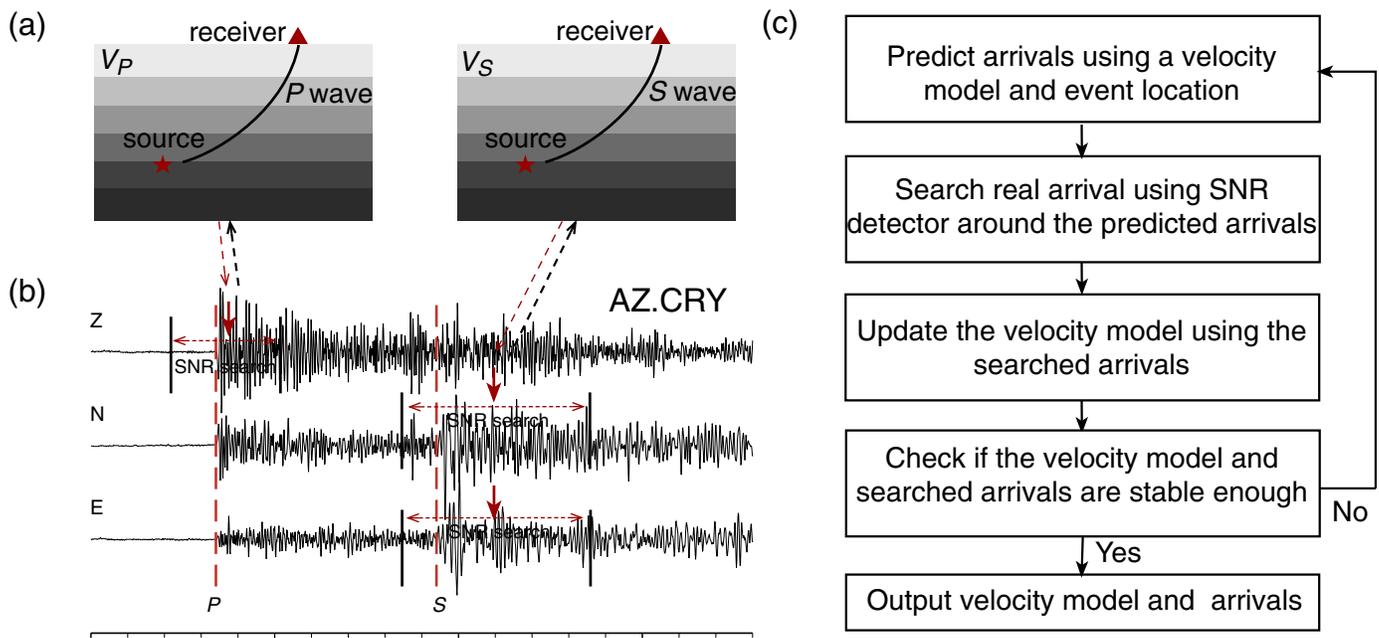
Online Material: Synthetic test results for noise-added waveforms.

INTRODUCTION

Automatic phase picking is a long-standing topic in observational seismology (e.g., Allen, 1982). Accurate identification of seismic

phase arrival times is of particular importance for determining event location and subsurface velocity structure. Although phase identification can be done manually, it is rather time consuming and subjective. As seismic data volume increases dramatically, many automatic phase-picking algorithms and characteristic functions have been developed in the past few decades. Beginning with the short-term average/long term average (STA/LTA) method (Allen 1978, 1982; Earle and Shearer, 1994), a few higher order statistical metrics have been proposed, including envelope functions (e.g., Baer and Kradolfer, 1987) and skewness and kurtosis functions (e.g., Saragiotis *et al.*, 2002; Kuperkoch *et al.*, 2010; Baillard *et al.*, 2014). Algorithms from the data-mining community are also introduced for phase picking purpose, such as neural networks (Dai and MacBeth, 1997; Wang and Teng, 1997; Zhao and Takano, 1999; Gentili and Michelini, 2006), autoregressive methods (Takanami and Kitagawa, 1988; Leonard and Kennett, 1999; Sleeman and van Eck, 1999; Zhang *et al.*, 2003), and nearest neighbors (Rawles and Thurber, 2015). Some studies combine these approaches to improve the robustness of the picking procedure (Sleeman and van Eck, 1999; Patane *et al.*, 2003; Diehl *et al.*, 2009; Ross and Ben-Zion, 2014). These methods are designed to identify sharp changes in waveform amplitudes without *a priori* information of seismic events. Therefore, they are essential at initial stages of seismic data processing, that is, building earthquake catalogs.

However, in some seismological studies we may encounter a different situation: starting with event catalogs in hand (generally from network centers or government agencies responsible for reporting earthquakes) but without phase arrival information at certain stations. One example is seismic travel-time tomography (e.g., Aki and Lee, 1976), which is widely used to image the Earth's 3D seismic-velocity structures. Another example is using waveform template matching technique to search missing events from standard catalogs, which requires phase arrivals to select proper windows for waveform correlation (e.g., Peng and Zhao, 2009). These studies mainly use



▲ **Figure 1.** Schematic diagram and flowchart for the Predict-Search-Invert-Repeat (PSIR) procedure. (a) 1D velocity models, source (star) and receiver (triangle) locations, and ray paths for P and S waves. (b) An example three-component seismogram recorded at station CRY in the ANZA seismic network. Vertical-dashed lines mark the P and S arrivals. Vertical arrows mark the predicted arrivals. Dashed arrows denote window ranges for searching the real P and S arrivals. Dashed arrows linking two parts suggest that the velocity model can predict arrivals and in return can be updated by the searched picks. (c) The detailed flowchart of the PSIR procedure. The color version of this figure is available only in the electronic edition.

earthquakes as listed in published catalogs and require reliable measurements of seismic phase arrival times for certain stations or networks. Even for the stations in which phase picks are routinely provided, certain phases, such as S picks, are usually not complete. On the other hand, with event locations in hand, simply using 1D or 3D velocity models to predict theoretical phase arrivals may not be satisfactory (e.g., Peng and Zhao, 2009), because of inaccurate velocity models. Another common way is to use characteristic functions, such as the aforementioned STA/LTA function, to pick phase arrivals. However, using characteristic functions alone may not be the optimal choice. For example, the performance of characteristic functions tends to decrease in the presence of high-amplitude local noise or seismic phases other than the targeted phases. For this reason, characteristic functions usually lead to low production of S picks, which are generally contaminated by P -wave coda and P/S converted phases.

