

Toward Fully Autonomous Seismic Networks: Backprojecting Deep Learning-Based Phase Time Functions for Earthquake Monitoring on Continuous Recordings

Wu-Yu Liao¹, En-Jui Lee^{*1} , Dawei Mu² , and Po Chen³ 

Abstract

Accurate and (near) real-time earthquake monitoring provides the spatial and temporal behaviors of earthquakes for understanding the nature of earthquakes, and also helps in regional seismic hazard assessments and mitigations. Because of the increase in both the quality and quantity of seismic data, an automated earthquake monitoring system is needed. Most of the traditional methods for detecting earthquake signals and picking phases are based on analyses of features in recordings of an individual earthquake and/or their differences from background noises. When seismicity is high, the seismograms are complicated, and, therefore, traditional analysis methods often fail. With the development of machine learning algorithms, earthquake signal detection and seismic phase picking can be more accurate using the features obtained from a large amount of earthquake recordings. We have developed an attention recurrent residual U-Net algorithm, and used data augmentation techniques to improve the accuracy of earthquake detection and seismic phase picking on complex seismograms that record multiple earthquakes. The use of probability functions of *P* and *S* arrivals and potential *P* and *S* arrival pairs of earthquakes can increase the computational efficiency and accuracy of backprojection for earthquake monitoring in large areas. We applied our workflow to monitor the earthquake activity in southern California during the 2019 Ridgecrest sequence. The distribution of earthquakes determined by our method is consistent with that in the Southern California Earthquake Data Center (SCEDC) catalog. In addition, the number of earthquakes in our catalog is more than three times that of the SCEDC catalog. Our method identifies additional earthquakes that are close in origin times and/or locations, and are not included in the SCEDC catalog. Our algorithm avoids misidentification of seismic phases for earthquake location. In general, our algorithm can provide reliable earthquake monitoring on a large area, even during a high seismicity period.

Cite this article as Liao, W.-Y., E.-J. Lee, D. Mu, and P. Chen (2022). Toward Fully Autonomous Seismic Networks: Backprojecting Deep Learning-Based Phase Time Functions for Earthquake Monitoring on Continuous Recordings, *Seismol. Res. Lett.* **93**, 1880–1894, doi: [10.1785/0220210274](https://doi.org/10.1785/0220210274).

Introduction

Earthquake monitoring is fundamental routine work in a seismic network. Accurate and complete earthquake activities are usually essential and critical for further studies in seismology, such as fault geometry, earthquake nucleation, earthquake triggering, postseismic deformation, and physical mechanisms of faulting (Peng and Zhao, 2009; Shelly and Hill, 2011; Ross, Idini, *et al.*, 2019; Ross, Trugman, *et al.*, 2019; Kato and Ben-Zion, 2021). Many studies reveal that small earthquakes can provide insights into related issues. For a more complete

catalog, the template matching filter (TMF), using the seismograms of known earthquakes to calculate their similarities with

1. Department of Earth Sciences, National Cheng Kung University, Tainan, Taiwan,  <https://orcid.org/0000-0003-1545-1640> (E-JL); 2. National Center for Supercomputing Applications, University of Illinois at Urbana-Champaign, Champaign, Illinois, U.S.A.,  <https://orcid.org/0000-0002-6354-6436> (DM); 3. Department of Geology and Geophysics, University of Wyoming, Laramie, Wyoming, U.S.A.,  <https://orcid.org/0000-0002-5148-9788> (PC)

*Corresponding author: rickli92@gmail.com

© Seismological Society of America

continuous data to find events, is a widely used technique to find missed small earthquakes (e.g., Peng and Zhao, 2009; Mu *et al.*, 2017; Ross, Trugman, *et al.*, 2019; Lee, Mu, *et al.*, 2020). However, significant limitations of TMF are that it can only detect events with high similarities and the need of an earthquake catalog for template seismograms. Therefore, it is necessary to provide a reliable earthquake catalog before using TMF for a more complete catalog. With the increase in data quality and quantity, automated earthquake detection and location algorithms in continuous recordings are needed. Recently, some automated algorithms for earthquake monitoring have been proposed (Tan *et al.*, 2019; Lee, Liao, *et al.*, 2020; Walter *et al.*, 2021; Zhou *et al.*, 2021), but there are still many challenges in automation, accuracy, and data scale.

Traditionally, earthquake detection, phase picking, and earthquake location are steps for cataloging earthquakes. Identifying waveforms and phases (e.g., *P* and *S* waves) generated by earthquakes from continuous seismic recordings are challenging. In continuous recordings, not only noises generated by other sources can interfere with earthquake detection, but it is also challenging to identify individual earthquake signals when the seismicity is high.

One of the commonly used approaches is to detect changes in the waveform features (e.g., energy, frequency content, and particle motions) that are different from the background noise (e.g., Allen, 1982; Lomax *et al.*, 2012; Baillard *et al.*, 2014). For example, the short-term average over long-term average (STA/LTA) can monitor the energy changes of incoming signals for earthquake signal or phases detections. The STA/LTA is straightforward and computationally efficient and, therefore, has been widely used. However, the generalization capabilities of traditional algorithms are usually poor, and many parameter settings need to be examined by trial and error for better performance (Akram and Eaton, 2016). Also, the traditional methods often fail when seismic activity is high (e.g., after-shock periods), that is, at the time when earthquake monitoring is the most needed. In general, traditional algorithms were designed without considering unexpected conditions, so they often fail when there are abnormal waveforms (e.g., waveform gaps) in continuous recordings. After picking available *P* and *S* arrivals from continuous recordings, a phase association algorithm is then used to cluster possible phase arrivals of an earthquake to invert for its source parameters. The grid search is one of the most commonly used methods and has been used in many earthquake monitoring systems (e.g., Weber *et al.*, 2007; Yeck *et al.*, 2019). When seismicity is high, many earthquakes with various magnitudes occurred close in time and location, and it is a challenging task to associate picked phases to individual earthquakes correctly.

In addition, backprojection-related methods, including waveform shifting and stacking processes, are also commonly used to study seismic sources, such as seismic sources of ambient noises, rupture processes of large earthquakes, and local

earthquake monitoring. For ambient noise sources, the wavefield received by a seismic network can be assumed to be the superposition of multiple plane waves. By shifting and stacking the recordings in the seismic network to grids of different incident azimuths and slownesses, the noise sources outside the network can be imaged. Because this approach is based on the assumption of plane waves, it is generally called beamforming (e.g., Nishida, 2017; L  er *et al.*, 2018). For imaging the rupture of a large earthquake, the range of the rupture can be hundreds to thousands of kilometers. Typically, the teleseismic recordings of a large earthquake in a dense array can be backprojected to the gridded source area to image the rupture process. Because of the conceptual intuitiveness of the backprojection and the few assumptions about the fault geometry and rupture speed, the backprojection is widely used to study the rupture process of major earthquakes (e.g., Ishii *et al.*, 2005; Kiser and Ishii, 2017). Unlike the rupture imaging applications, the backprojection in local earthquake monitoring usually assumes that earthquakes are point sources and located in or near the region covered by the seismic network. Just like you can find the source of ripples in the water by rewinding, in most applications, the functions (e.g., envelope, STA/LTA, or kurtosis) of seismic recordings in the network that may reflect the arrivals of *P* and *S* waves are backprojected according to *P* and/or *S* wave speeds of the gridded study area to find potential earthquake sources (e.g., Langet *et al.*, 2014; Tan *et al.*, 2019; Lee, Liao, *et al.*, 2020). By selecting the corresponding *P* and *S* arrivals of potential earthquakes, the earthquake source parameters can be further improved. The backprojection is simple in logic and provides a more comprehensive consideration for the recordings in a seismic network, but there remain considerable challenges in the applications. First, the functions used for backprojection in previous studies are not the time functions of *P* and *S* arrivals. Those functions mainly reflect changes in seismograms, including the factors other than seismic phases such as random noises or recording gaps. This may lead to false detections or large uncertainties on source parameter estimates. Second, when the waveform changes caused by small earthquakes are not significant enough or the waveform changes are large during a high seismicity period, those functions often fail to reflect the phase arrivals of earthquakes. Therefore, the earthquake detection capability of backprojection is limited because of the drawbacks of the backprojected functions. Furthermore, in previous studies, the backprojection is usually applied to monitor earthquakes in a small area (e.g., Langet *et al.*, 2014; Tan *et al.*, 2019). To ensure consistency, all functions need to be backprojected to all grid points of a study area. When the study area and the number of stations are large, the required computational resources increase significantly.