Here, we show that additional information from event catalogs may be exploited to enhance picking accuracy. A simple algorithm is introduced to combine characteristic function search for phase arrivals and velocity model prediction, in an iterative sense. This method ensures that a detector function is applied within a time window allowed by initial velocity models, thereby avoiding erroneous picks clearly outside typical P and S windows. It also updates the velocity models and makes more accurate prediction of phase arrivals in the next iteration. The resulting velocity models derived in this way can be used

for other research purposes, such as earthquake location and tomography. We test this method with both synthetic and real data sets and demonstrate its capability to produce reliable phase picks and improved 1D velocity models.

METHOD

Figure 1 shows the schematic diagram and flowchart of the method. As mentioned before, we first take event locations from a given event catalog and a preliminary velocity model as inputs. Next, given the location of an event and the preliminary velocity model, the phase arrivals are predicted at several stations. We then apply a standard signal-to-noise ratio (SNR) detector to search the actual arrival within a window around the predicted arrival. The searched picks are used as input to the velocity model inversion. The updated velocity model is used to predict theoretical arrivals again, and thus the entire procedure can be iterated. The use of velocity models ensures that the detector function is applied within a reasonable time window, and in return the newly identified picks improve the velocity model. The above process is repeated until the picks and the velocity model become stable, which usually takes no more than three iterations. The method is essentially a joint search for the picks and the velocity model based on predetermined event locations. We term this procedure as Predict-Search-Invert-Repeat (PSIR, which could be pronounced as SIRE or SIR), and we describe the details of each step below.

The travel-time calculation follows the code in the frequency–wavenumber synthetic seismogram package developed by [Zhu and Rivera \(2002\)](#), which computes the ray parameter by a shooting method. To determine the window for searching the real phase arrival, we assume that the actual velocity deviates within a percentage of the velocity model

$$\epsilon = \Delta V_{\max}/V, \quad (1)$$

in which V is a given velocity model and ΔV_{\max} is the maximal allowed difference between the true velocity structure and the given velocity model. We note that ϵ is an average perturbation of the whole velocity model, but in reality, the velocity perturbation may vary within each layer. The parameter ϵ can be preset by users, based on a rough estimate of the accuracy of the given velocity model. By propagating the velocity perturbation to arrival-time perturbation, the search time window becomes

$$\left[t_O + \frac{t_T}{1 + \epsilon}, \quad t_O + \frac{t_T}{1 - \epsilon} \right], \quad (2)$$

in which t_O is the event origin time and t_T is the theoretical travel time. It is worth noting that this search window may not include all true arrivals, due to underestimation of the inaccuracy of the velocity model and/or lateral heterogeneity, which is usually the case for some initial velocity models. However, as long as actual arrivals falling within the search window are sufficient for stable inversion of the velocity model, the inverted velocity model should shift toward the true one. More importantly, based on the updated velocity model, the predicted arrival and the search window are revised to be more accurate, and the number of successful picks should increase in the next iteration.

To search the true picks within the window, we use the following SNR function:

$$\text{SNR}(t) = \frac{\int_t^{t+T} A^2(\tau) d\tau}{\int_{t-T}^t A^2(\tau) d\tau}, \quad t_O + \frac{t_T}{1 + \epsilon} \leq t \leq t_O + \frac{t_T}{1 - \epsilon}, \quad (3)$$

in which A is the seismic waveform amplitude and T is the predefined window length for SNR calculation. The true phase arrival is simply picked at the maximum of SNR function in the window, which is different from conventional characteristic function pickers that widely use preset trigger thresholds. Because the maximum within a window always exists, a pick is always identified. Hence, it is important to use SNR to control the pick quality. A pick would be discarded when the maximum is found at the window edge. This is to avoid flat/monotonously increasing/decreasing amplitude or other erroneous data artifacts within the window. Other exceptions, such as incomplete waveform header information or missing data points, could also result in picking failure.

After obtaining arrivals from the SNR picker, we update the velocity model via 1D velocity inversion, which is a

common technique in seismology. Here, we follow [Shearer \(2009\)](#) to formulate the problem

$$\delta t_i = \sum_j t_{ij} \epsilon_j, \quad (4)$$

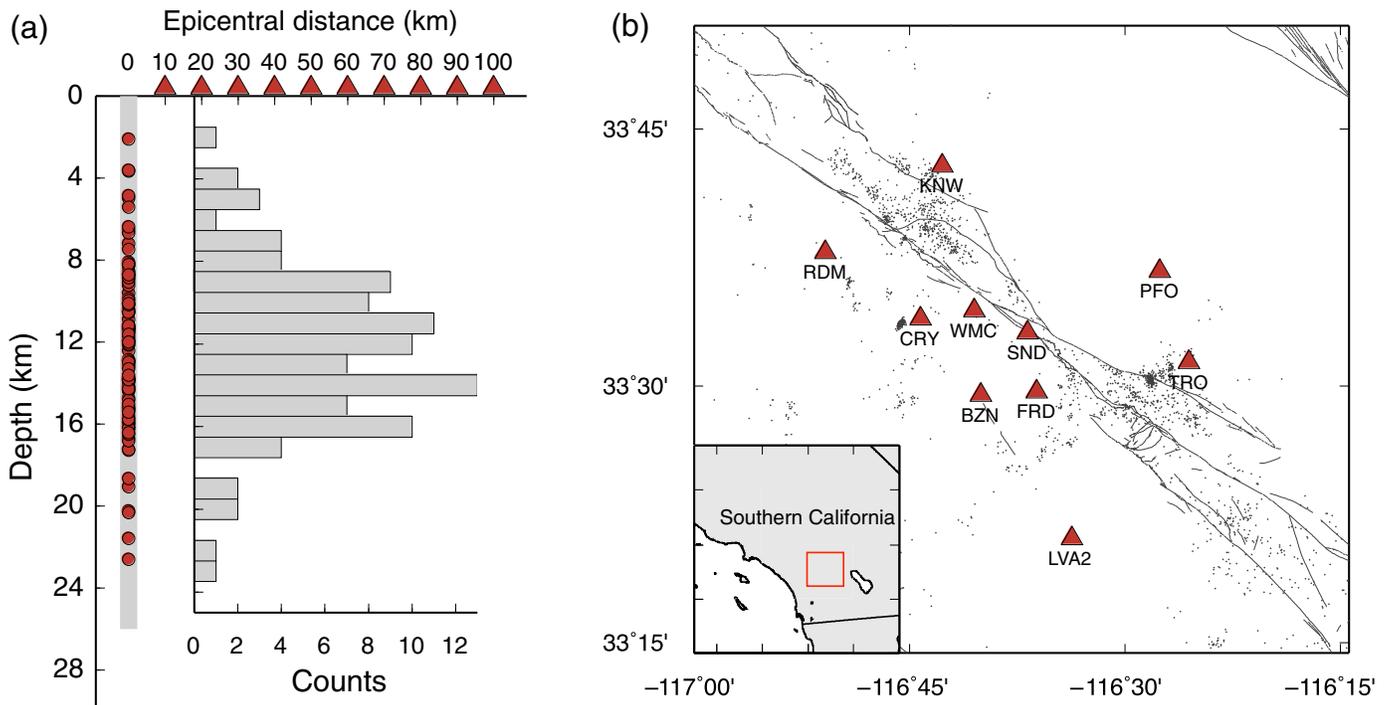
in which i denotes the i th travel-time measurement and j denotes the j th layer. Therefore, δt_i is the residual between the i th searched arrival and the i th predicted arrival, t_{ij} is the travel time in the j th layer of the i th measurement, and ϵ_j is the fraction of velocity perturbation in the j th layer. This is a linear equation and is typically ill posed, so we use zero-order Tikhonov regularization to solve it ([Aster et al., 2013](#)). The regularization parameter is chosen to be 10 for our test cases, but is available for change at users' choice.

In our procedure, P phase is picked on the vertical component, and the SNR measured at the pick is used as a quality metric. S phase can be picked on east and north horizontal components separately, and the final S pick is their SNR-weighted average. Because the event location is available, the S pick could be made simply on the transverse component. The later approach is not elaborated here but is included as a separate version in the downloadable package. Only the picks with high enough SNR are used in the velocity inversion (in the following tests we define the threshold to be $\text{SNR} > 5$). The algorithm allows users to set the velocity perturbation ϵ , based on their evaluation of the initial velocity model accuracy. However, if ϵ is set too large, for example, $\epsilon > 0.27$ for $V_P/V_S = \sqrt{3}$, the search window for P wave overlaps with the one for S wave. If overlap does occur, the algorithm automatically sets the start of the S window to be after the searched P arrival. Users can also set the window length T for SNR calculation, according to the sampling rate and noise level of a specific data set. Typically a large T is more noise-robust, but it is less sensitive to localized amplitude change. By default, the algorithm uses 0.1 and 0.2 s for P and S waves, respectively.

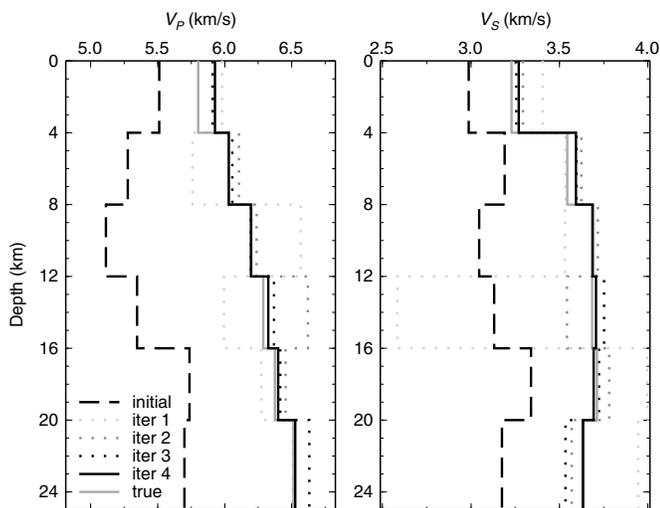
TEST WITH SYNTHETIC DATA

In the test, we use the frequency–wavenumber synthetic seismogram package ([Zhu and Rivera, 2002](#)) to compute synthetic waveforms. The setting includes 10 receivers at the surface and 100 events at depth (Fig. 2a). The epicenters of the 100 events are placed at $X = 0$ km, whereas their depths follow a pseudo-Gaussian distribution with mean at 12.5 km and truncated at the edges of 0 and 25 km. The 10 receivers are located at $X = 10, 20, \dots, 90, 100$ km. The focal mechanism parameters (i.e., strike, dip, and rake) of these events are randomly set with even chance in the parameter's allowable ranges. P and S phases are searched on vertical and horizontal components, respectively.

Using the SCEC community velocity model ([Magistrale et al., 1996](#)), we extract the velocity profiles at the locations of 10 stations in the ANZA seismic network (Fig. 2b). These 10 velocity profiles are then averaged into a 1D model with each layer of 4 km thickness (gray lines in Fig. 3), which is used for synthetic seismogram generation. Next, we apply a



▲ **Figure 2.** (a) The synthetic test setup. Triangles represent 10 surface receivers from 10 to 100 km with 10 km spacing. Circles represent 100 events with depths following a pseudo-Gaussian distribution and truncated at the edges of 0 and 25 km. The histogram shows the counts of events in 1 km bins in depth. (b) Map of the San Jacinto fault zone in southern California. Triangles represent stations in the ANZA seismic network used in the test. Background seismicity is between 1 January 2014 and 1 June 2014 from the Southern California Seismic Network (SCSN) catalog. The rectangle in the inset marks the study region in a larger map of southern California. The color version of this figure is available only in the electronic edition.