In seismology, the earthquake detection and seismic phase picking have also benefited from recent significant development in deep learning algorithms (e.g., Ross *et al.*, 2018; Zhu and Beroza, 2018; Mousavi *et al.*, 2020; Liao *et al.*, 2021). Many

studies have shown that deep learning-based methods have high accuracy and stable results in both earthquake detection and seismic phase picking. An essential difference between traditional methods and deep learning methods is that the deep learning methods use the features extracted from the training data rather than empirical criteria based on processed seismograms (e.g., STA/LTA) selected by users. However, most deep learning methods use confirmed earthquake recordings to evaluate model performances. When applying deep learning-based models to continuous recordings, different situations need to be considered. Most of the time, continuous data mostly record background noise, but when the seismic activity is high, the waveforms of earthquakes with various magnitude and occurred close in origin times may dominate the recordings. In the deep learning methods, features are not only limited to the targets (e.g., earthquake signal, *P*, and *S* phases) but also include features of nontargets (e.g., background noises and other seismic waves), so the abnormal recordings (e.g., interrupted recordings) can be included in training data to reduce the probability of false detections. In addition, deep learning has significant advantages in dealing with multiple tasks and complex situations by including additional tasks in the model and data of complex cases in the training dataset.

In this study, we have developed a workflow that combines the advantages of deep learning in seismic phase picking and the straightforward principles of backprojection for earthquake monitoring (Fig. 1). Our updated attention recurrent residual U-Net (ARRU) phase picker is used for detecting earthquake signal, and picking *P* and *S* arrivals in the workflow. Data augmentation techniques have been applied to our training dataset to improve the performance on picking seismic phases on recordings with multiple earthquakes. Our backprojection of *P* and *S* arrival probability functions has increased the capability of detecting small earthquakes and reduced the demand for computing resources for earthquake monitoring in a large area using a large number of recordings. The workflow has been successfully applied to southern California (Fig. 2) and provides a more complete earthquake catalog during a high seismicity period.

Methods

Multitask ARRU model

Identifying the earthquake waveforms on continuous seismic recordings and picking available *P* and *S* phase arrivals of an earthquake are essential for earthquake location. We, therefore, improved the ARRU phase picker (Liao *et al.*, 2021) for the earthquake identification and seismic phase picking tasks. The ARRU phase picker is an encoder–decoder model, so the features of seismic recordings could also be used for detecting waveforms of earthquakes. The earthquake waveform identification module could be incorporated by simply branching a convolution block from the last recurrent residual convolution block (Fig. 3a). The truncated Gaussian functions for *P* and *S*

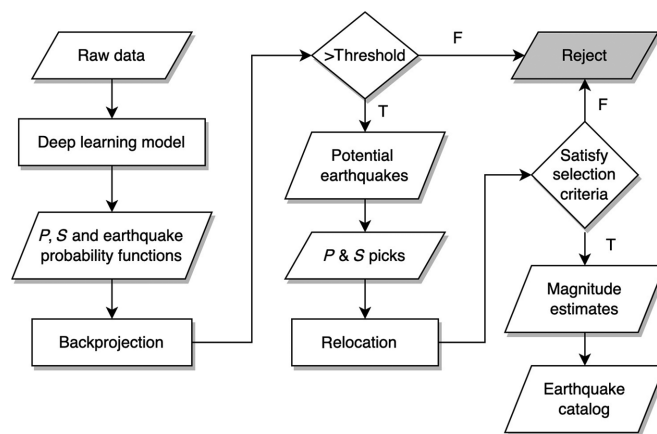


Figure 1. Workflow of our earthquake monitoring system.

arrivals (Liao *et al.*, 2021) and box-like functions wrapping the *P* and *S* arrivals (earthquake signal) are used as labels in training data (Fig. 3b). The model inputs are three-component seismograms, and outputs are point-to-point probability density functions of earthquake signals and seismic phase arrivals.

Model training for continuous recordings

For earthquake monitoring, a reliable machine learning-based phase picker must perform well both on continuous recordings with only background noise and with seismic phases of multiple earthquakes. When there is only background noise, the false detection rate of a model must be low; when there are multiple earthquakes, a model should be able to distinguish the phases of different events. The training, validation, and test datasets of most machine learning-based seismic phase picking models are earthquake recordings containing only one single pair of *P* and *S* arrivals, so the seismic phase detection capability on waveforms containing multiple earthquakes may be limited. To improve the performance of our ARRU phase picker in various situations, we included waveforms of background noise recordings and applied data augmentation to earthquake recordings. The data augmentation method has been widely used in machine learning to increase the complexity and variations of data when labeled training data are limited. In addition, models trained using complex datasets usually show better performance at various situations. The data augmentation method has been adopted for seismic phase picking and shown improvements on model performance (Mousavi *et al.*, 2020; Zhu *et al.*, 2020).

In this study, we select the earthquake recordings with signal-to-noise ratio (SNR) greater than 3 for data augmentation. We estimated the SNR on high-passed waveforms at 2 Hz, but using raw waveforms in training, validation, and test. Here, the waveforms spanning from 0.5 s before the *P* arrival to 5 s after the *S* arrival are defined as earthquake recordings, and the waveforms 10 s before the *P* arrival are treated as background

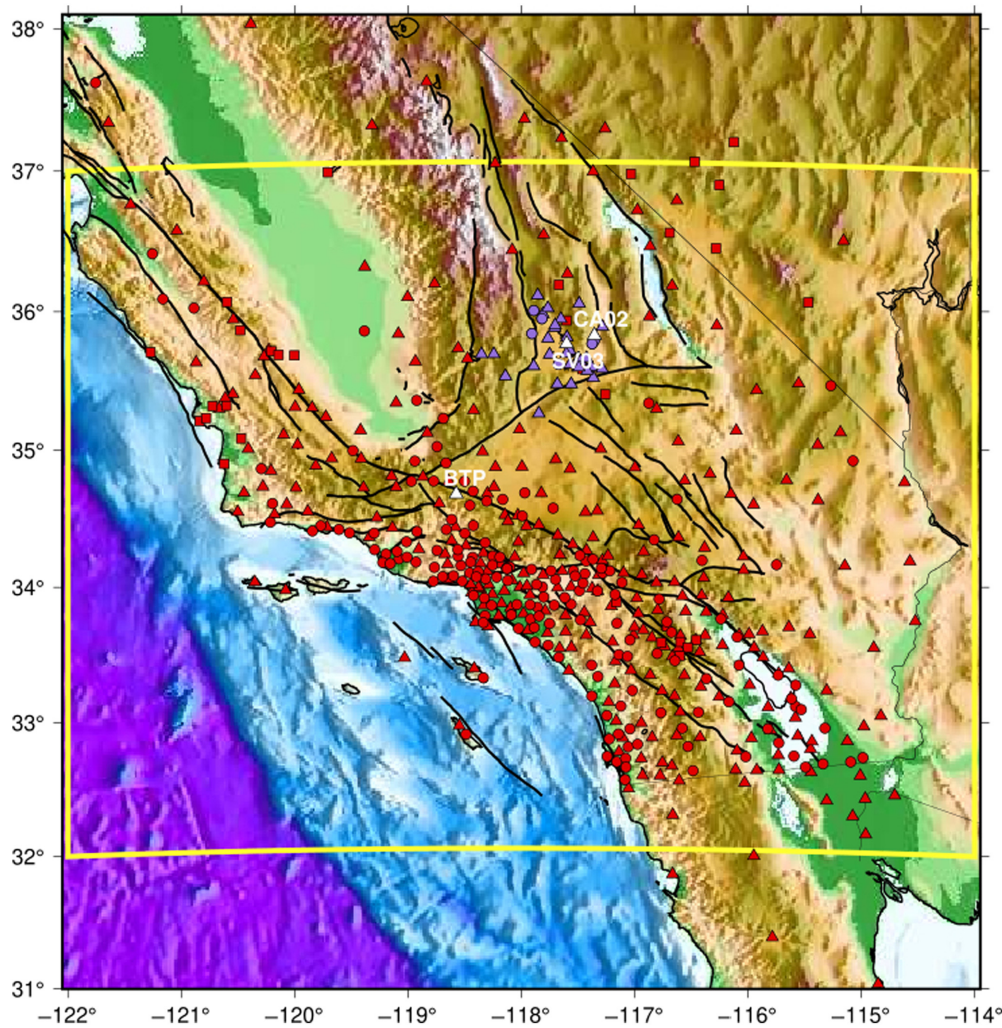


Figure 2. Distribution of the 507 stations used for earthquake location in this study. Triangles, broadband stations; circles, accelerometer; squares, short period stations; box, region for earthquake monitoring; and black-solid lines, major faults. The color version of this figure is available only in the electronic edition.

noise. To generate semisynthetic recordings of multiple earthquakes, data augmentation techniques, including random shift, superposition, and scaling, have been applied to earthquake waveforms recorded at the same station. We randomly select recordings of an earthquake as the first waveform, and the other earthquake waveforms are superimposed on it. A semisynthetic waveform of multiple earthquakes can be denoted as

$$W^c(t) = \{W_1^c(t) + \sum_i a W_i^c(t_{\tau_i}) : t_{\tau_i} \in [x_i, y_i]\}, \quad (1)$$

in which c is the component of seismogram; a is a random scaling factor between 0.5 and 1.0. The scaling factor is multiplied on the i th earthquake recording, and the SNR of the scaled earthquake waveforms must be larger than 3 to ensure that it is discernible from background noise. Let t_p and t_s

denote the P and S arrivals. The i th superimposed earthquake waveform starts at $x_i = t_p - 0.5 - \alpha_i$ and ends at $y_i = t_s + 0.5 * (t_s - t_p) + \beta_i$. We keep at least 0.5 s before P arrival and 0.5 times of $(t_s - t_p)$ to make sure the earthquake waveforms are recognizable during data augmentation processes. The α_i and β_i are random time shifts between 0 and 1.5 s and $2 \leq i \leq 4$ (events). The length of semisynthetic waveforms is 30 s, and we then truncated three 20 s waveforms with random time shifts between 1.0 and 7.0 s for datasets.