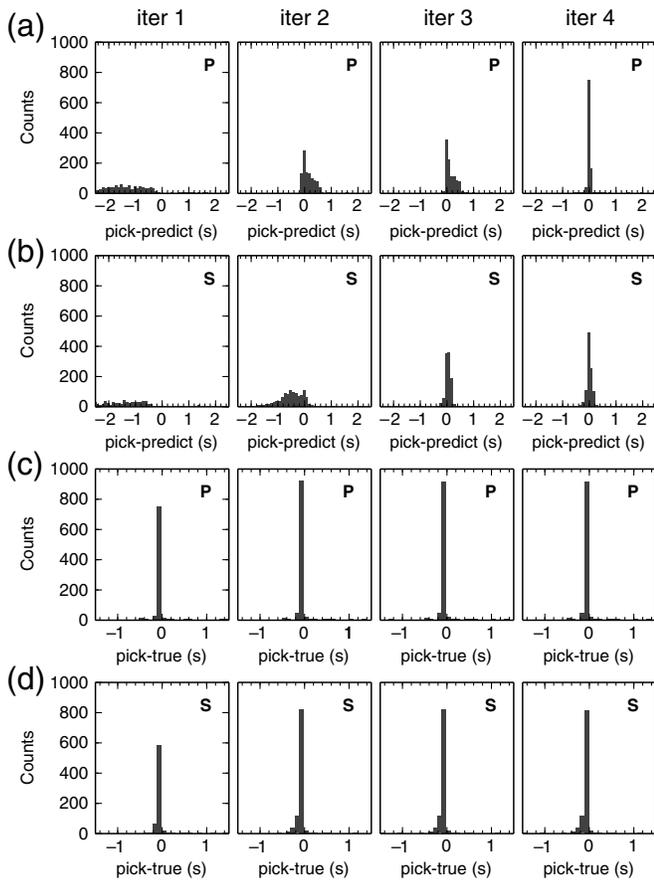


▲ **Figure 3.** 1D P - and S -velocity models in the synthetic test. Dashed lines and solid gray lines represent initial and true models, respectively. The light gray, moderate gray, and dark gray dotted lines represent the inverted velocity models in the first, second, and third iteration, respectively. The solid black line is the resulting velocity model in the fourth iteration.

5%–20% random velocity perturbation in each layer of the true model in order to obtain a perturbed velocity model serving as the starting model for the test (dashed lines in Fig. 3).

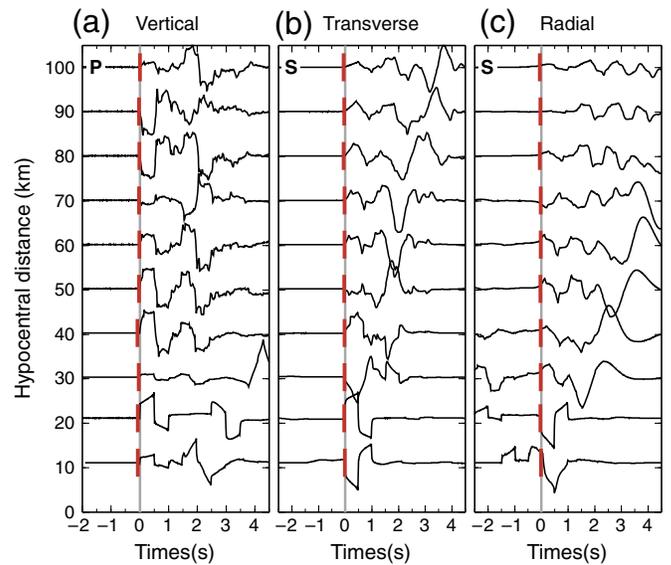
Although the actual maximum velocity perturbation could reach up to 20%, we set the maximum allowed perturbation $\epsilon = 15\%$ for all iterations. In the synthetic test, we perform four iterations. The resulting P - and S -velocity models converge well to true velocity models, with the relative inaccuracy about 1% or less in each layer (Fig. 3). Figure 4a and 4b shows the difference of the searched picks relative to the predicted arrivals, based on the initial velocity models (iteration 1) and updated models (iterations 2–4). It is clear that the discrepancy between the searched picks and predicted arrivals narrows down during four iterations. Because the initial models are significantly slower than the true models, the travel-time residuals are clustered at the negative sides. Although some of the picks cannot be searched because the preset maximum perturbation is less than the actual maximum perturbation, many picks within the 15% range can still be obtained. These picks are then used to update the velocity models, which help in reducing differences between the predicted arrivals and the true arrivals in the later iterations.

After four iterations, the searched picks are very consistent with the prediction from the updated velocity models (iteration 4 in Fig. 4a,b). Figure 4c and 4d shows the residuals



▲ **Figure 4.** Histograms of *P* and *S* pick residuals in four iterations of the synthetic test. (a) *P* and (b) *S* arrival-time residuals between searched and predicted arrivals from the initial (iter 1) and updated (iter 2, 3, 4) velocity models. (c) *P* and (d) *S* travel-time residuals between the searched and true arrivals from the true velocity models. The bin size is 0.1 s.

between the searched picks and the true arrivals. In the first iteration, because the velocity inaccuracy is underestimated, a number of arrivals are excluded in the search range and thus cannot be obtained. However, after the first iteration of velocity update, the number of successful picks increases significantly. As shown in Figure 4c,d, *P* picks increase from 821 to 1000 and *S* picks increase from 749 to 1000. From iterations 2–4, the residual histograms do not change much. This is because, after the first iteration, the velocity models converge within 15% perturbation of the true velocity models, thus most picks can be searched using the preset 15% threshold. Figure 5 shows a few automatic pick examples in comparison to true arrivals. To check for the performance involving noisy data, we also add zero-mean Gaussian noise with standard deviation at 20% and 50% of first arrival peak amplitude. Although there is a decrease in successful picks and greater error in the inverted velocity model, as the noise increases, the results generally agree with the true arrivals and velocity models (© Figs. S1–S6, available in the electronic supplement to this article).