We used the recordings of earthquakes in Taiwan from 2012 to 2017 to train and validate our ARR model. There are 250 K earthquake recordings used as training dataset and 50 K earthquake recordings used as validation dataset. About 211 K earthquake recordings in Taiwan from 2018 to 2019 are used as a test dataset. The Z-score standardization, which removes the mean and then divides the waveform by the standard deviation of the waveform, is applied to all training, validation, and test data to reduce the variances in amplitude (Liao *et al.*, 2021).

We trained two models using the waveforms in the training dataset. Model A was trained with data containing waveforms from single earthquake. Model B was trained using augmented data containing waveforms from two to four different events and various scaling factors for waveforms from different events.

To confirm that the model trained by augmented data has not deteriorated on detecting seismic phases on recordings of P and S phase pair from single event, we applied the two models to the test dataset of Taiwan (Table 1), and the performances of the two models are at the same level. A global waveform dataset, STEAD (Mousavi *et al.*, 2019), has also been used to evaluate the generalization capabilities of the two models (Table 1). The result shows that the two models have good model generalization and similar capabilities on detecting seismic phases of recordings of single earthquake.

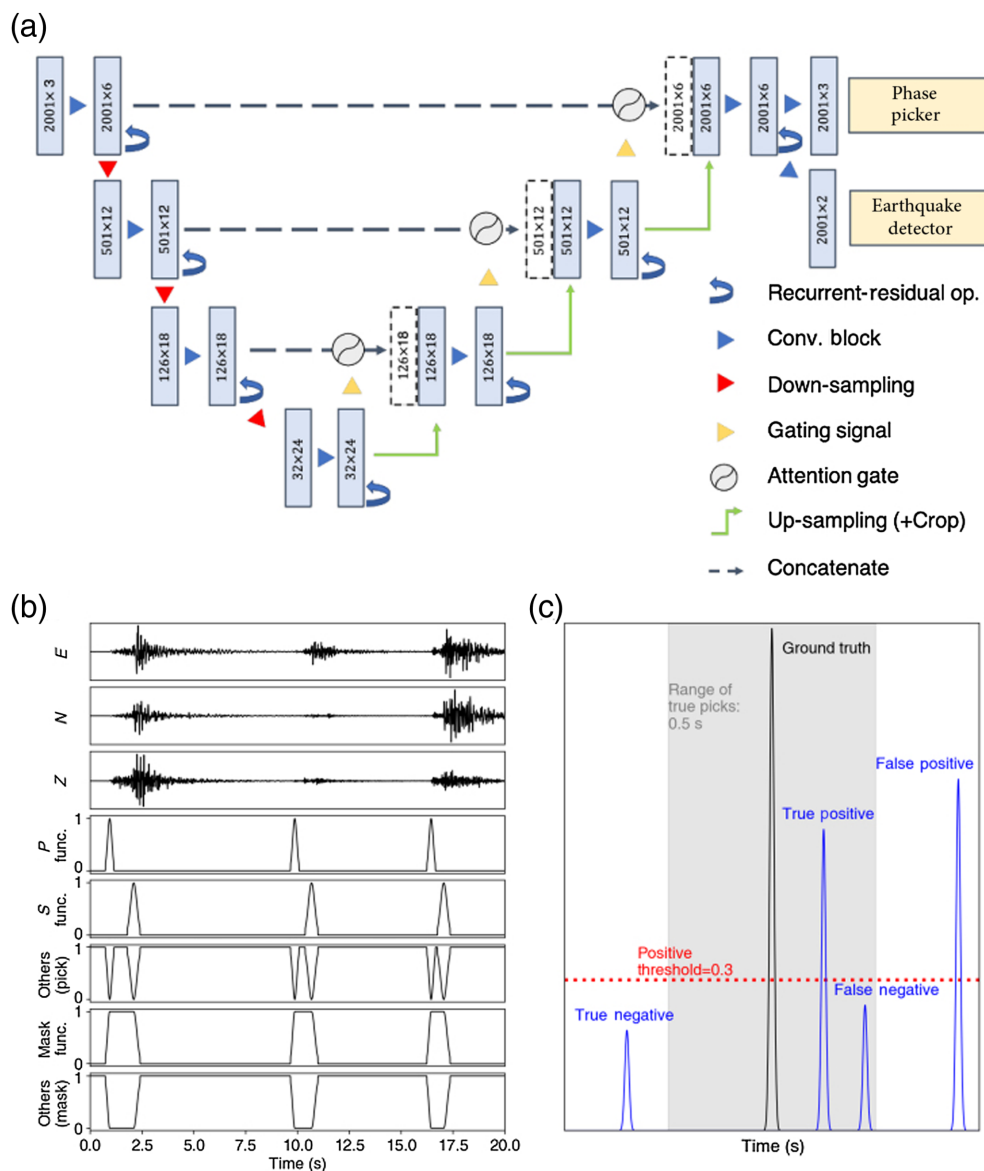


Figure 3. (a) Schematic diagram of our updated attention recurrent residual U-Net (ARRU) model used in this study. The ARRU model is adapted to multitask learning model. The model receives seismograms in shape of 2001×3 and outputs 5 vectors in total that represent the probability functions of *P*, *S*, and others phases and earthquake and other masks. (b) An example of semi-synthetic waveforms and the corresponding target functions. The model optimization relies on the weighted softmax loss separately estimated from target functions of phase arrivals and detection masks. (c) The criteria and classification of model predicted picks in confusion matrix. Here, the true pick is counted when the arrival residual between the labeled arrival and predicted arrival is less than 0.5 s. A positive pick is with predicted probability larger than 0.3. The color version of this figure is available only in the electronic edition.

To evaluate performances of the two models on waveforms of multiple earthquakes, data augmentation techniques have been applied to the test dataset of Taiwan to construct three 90 K waveform datasets of 2, 3, and 4 earthquakes. The true positive picks must satisfy that the phase probabilities larger than 0.3 and time differences with the ground truth less than 0.5 s (Fig. 3c). True positive detection requires that the

probabilities of entire detection mask to be greater than 0.3. Only the detection rate, true positives divided by all samples, is available for earthquake detection of Taiwan dataset, because noise label is not included in the dataset to compute confusion matrix. Figure 4 shows the model performances of the two models on different test datasets. Clearly, model B that was trained by augmented waveforms shows much better performances on both phase picking and earthquake detection for waveforms of multiple earthquakes. The results suggest that the use of model B could provide more reliable phase picking when seismicity is high.

Backprojection

In previous backprojection-based applications, functions based on energy or waveform changes on seismograms (e.g., envelope, STA/LTA, or kurtosis) are treated as time functions of *P* and/or *S* arrivals, and are backprojected to find potential earthquakes (e.g., Liao *et al.*, 2012; Drew *et al.*, 2013; Langet *et al.*, 2014). However, such an assumption may have the following drawbacks that would limit the performance of backprojection in local earthquake monitoring: (1) Variations in these functions are not solely caused by seismic wave phases. Complex seismic wave propagation effects, random noises, or problematic seismic record-

ings can all cause changes in functions (Fig. 5a,b). Even the earthquake source radiation patterns can cause significant changes in the functions for the recordings on different components (Fig. 5a). (2) For small earthquakes or when earthquakes with different magnitudes occur close in time, those functions often perform poorly in detecting seismic phases (Fig. 5a). (3) Usually it is not clear from the functions that

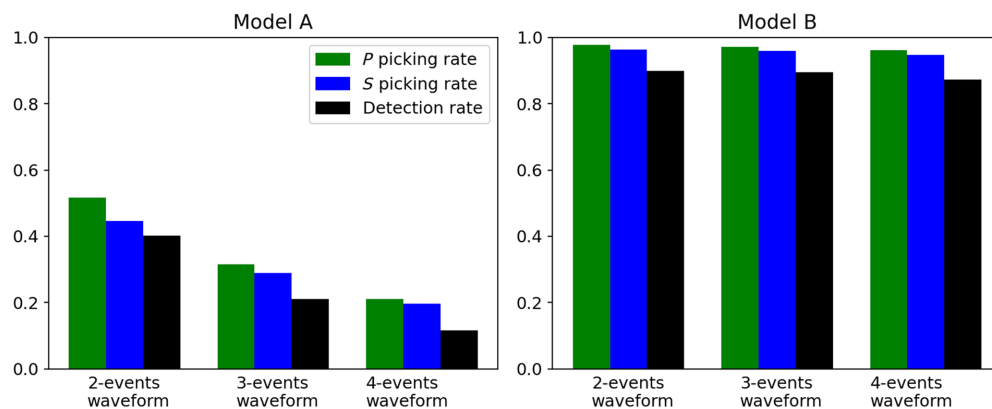


Figure 4. Model performances of model A (trained by recordings of single earthquake) and model B (trained by augmented recordings) on augmented test datasets of multiple events. The color version of this figure is available only in the electronic edition.

the changes are caused by *P* or *S* waves. If the possibilities of the functions changing in different components and *P* and *S* waves are all considered, the amount of computation will be greatly increased for backprojection. High computational demands limit the potential of backprojection for earthquake monitoring in a large region or a network with many stations. (4) The peak values of some functions (e.g., envelope and STA/LTA) are often not the arrivals of the *P* and *S* waves (Fig. 5a,b), thus

increasing the uncertainties of the source parameter estimations. (5) The value ranges of some functions are large (e.g., envelope and kurtosis), and it is difficult to set the threshold for selecting potential earthquakes.

Unlike previous studies, more reliable *P* and *S* arrival probability functions obtained by ARRUG in continuous recording are used for our backprojection application to improve the performance of local earthquake monitoring. The backprojection method used in this study for estimating potential earthquake location consists of the following steps:

1. Compute travel-time tables of grids. In this work, the study area, southern California, is divided into 4×4 km horizontally and 3 km vertically grids, and the CVM-S4.26 (Lee *et al.*, 2014) is used as the velocity model for computing phase travel times. The *P* and *S* travel times of a grid point

TABLE 1
Benchmarking Results on Taiwan and STEAD Test Datasets

STEAD	Model	Picking or Detection Rate	Mean (s)	Standard Deviation (s)	Precision	Recall	F1 Score	MAE	Test Samples
P phase	A	0.9854	−0.0102	0.0599	0.9963	0.9906	0.9934	0.0309	81 K EQ (20 s)
	B	0.9829	−0.0057	0.0631	0.9955	0.9896	0.9926	0.0347	
S phase	A	0.9687	−0.0169	0.1040	0.9832	0.9890	0.9861	0.0661	
	B	0.9655	−0.0002	0.1081	0.9821	0.9863	0.9842	0.0685	
Detection	A	0.9921	−	−	0.9944	0.9921	0.9933	−	81 K EQ and 23.5 K Nz (20 s)
	B	0.9915	−	−	0.9942	0.9915	0.9928	−	
Taiwan Data									
P phase	A	0.9914	0.0002	0.0405	0.9974	0.9948	0.9961	0.0205	211 K EQ
	B	0.9875	0.0011	0.0493	0.996	0.9927	0.9943	0.0274	
S phase	A	0.9781	0.0070	0.0927	0.9887	0.9918	0.9902	0.0583	
	B	0.9736	0.0297	0.1005	0.9844	0.9920	0.9882	0.0687	
Detection	A	0.9882	−	−	−	−	−	−	
	B	0.9828	−	−	−	−	−	−	

EQ, earthquake; MAE, mean absolute error; Nz, noise.

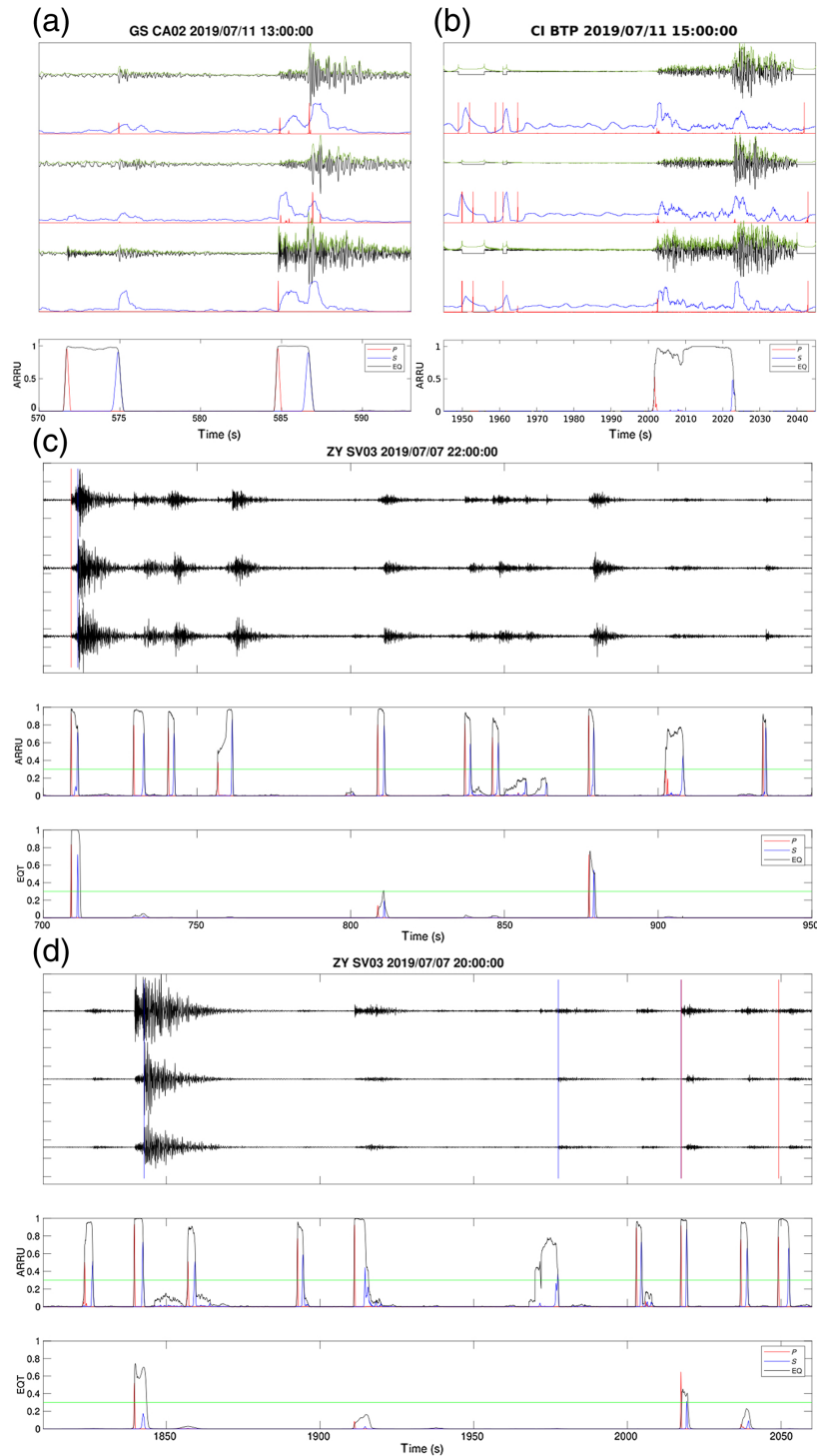


Figure 5. Examples of ARR model predictions compared with the results of other phase pickers. (a,b) Examples of three-component seismograms (black solid lines), their waveform envelopes (green solid lines), short-term average/long-term average functions (blue solid lines), positive time derivatives of kurtosis functions (red solid lines), and P (red) and S (blue) arrival predictions of ARR model. (c,d) Solid lines on seismograms indicate the P (red) and S (blue) phase arrivals picked by Southern California Earthquake Data Center (SCEDC) for earthquake location. Comparisons of P (red curves), S (blue curves), and earthquake signal (black curves) probability functions made using the ARR model and earthquake transformer (EQT). The green lines indicate the threshold for phase selection. The color version of this figure is available only in the electronic edition.

to its nearest 20 stations are calculated using simul2000 (Thurber and Eberhart-Phillips, 1999) and stored according to distance from the nearest station to the farthest station for backprojection.

2. Backproject probability functions of P and S arrivals. The probability functions of P arrivals, $\pi(t)$, S arrivals, $\sigma(t)$, and earthquake detection, $E(t)$, obtained from our ARR model are used for backprojection. To fully utilize the information of phase pairs, we regard a pair of P and S picks with a local maximum larger than a given threshold within a mask of $E(t)$ above another threshold as a phase pair. Here, the threshold values are all 0.3 for earthquake detection, P pick, and S pick. Before backprojecting the $\pi(t)$ and $\sigma(t)$ to a given grid point, we reject the phase pairs that are unlikely to be generated from the grid point to the station. When the observed and calculated P and S phase residuals are larger than a given time, we replace the $\pi(t)$ and $\sigma(t)$ in the earthquake detection mask, $E(t)$, written as $\pi'(t)$ and $\sigma'(t)$, with zero to reject the phase pairs that are unlikely to be generated from the grid to the station. The time difference threshold here is 0.5 s, but the threshold is related to the velocity model and grid sizes. Our backprojection results on a grid j can be expressed as follows:

$$B(j, t) = \sum_{k=1}^N \sigma'(t - t_{s_{jk}}) + \pi'(t - t_{p_{jk}}), \quad (2)$$

in which k indicates the available stations sorted from the nearest to the farthest in the time table of the grid j . The t_{sjk} and t_{pjk} are calculated P and S travel times from grid j to station k . The minimum number of stations required for backprojection is 8 in this study.