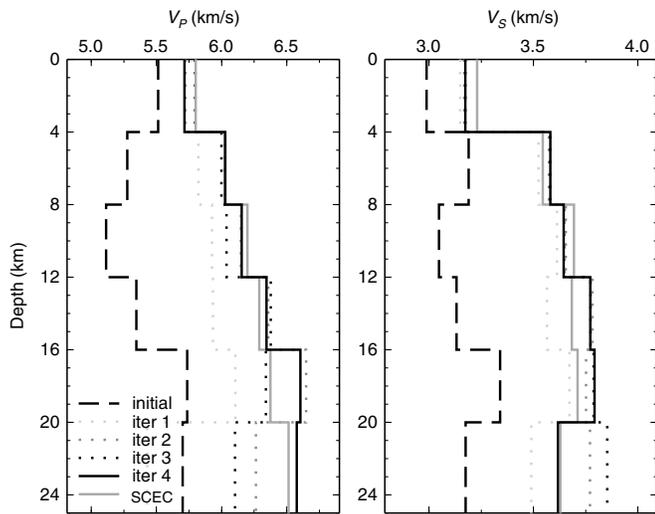


▲ **Figure 5.** Examples of *P* and *S* picks on synthetic seismograms. (a) The vertical waveforms are aligned at true *P* arrivals (gray vertical line), and the automatic *P* picks are marked by bars. The (b) transverse and (c) radial waveforms are aligned at true *S* arrivals (gray vertical lines), and bars mark the automatic *S* picks. The color version of this figure is available only in the electronic edition.

TEST WITH REAL DATA

We test the method using six months of seismic data recorded by the ANZA seismic network in southern California (Fig. 2a). The data set consists of 2815 events recorded by 10 stations, resulting in 23,932 event–receiver pairs. We download the event catalog, waveforms, and phase picks from the Southern California Earthquake Data Center (SCEDC). We use the same starting velocity s as the ones in the synthetic test (Fig. 3). Similarly, the maximum allowed velocity perturbation ϵ is set as 15%. Although the exact velocity structure in the region is not known, we use the average SCEC community velocity model as a reference for the test output.

We perform four iterations and check the pick and velocity model changes (Fig. 6). Although with some minor differences, the resulting velocity models in the iterations are generally consistent with the average SCEC models, indicating the SCEC models agree well with those derived from our test data. Table 1 gives the number of picks in each iteration. At the beginning of the first iteration, because the initial velocity models are not accurate, many of the predicted arrivals are far from the true ones, so that the true arrivals are outside the search window. Therefore, we obtain a relatively low number of high-SNR picks in this iteration. By the end of the fourth iteration, we have 22,516 $\text{SNR} > 5$ and 19,638 $\text{SNR} > 10$ *P* picks, 21,494 $\text{SNR} > 5$, and 19,522 $\text{SNR} > 10$ *S* picks. In comparison, the SCEDC provides 22,101 *P* picks and 18,093 *S* picks.



▲ **Figure 6.** 1D P - and S -velocity models in the real data test. Dashed lines and solid gray lines represent initial and average Southern California Earthquake Center (SCEC) models, respectively. The light gray, moderate gray, and dark gray dotted lines represent the inverted velocity models in the first, second, and third iteration, respectively. The solid black line is the resulting velocity model in the fourth iteration.

Figure 7a and 7b shows the residuals between our automatic P and S picks and predicted picks. Similar to the synthetic test, the residuals between the searched arrivals and the predicted ones are concentrated at the negative side in the first iteration, which suggests that the initial models are systematically too slow. The residuals decrease quickly and concentrate at zero in the later iterations, which is largely due to the improvement of the velocity models. Figure 7c and 7d compares the obtained picks with the SCEDC phase picks. By the 4th iteration, 90% of P picks and 80% of S picks have residuals within ± 0.15 s. Figure 8 shows a few waveform pick examples of events at different hypocentral distances with varying SNR. Even for low SNR signals, the picker still performs well. Figure 8b also shows that the picks are generally within 0.2 s from the phase picks obtained from SCEDC. In general, the consistency between the automatic picks and the cataloged ones tends to decrease with decreasing SNR (Fig. 9a,b). The picks that are made by our method but not in the catalog are also with relatively low SNR (Fig. 9c,d). This partly demonstrates the method's picking ability in noisy data, although further selection criteria could be applied for quality control purpose.

DISCUSSIONS

In this study, we presented a new method (PSIR) to accurately pick the phase arrivals when preliminary earthquake locations and a 1D velocity model are available. This method essentially links a traditional characteristic function search with simple 1D velocity model inversions. By allowing a certain degree of