3. Find potential earthquakes. The backprojection results, $B(j,t)$, store the summation of P and S arrival probabilities close to the grid point. When a local maximum on a grid point is larger than a given threshold (2.0 in this study), the grid point can then be considered as a potential hypocenter. The time point of local maximum is taken as the origin time of the potential earthquake (Fig. 6e,f). To avoid grid points that are close to the potential hypocenter being selected, we also set thresholds in space and time for selecting the potential earthquakes (Fig. 6). The thresholds are related to the velocity model and grid size used in the backprojection to some extent. In this study, the largest stacking values of grids with time differences in 2.0 s and location difference in 20 km are selected as potential earthquakes.

Refine phase picks, earthquake relocation, and magnitude estimates

We calculate the P and S arrivals of potential earthquakes for all stations in the network as references for selecting seismic phases of the events. In this study, the local maximum of phase probabilities larger than 0.3 and time residuals of the P and S arrivals within 1.0 and 1.2 s, respectively, are selected. The Hypoinverse (Klein, 2002) is used to relocate potential earthquakes with at least 8 P and/or S phases. We select earthquake location results with a location quality of C or higher to our earthquake catalog. If a phase arrival is used to locate multiple earthquakes, we only retain the earthquake with better location quality. Finally, the local magnitude, M_L , of qualified earthquakes are estimated based on Richter scale (Richter, 1935; Boore, 1989; Stein and Wyssession, 2002).

Results

Phase-picking results on continuous recordings

The improved ARRUPHASE picker trained by augmented data shows capabilities for detecting P and S arrivals and earthquake signals in various continuous recordings. Figure 5c,d shows the performances of our improved ARRUPHASE picker and the earthquake transformer (EQT) model (Mousavi et al., 2020) that was trained using the STEAD dataset, a global waveform dataset (Mousavi et al., 2019), on the continuous recordings of stations SV03 in southern California. EQT has also used data augmentation techniques in the training dataset to improve the model performance, and the waveforms are band-passed at 1–45 Hz for EQT (Mousavi et al., 2020). If the data augmentation techniques used in this study are also applied to the training data of EQT, it will be possible to improve the detection ability of EQT in the seismic phases of small earthquakes.

The SV03 station was deployed by U.S. Geological Survey after the 2019 Ridgecrest earthquake (Cochran et al., 2020). The training dataset of the two models did not contain the recordings of the stations, so the results also demonstrate the generalization of the two models. In the case of multiple small earthquakes that occurred in a short period of time, the ARRUPHASE model can identify most earthquakes, and that can be recognized by the naked eye (Fig. 5c). When there are many small earthquakes occurred before and after a larger earthquake, the phases of small events are challenging to identify using traditional methods (e.g., STA/LTA). The ARRUPHASE phase picker can identify most small earthquakes and is less affected by the amplitude of the large earthquake (Fig. 5d). When using traditional methods (e.g., STA/LTA, kurtosis) for phase picking, the unexpected changes that are not caused by earthquakes in continuous recordings, such as padding zeros when data missing, can easily cause false picks. Deep learning-based methods (e.g., ARRUPHASE, EQT) can include exceptions to the training dataset to reduce the probability of false phase picks (Fig. 5b).

Comparison of earthquake catalogs

Our workflow has been applied to southern California where the number of stations is as high as 500, and the earthquake monitoring area is about 500×600 km (Fig. 2). We applied the workflow to the available continuous seismic recordings in July 2019 that included the 2019 Ridgecrest earthquake sequence to demonstrate the capability of earthquake monitoring during a high seismicity period. The two catalogs have high consistency in the distribution of earthquakes, mostly concentrated in the Ridgecrest regions and the Peninsular Ranges (Fig. 7a,b). In the Peninsular Ranges, the distribution of earthquakes in our catalog coincides with the San Jacinto fault (Fig. 7g). The number of earthquakes detected by our workflow is more than three times that of the Southern California Earthquake Data Center (SCEDC) Catalog (Fig. 7c). Most of the earthquakes in Ridgecrest regions are aftershocks of the 2019 M_w 6.4 and 7.1 earthquakes or earthquakes triggered by the mainshock. Our results provided a more complete catalog of aftershocks and triggered events, which may be attributed to our ARRUPHASE phase picker that can identify more seismic phases of small earthquakes (Fig. 5c,d). The results show that our workflow can be used to monitor seismic activity in a large area with a large number of seismometers, and it can still provide reliable seismicity monitoring when seismicity is high.

Events close in origin times and/or hypocenters

Our ARRUPHASE backprojection method has the potential to locate earthquakes that are likely unidentified using traditional methods. In earthquake location, it is challenging to locate earthquakes with close origin times, and hypocenters due to the difficulties in phase picks and phase associations. In our method, when most of the seismic phases of earthquakes have been identified, the backprojection could associate the

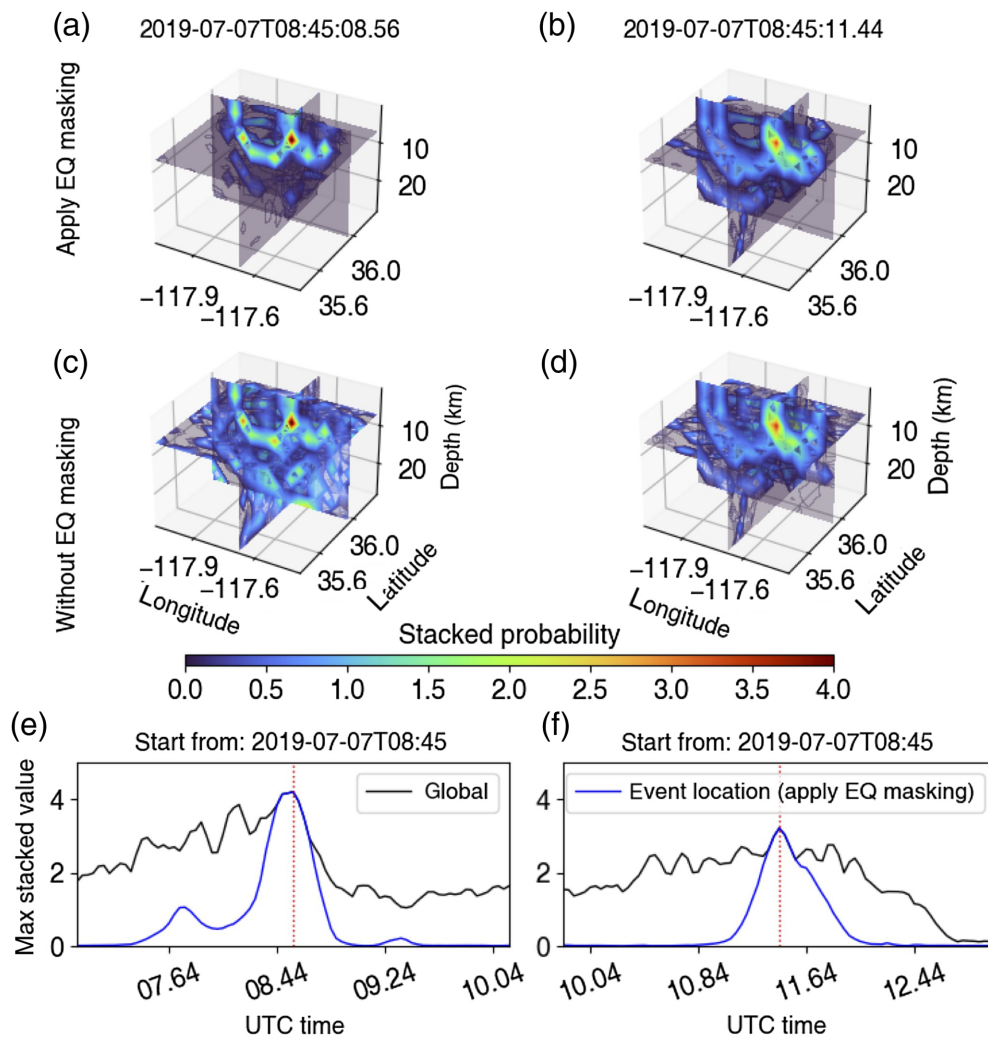


Figure 6. (a,b) Snapshots of the results of our backprojection with earthquake mask applied at the initial origin time of events 1 and 2. (c,d) The same snapshots as panels (a) and (b) for events 1 and 2 but the backprojection without earthquake mask applied. The color bar indicates the stacked probability. (e,f) The black curves indicate the maxima of global stacking values; the blue curves are the stacking values at the optimal grid point of events 1 and 2. The red dotted line indicates the optimal earthquake origin time estimate. The color version of this figure is available only in the electronic edition.

probabilities of phases to find the potential hypocenters. Figure 8c shows an example of earthquakes with close hypocenters, and the origin time difference is 3.08 s detected by our method (origin times of events 1 and 2 are 7 July 2019 08:45:08.72 and 7 July 2019 08:45:11.80, respectively). These two earthquakes are not included in the SCEDC catalog. In traditional backprojection based method for local earthquake monitoring, usually only the location with the maximum stacking value at the time point is considered a potential earthquake (e.g., Langet *et al.*, 2014). For the earthquakes that occurred at about the same time, even if the two earthquakes have a discernible distance, one of them may not be detected. Figure 8d,e shows two earthquakes with origin time differences less than 0.1 s (origin time of the event 3 on Fig. 8d is 6 July

2019 09:36:39.20, and that of event 4 on Fig. 8e is 6 July 2019 09:36:39.27), and the location difference that is about 300 km have been identified by our method. Of the two, only the larger earthquake in the Ridgecrest region is listed in the SCEDC catalog.

Discussion

The application of our workflow in southern California shows the capability of earthquake monitoring in a large area under high seismicity. The following are the advantages of combining our ARRUPHASE picker and backprojection in earthquake monitoring.

1. The use of individual *P* and *S* probability functions can improve the accuracy and computational efficiency of backprojection. When using the backprojection for detecting potential earthquakes, if the cause of the waveform changes (e.g., energy) is not known to be a *P* or *S* wave, the shift-and-stack processes need to be applied to the recordings of three components for both *P* and *S* travel times to consider all the possibilities (Lee, Liao, *et al.*, 2020). When the specific

inputs of *P* and *S* waves are known, the shift-and-stack only needs to be applied to the functions of *P* and *S* waves according to their travel times. Fewer shift-and-stack operations in backprojection not only improve the calculation efficiency but also improve the accuracy of earthquake detection due to the reduction of waveforms with incorrect phase shifts on stacking results.

2. The fixed range (between 0 and 1) probabilities provide same weights for different phases in backprojection. Different from other data processing results (e.g., envelope or kurtosis of seismograms) that are greatly affected by various factors (e.g., earthquake magnitude, source-receiver distance, and background noise), our ARRUPHASE picker provides probabilities between 0 and 1 of *P* and *S* arrivals, so that the weights of

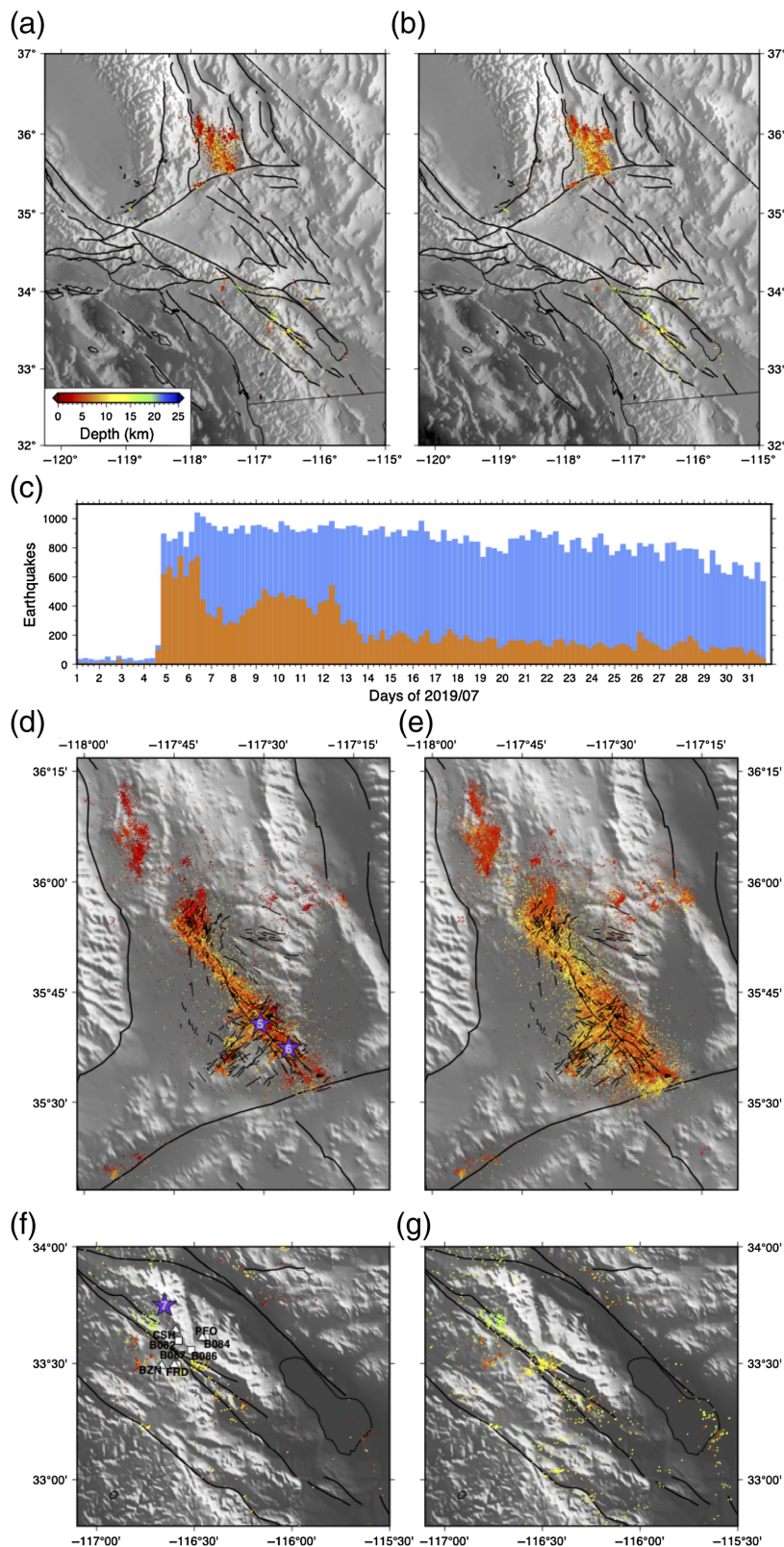


Figure 7. (a,d,f) Earthquake distributions of SCEDC and (b,e,g) our catalogs in July 2019. (c) Histograms show the number of earthquakes in SCEDC (orange bars) and our (blue bars) catalogs. Black lines, major faults, and surface fractures of the 2019 Ridgecrest earthquakes (Xu et al., 2020). The color version of this figure is available only in the electronic edition.

phase arrivals at different stations are the same. Therefore, when using backprojection to evaluate potential earthquakes at each grid point, only the same number of phase arrival functions are required to be stacked to achieve consistency, and there is no need to shift-and-stack the phase arrival functions of all the stations to each grid point. This not only greatly reduces required calculations but also opens the possibilities of parallel computing and earthquake monitoring in large area. The fixed probability range also allows the threshold value to be set according to the expected stacking value of the *P* and *S* phases of an earthquake, which is different from the difficulty in setting a suitable threshold of back-projections using other data (e.g., envelope). In addition, the use of the data close to a grid point in backprojection is also more advantageous for the monitoring of small earthquakes and earthquakes that occur close in time.

3. Excluding improper *P* and *S* pairs in backprojection can increase the accuracy of earthquake detection (Fig. 6a,b). When earthquakes occur continuously in a short period of time (e.g., swarms or aftershocks), if directly backproject all probability functions of *P* and *S* arrivals, some high values in stacking results may not be earthquakes (Fig. 6c–f). Removing the *P* and *S* pairs determined by ARRUPHASE picker that are unlikely from a given grid to the station can increase the accuracy in earthquake detection.