Table 1
Number of Picks in the Real Data Test

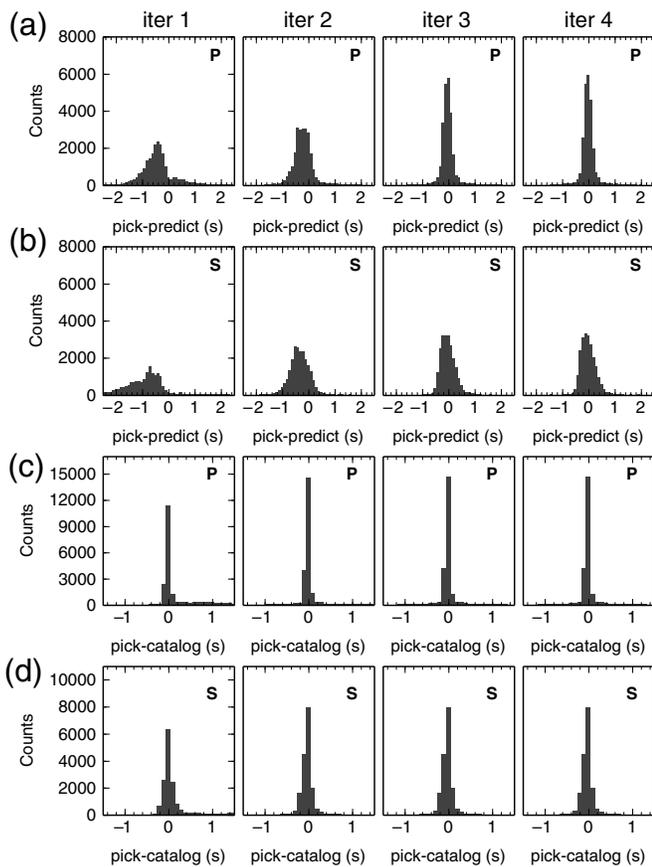
	Iter 1	Iter 2	Iter 3	Iter 4	Pick Rate (%)
P (SNR > 1)	21,002	23,385	23,784	23,770	99.3
P (SNR > 5)	18,207	22,091	22,516	22,516	94.1
P (SNR > 10)	14,843	19,193	19,643	19,638	82.1
S (SNR > 1)	16,785	21,936	21,938	21,935	91.7
S (SNR > 5)	15,218	21,490	21,498	21,494	89.8
S (SNR > 10)	13,274	19,499	19,519	19,522	81.6

SNR, signal-to-noise ratio.

velocity inaccuracy, the characteristic function is confined to search within a range, thereby avoiding false picks that are far from the theoretical expectation. Our synthetic tests (Figs. 3–5 and ⑥ Figs. S1–S6) demonstrate its ability to obtain accurate picks and converge from an inaccurate velocity model toward the true one. The test with ANZA network data (Figs. 6–8) shows a significant increase of S -phase picks in comparison with the existing catalog picks. The method could be useful for travel-time tomography, in which picks made outside reasonable P and S windows may introduce unreal velocity anomalies. In addition, the updated velocity model obtained using searched arrivals can also serve as a starting point for 3D seismic tomography.

Specially designed for the case in which an event catalog is available, our method does not compete with many automatic arrival pickers for phase identification from scratch (e.g., Ross and Ben-Zion, 2014). Instead, with *a priori* knowledge of event locations, the method increases the phase-pick success rate (especially S phases) and reduces false picks (picks that are significantly outside the typical arrival window or erroneous phases). Both are advantageous over characteristic functions alone. In addition to the tests above, this method has been applied to sizeable data sets, for example, in the San Jacinto fault zone (Li *et al.*, 2015) and in the Costa Rica subduction zone (Yao *et al.*, 2016). We find that the method generally achieves good accuracy in these data sets and can significantly reduce the workload of manual phase picking.

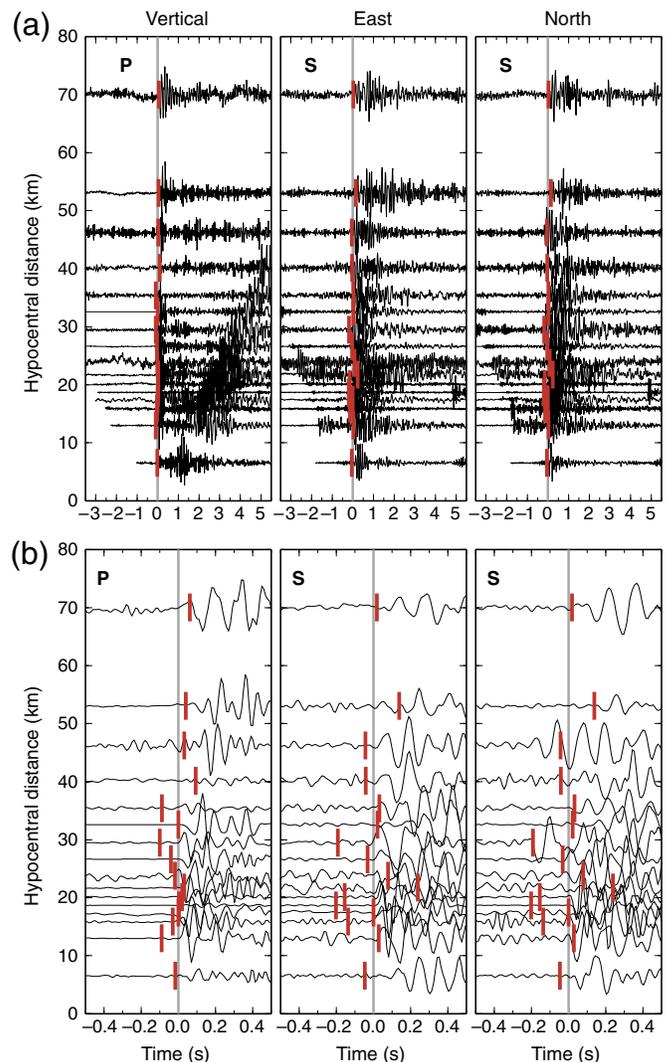
Our method considers the maximal SNR within the search range as the direct P and S arrivals. This may not always be true because other phases, such as P/S conversion, refracted (e.g., Pn or fault zone head waves), and/or reflected phases, may be more impulsive than direct phases and could also occur within the search range. Therefore, if a better velocity model is available, decreasing the search range parameter ϵ could help to reduce the possibility of erroneous phase picking. In particular, a good choice of a starting velocity model could be the one used for locating the events. An alternative would be to replace the STA/LTA with other characteristic functions that are capable of identifying additional phases (e.g., Ross and Ben-Zion, 2014; Li and Peng, 2016). On the other hand, a threshold can be applied to select the picks with relatively higher SNRs, which tend to have sharper phase onset. Another assumption of the



▲ **Figure 7.** Histograms of *P* and *S* pick residuals at the end of each iteration in the real data test. (a) *P* and (b) *S* arrival-time residuals between the searched and predicted arrivals from the initial (iter 1) and updated (iter 2, 3, 4) velocity models. (c) *P* and (d) *S* arrival-time residuals between the searched and the referenced Southern California Earthquake Data Center (SCEDC) phase arrivals. The bin size is 0.1 s.

method is that event locations should be accurate enough and that the differences between the actual arrivals and theoretical predictions are simply attributed to inaccuracy of the velocity model. For a data set with significant location errors, the proposed method inherently treats the location error as part of the velocity model error, which means the search-range parameter ϵ should be increased to reflect both event location error and velocity model error. Once the arrival search window is set large enough, the total error can be handled, and picks can be still made. However, due to the coupled problem of event location and medium velocity, the inverted velocity model will be negatively impacted.

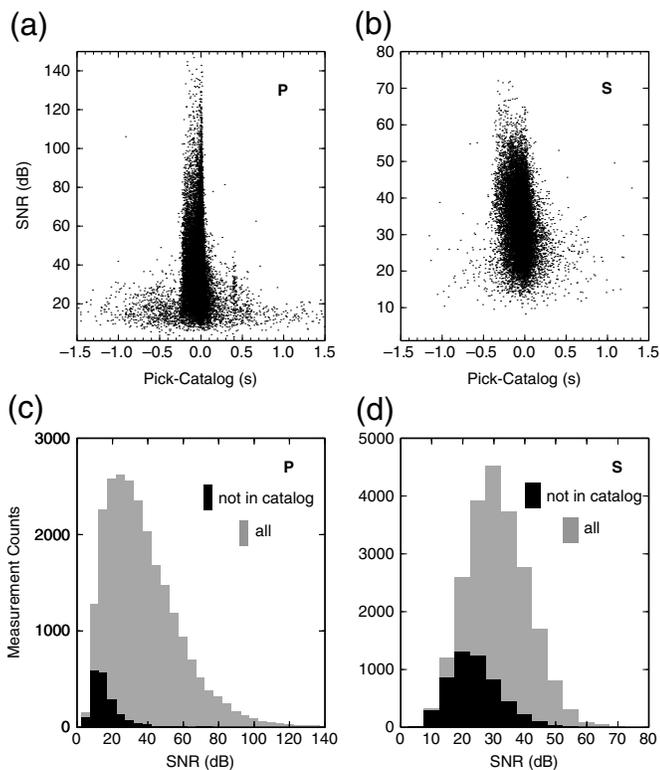
In the synthetic test, we obtained most of the picks after the first iteration (Fig. 4c,d), whereas the velocity models became close to the true models at the later iterations (Fig. 3). This is because once the velocity model differences decrease to 15%, most arrivals can be obtained and remain unchanged, whereas the velocity inaccuracy is further reduced. Therefore, although the iterative process in principle can be implemented many times as necessary, based on our experience, two or three



▲ **Figure 8.** Examples of *P* and *S* picks on the ANZA network data. (a) (Left) the vertical waveforms are aligned at SCEDC *P* picks (gray vertical line), and bars mark the automatic *P* picks. (Middle and right) The east-component and north-component waveforms are aligned at SCEDC *S* picks (gray vertical lines), and the automatic *S* picks are marked by bars. (b) The enlarged plot of the same waveforms within ± 0.5 s of the SCEDC picks. The color version of this figure is available only in the electronic edition.

iterations are sufficient to obtain the phase picks, but more iterations may be needed to update the velocity models. Too many iterations are not recommended because of relatively low gain relative to computational expense. Another possibility is to generalize the method to 3D velocity model inversion. However, compared to 1D velocity model inversion, this does not provide further benefit for the phase picking, because the search window defined by velocity perturbation already handles a certain amount of lateral heterogeneities.

With the amount of available seismic data increasing exponentially at Incorporated Research Institutions for



▲ **Figure 9.** Distribution of signal-to-noise ratios (SNRs) for all the picks. (a) and (b) SNR versus time differences of automatic and cataloged picks for *P* and *S* waves, respectively. (c) and (d) Histograms of SNR distribution of all the automatic picks (gray bars) and those not in the SCEDC catalog (black bars), for *P* and *S* waves, respectively.

Seismology Data Management Centers and other data centers, we expect that our procedure could be useful for other types of seismological research in which event catalogs are already available.

DATA AND RESOURCES

The event catalogs, waveforms, and phase picks in the Anza region were downloaded from the Southern California Earthquake Data Center (SCEDC) via the Seismogram Transfer Program. The velocity profiles beneath the ANZA network stations were from the Southern California Earthquake Center (SCEC) community velocity model (<http://www.data.sceec.org/research-tools/3d-velocity.html>, last accessed November 2015). The synthetic waveforms were generated by the frequency-wavenumber synthetic seismogram package developed by Lupei Zhu at the Saint Louis University (<http://www.eas.slu.edu/People/LZhu/home.html>, last accessed November 2015). The MATLAB code of the PSIR procedure can be downloaded at http://geophysics.eas.gatech.edu/people/zli/PSIRpicker/PSIRpicker_v1.0.tar.gz (last accessed February 2016). The related manual is available at <http://geophysics.eas.gatech.edu/people/zli/PSIRpicker/PSIRpicker.html> (last accessed February 2016). ☒

ACKNOWLEDGMENTS

We thank Electronic Seismologist column Editor Jeremy D. Zechar and four anonymous reviewers for helpful comments that improve the quality of this article significantly. The article also benefited from feedback from Chastity Aiken and Xiaofeng Meng. This work was supported by National Science Foundation (NSF) Grant Numbers EAR-1447091, EAR-1551022, and Southern California Earthquake Center (SCEC, Contribution Number 6547). SCEC was funded by NSF Cooperative Agreement EAR-1033462 and U.S. Geological Survey Cooperative Agreement G12AC20038.