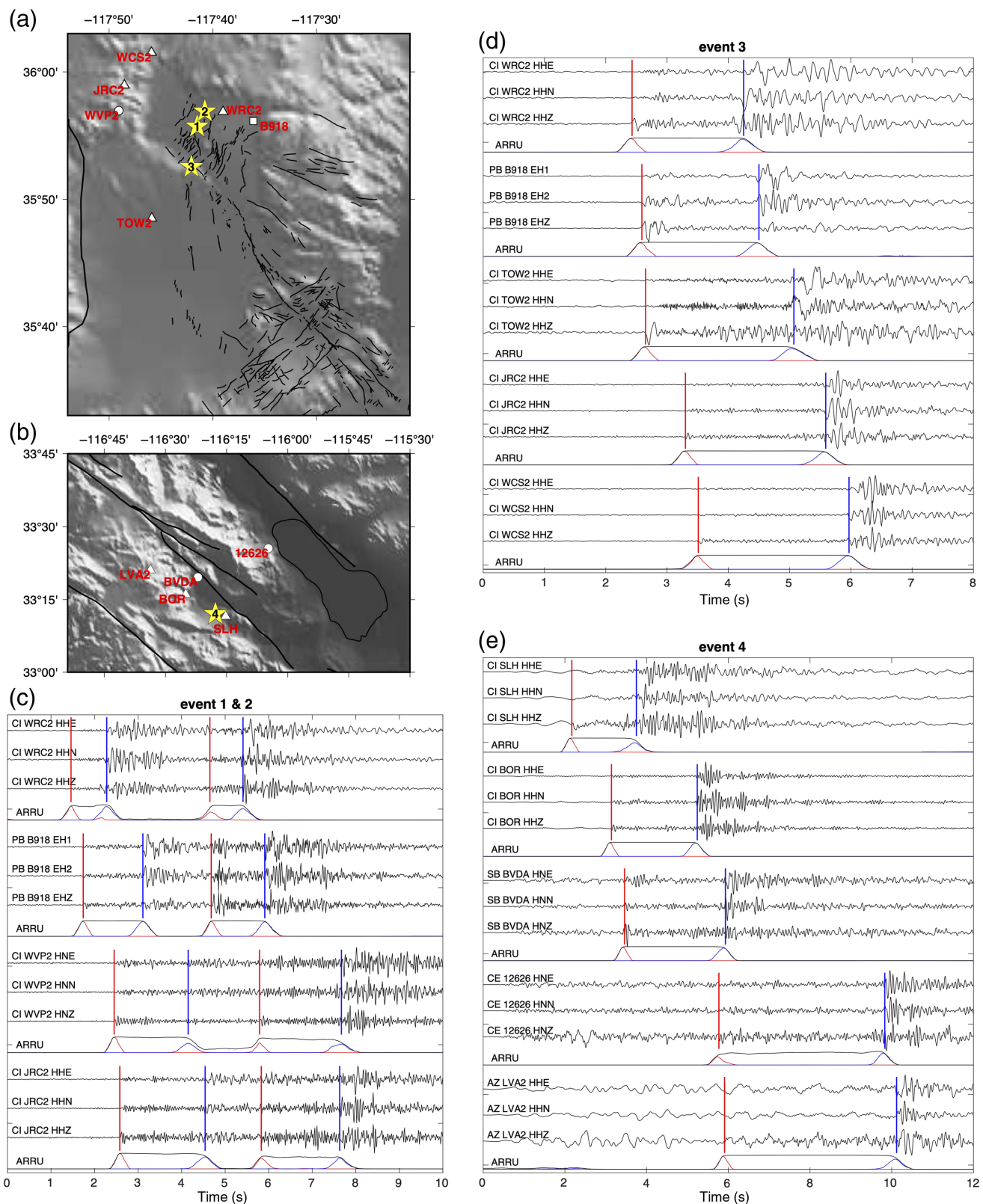


Figure 8. (a,b) Maps of examples of earthquakes and stations. (c) Seismograms of events 1 and 2 with close origin times and locations detected by our method but missed in SCEDC catalog. (d,e) seismograms of events 3 and 4 occurred almost simultaneously, but the locations differed by 300 km detected by our

workflow. Solid lines on seismograms indicate the *P* (red) and *S* (blue) arrivals picked for earthquake location. The rows of ARRUs show the *P* (red curves), *S* (blue curves), and earthquake signal (black curves) probability functions predicted by our ARRUs model. The color version of this figure is available only in the electronic edition.

Our ARRUC phase picker has high computational efficiency and has the potential for near real-time earthquake monitoring. It only takes 0.73 s to process 60 s of continuous seismograms using our ARRUC model using NVIDIA RTX 2080Ti. Taking backprojection in the southern California as an example (in a case of 480 available stations and 181,500 grid points), it takes 1657.29 s to backproject 1 hr data with 1 core and 456.12 s on 6 cores of the Intel Xeon Silver 4210R CPU. Using the results of ARRUC phase picker to perform backprojection, each grid only needs the results of its surrounding stations rather than all stations, which is conducive to parallel computing. Rapid and accurate earthquake monitoring is not only helpful for studies of earthquake physics but also for disaster relief after earthquakes and can even be used to assess the potential of large earthquakes. For example, a *b*-value based method for near real-time evaluates the possibilities of another major earthquake after a large earthquake has also been successfully applied to the 2019 Ridgecrest sequence (Gulia *et al.*, 2020).

The earthquakes determined by our method and the earthquakes in SCEDC catalog have high consistency in the distributions (Fig. 7). However, our earthquake distribution is more scattered than that in SCEDC catalog in the Ridgecrest region (Fig. 7d,e). This may be due to our method's higher ability to detect small earthquakes (Fig. 8). The scattered distribution of earthquakes in the intersection area of surface ruptures of the M_w 6.4 event on 4 July 2019, and the later M_w 7.1 Ridgecrest earthquake (Fig. 7e) may reflect the damages in fault zones. The surface fractures of the 2019 Ridgecrest earthquakes obtained by Interferometric Synthetic Aperture Radar (Xu *et al.*, 2020) also show a dense fracture in the region.

When monitoring earthquakes in large areas or areas with complex geological structures, seismic phases that propagate through different paths may be misjudged and used to cause errors in earthquake location. For example, Figure 9a shows an earthquake that occurred in the Ridgecrest region (SCEDC event ID: 38488719, M_w 3.83), and the *Pn* and *Pg* phases were recorded at a station in the Anza region about 300 km away. Because the *S* and surface waves arrive late, the *Pn* and *Pg* may be wrongly judged as other wave phases and used for earthquake location. Figure 9b shows the seismograms in the Anza region of an earthquake (SCEDC event ID: 38488767, M_L 3.48) that occurred in Ridgecrest, and the *Pg* phase was incorrectly identified as *P* arrivals of the other earthquake (SCEDC event ID: 38488775, M_L 1.60) in the Anza area.

Conclusion

Because most seismic networks currently record ground motions continuously, a reliable automated workflow is needed for earthquake monitoring on continuous recordings. In this study, the advantages of our automated workflow can be summarized as follows:

1. We have improved our ARRUC model for seismic phase picking and earthquake signal detection. The results of earthquake detection can be used to pair the *P* and *S* arrivals of the same earthquake.
2. The use of data augmentation techniques has significantly improved the capability of ARRUC phase picker in detecting multiple earthquakes on continuous recordings.
3. The combination of accuracy probability functions of our ARRUC phase picker and travel time-based backprojection has improved significantly the spatial and temporal resolution of earthquake monitoring.
4. Our workflow has been successfully applied to the recordings of more than 500 seismic stations throughout southern California (about 500 × 600 km) and shows the capability of earthquake monitoring in a large area.
5. During periods of high seismicity, our workflow can provide a more complete earthquake catalog. Our workflow has detected many earthquakes missed during the 2019 Ridgecrest sequence, and the number of earthquakes in our catalog is more than three times than that in SCEDC catalog.

Data and Resources

The earthquake recordings in Taiwan used in this work are from the Geophysical Database Management System (GDMS) operated by the Central Weather Bureau (CWB) and the Broadband Array in Taiwan for Seismology (BATS) operated by the Institute of Earth Sciences, Academia Sinica (IES). The Southern California Earthquake Data Center (SCEDC) earthquake catalog website is <https://scedc.caltech.edu/data/eq-catalogs.html> (last accessed October 2021). The seismograms in southern California used in this study were obtained from SCEDC through the Seismogram Transfer Program (STP), and the website of STP is <https://scedc.caltech.edu/data/downloads.html> (last accessed October 2021). The source code and model of our attention recurrent residual U-Net (ARRUC) phase picker, source code and examples of our backprojection, and our southern California earthquake catalog are available for download from https://github.com/tso1257771/ARRUC_seismic_backprojection (last accessed February 2022).

Declaration of Competing Interests

The authors acknowledge that there are no conflicts of interest recorded

Acknowledgments

En-Jui Lee and Wu-Yu Liao are supported by the Ministry of Science and Technology, Republic of China, under Contract MOST 109-2116-M-006-016. The authors would like to acknowledge Science College of National Cheng Kung University (NCKU Science) and Ministry of Science and Technology (MOST), Taiwan, Republic of China, for a fellowship to support Wu-Yu Liao's Ph.D. study. Po Chen acknowledges support from the Nielson Energy Fellowship provided by the School of Energy Resources, University of Wyoming. The authors thank Southern California Earthquake Data Center (SCEDC) for