REFERENCES

- Aki, K., and W. Lee (1976). Determination of three-dimensional velocity anomalies under a seismic array using first *P* arrival times from local earthquakes: 1. A homogeneous initial model, *J. Geophys. Res.* **81**, 4381–4399.
- Allen, R. (1978). Automatic earthquake recognition and timing from single traces, *Bull. Seismol. Soc. Am.* **68**, 1521–1532.
- Allen, R. (1982). Automatic phase pickers: Their present use and future prospects, *Bull. Seismol. Soc. Am.* **72**, S225–S242.
- Aster, R. C., B. Borchers, and C. Thurber (2013). *Parameter Estimation and Inverse Problems*, Second Ed., Elsevier, Boston, Massachusetts, ISBN 978-0-12-385048-5.
- Baer, M., and U. Kradolfer (1987). An automatic phase picker for local and teleseismic events, *Bull. Seismol. Soc. Am.* **77**, 1437–1445.
- Baillard, C., W. Crawford, V. Ballu, C. Hibert, and A. Mangeny (2014). An automatic kurtosis-based *P*- and *S*-phase picker designed for local seismic networks, *Bull. Seismol. Soc. Am.* **104**, 394–409, doi: [10.1785/0120120347](https://doi.org/10.1785/0120120347).
- 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, 15,105–15,113, doi: [10.1029/97JB00625](https://doi.org/10.1029/97JB00625).
- Diehl, T., E. Kissling, S. Husen, and A. Aldersons (2009). Consistent phase picking for regional tomography models: Application to the greater Alpine region, *Geophys. J. Int.* **176**, 542–554.
- Earle, P. S., and P. M. Shearer (1994). Characterization of global seismograms using an automatic picking algorithm, *Bull. Seismol. Soc. Am.* **84**, 366–376.
- Gentili, S., and A. Michelini (2006). Automatic picking of *P* and *S* phases using a neural tree, *J. Seismol.* **10**, 39–63.
- Kuperkoch, L., T. Meier, J. Lee, W. Friederich, and EGEADOS Working Group (2010). Automated determination of *P*-phase arrival times at regional and local distances using higher order statistics, *Geophys. J. Int.* **181**, 1159–1170.
- Leonard, M., and B. L. N. Kennett (1999). Multi-component autoregressive techniques for the analysis of seismograms, *Phys. Earth Planet. Int.* **113**, 247–263.
- Li, Z., and Z. Peng (2016). Automatic identification of fault zone head waves and direct *P* waves and its application in the Parkfield section of the San Andreas fault, California, *Geophys. J. Int.* **250**, 1326–1341, doi: [10.1093/gji/ggw082](https://doi.org/10.1093/gji/ggw082).
- Li, Z., Z. Peng, Y. Ben-Zion, and F. Vernon (2015). Spatial variations of shear-wave anisotropy near the San Jacinto Fault Zone in southern California, *J. Geophys. Res.* **120**, no. 12, 8334–8347, doi: [10.1002/2015JB012483](https://doi.org/10.1002/2015JB012483).
- Magistrale, H., K. McLaughlin, and S. Day (1996). A geology based 3-D velocity model of the Los Angeles basin sediments, *Bull. Seismol. Soc. Am.* **86**, 1161–1166.
- Patane, D., F. Ferrari, E. Giampiccolo, and S. Gresta (2003). A PC-based computer package for automatic detection and location of earth-

- quakes: Application to a seismic network in eastern Sicily (Italy), methods and applications of signal processing in seismic network operations, in *Lecture Notes in Earth Sciences*, Vol. 98, Springer, Berlin, Germany, 89–129.
- Peng, Z., and P. Zhao (2009). Migration of early aftershocks following the 2004 Parkfield earthquake, *Nature Geosci.* **2**, 877–881, doi: [10.1038/ngeo697](https://doi.org/10.1038/ngeo697).
- Rawles, C., and C. Thurber (2015). A non-parametric method for automatic determination of *P*-wave and *S*-wave arrival times: Application to local micro earthquakes, *Geophys. J. Int.* **202**, no. 2, 1164–1179, doi: [10.1093/gji/ggv218](https://doi.org/10.1093/gji/ggv218).
- Ross, Z. E., and Y. Ben-Zion (2014). Automatic picking of direct *P*, *S* seismic phases and fault zone head waves, *Geophys. J. Int.* **199**, 368–381.
- 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**, 1395–1404.
- Shearer, M. P. (2009). *Introduction to Seismology*, Second Ed., Cambridge University Press, New York, 119 pp.
- 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. Int.* **113**, 265–275.
- Takanami, T., and G. Kitagawa (1988). A new efficient procedure for the estimation of onset times of seismic waves, *J. Phys. Earth* **36**, 267–290.
- Wang, J., and T. Teng (1997). Identification and picking of *S*-phase using an artificial neural network, *Bull. Seismol. Soc. Am.* **87**, 1140–1149.
- Yao, D., X. Meng, Z. Peng, A. V. Newman, J. Walter, S. Schwartz, and M. Protti (2016). Detailed spatio-temporal evolution of aftershocks and repeating earthquakes following the 2012 M_w 7.6 Nicoya earthquake, *Seismol. Res. Lett.* **87**, no. 2B, 463–464.
- Zhang, H., C. Thurber, and C. Rowe (2003). Automatic *P*-wave arrival detection and picking with multiscale wavelet analysis for single-component recording, *Bull. Seismol. Soc. Am.* **93**, 1904–1912.
- Zhao, Y., and K. Takano (1999). An artificial neural network approach for broadband seismic phase picking, *Bull. Seismol. Soc. Am.* **89**, no. 3, 670–680.
- Zhu, L., and L. A. Rivera (2002). A note on the dynamic and static displacements from a point source in multi-layered media, *Geophys. J. Int.* **148**, 619–627.

Zefeng Li
Zbigang Peng
School of Earth and Atmospheric Sciences
Georgia Institute of Technology
311 Ferst Drive
Atlanta, Georgia 30332-0340 U.S.A.
zli354@gatech.edu

Published Online 24 August 2016