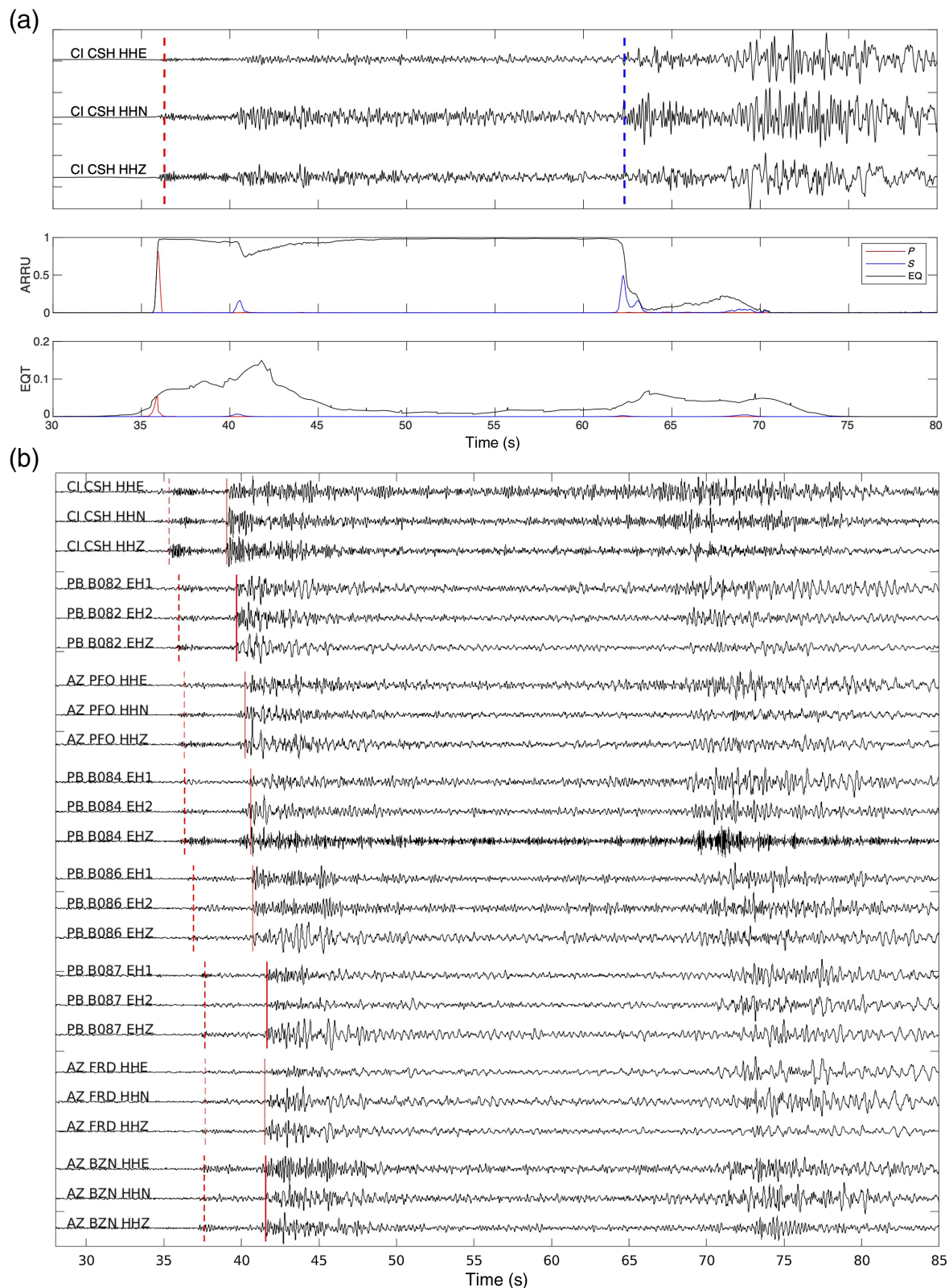


Figure 9. (a) Seismograms of the earthquake occurred in the Ridgecrest region (SCEDC event ID: 38488719, M_w 3.83, event 5 on Fig. 7d) recorded in the Peninsular Ranges region. Dashed lines on seismograms indicate the predicted first arrived P (red) and S (blue) phases. The middle and bottom panels show P (red curves), S (blue curves), and earthquake signal (black curves) probability functions made using the ARR model and EQT. (b) Seismograms

of the earthquake occurred in the Ridgecrest region (SCEDC event ID: 38488767, M_L 3.48, event 6 on Fig. 7d) recorded in the Peninsular Ranges region. The red dashed lines indicate predicted first P arrivals of the event at the stations, and the red lines are P picks used for locating the earthquake (SCEDC event ID: 38488775, M_L 1.60, event 7 on Fig. 7f) in SCEDC catalog. The color version of this figure is available only in the electronic edition.

providing seismic recordings for this study. This work utilizes resources supported by the National Science Foundation's Major Research Instrumentation program, Grant Number 1725729, as well as the University of Illinois at Urbana-Champaign. The authors also thank National Center for High-performance Computing (NCHC) in Taiwan, for providing computational and storage resources.

References

- Akram, J., and D. W. Eaton (2016). A review and appraisal of arrival-time picking methods for downhole microseismic data, *Geophysics* **81**, no. 2, KS71–KS91, doi: [10.1190/geo2014-0500.1](https://doi.org/10.1190/geo2014-0500.1).
- Allen, R. (1982). Automatic phase pickers: Their present use and future prospects, *Bull. Seismol. Soc. Am.* **72**, no. 6B, S225–S242.
- 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](https://doi.org/10.1785/0120120347).
- Boore, D. M. (1989). The Richter scale: Its development and use for determining earthquake source parameters, *Tectonophysics* **166**, no. 1, 1–14, doi: [10.1016/0040-1951\(89\)90200-X](https://doi.org/10.1016/0040-1951(89)90200-X).
- Cochran, E. S., E. Wolin, D. E. McNamara, A. Yong, D. Wilson, M. Alvarez, N. van der Elst, A. McClain, and J. Steidl (2020). The U.S. Geological Survey's rapid seismic array deployment for the 2019 Ridgecrest earthquake sequence, *Seismol. Res. Lett.* **91**, no. 4, 1952–1960, doi: [10.1785/0220190296](https://doi.org/10.1785/0220190296).
- Drew, J., R. S. White, F. Tilmann, and J. Tarasewicz (2013). Coalescence microseismic mapping, *Geophys. J. Int.* **195**, no. 3, 1773–1785, doi: [10.1093/gji/ggt331](https://doi.org/10.1093/gji/ggt331).
- Gulia, L., S. Wiemer, and G. Vannucci (2020). Pseudoprospective evaluation of the foreshock traffic-light system in Ridgecrest and implications for aftershock hazard assessment, *Seismol. Res. Lett.* **91**, no. 5, 2828–2842, doi: [10.1785/0220190307](https://doi.org/10.1785/0220190307).
- Ishii, M., P. M. Shearer, H. Houston, and J. E. Vidale (2005). Extent, duration and speed of the 2004 Sumatra–Andaman earthquake imaged by the Hi-Net array, *Nature* **435**, no. 7044, 933–936, doi: [10.1038/nature03675](https://doi.org/10.1038/nature03675).
- Kato, A., and Y. Ben-Zion (2021). The generation of large earthquakes, *Nat. Rev. Earth Environ.* **2**, no. 1, 26–39, doi: [10.1038/s43017-020-00108-w](https://doi.org/10.1038/s43017-020-00108-w).
- Kiser, E., and M. Ishii (2017). Back-projection imaging of earthquakes, *Ann. Rev. Earth Planet. Sci.* **45**, no. 1, 271–299, doi: [10.1146/annurev-earth-063016-015801](https://doi.org/10.1146/annurev-earth-063016-015801).
- Klein, F. W. (2002). User's guide to HYPOINVERSE-2000, a Fortran program to solve for earthquake locations and magnitudes, *U.S. Geol. Surv. Open-File Rept. USGS Numbered Series 2002–171*, 123 pp., doi: [10.3133/ofr02171](https://doi.org/10.3133/ofr02171).
- Langet, N., A. Maggi, A. Michelini, and F. Brenguier (2014). Continuous kurtosis-based migration for seismic event detection and location, with application to Piton de la Fournaise Volcano, La Reunion, *Bull. Seismol. Soc. Am.* **104**, no. 1, 229–246, doi: [10.1785/0120130107](https://doi.org/10.1785/0120130107).
- Lee, E., W. Liao, D. Mu, W. Wang, and P. Chen (2020). GPU-Accelerated Automatic Microseismic Monitoring Algorithm (GAMMA) and its application to the 2019 Ridgecrest earthquake sequence, *Seismol. Res. Lett.* **91**, no. 4, 2062–2074, doi: [10.1785/0220190323](https://doi.org/10.1785/0220190323).
- Lee, E., D. Mu, W. Wang, and P. Chen (2020). Weighted Template-Matching Algorithm (WTMA) for improved foreshock detection of the 2019 Ridgecrest earthquake sequence, *Bull. Seismol. Soc. Am.* **110**, no. 4, 1832–1844, doi: [10.1785/0120200020](https://doi.org/10.1785/0120200020).
- Lee, E.-J., P. Chen, T. H. Jordan, P. B. Maechling, M. A. M. Denolle, and G. C. Beroza (2014). Full-3-D tomography for crustal structure in southern California based on the scattering-integral and the adjoint-wavefield methods, *J. Geophys. Res.* **119**, no. 8, 6421–6451, doi: [10.1002/2014JB011346](https://doi.org/10.1002/2014JB011346).
- Liao, W.-Y., E.-J. Lee, D. Mu, P. Chen, and R.-J. Rau (2021). ARRU phase picker: Attention recurrent-residual U-Net for picking seismic P- and S-phase arrivals, *Seismol. Res. Lett.* doi: [10.1785/0220200382](https://doi.org/10.1785/0220200382).
- Liao, Y.-C., H. Kao, A. Rosenberger, S.-K. Hsu, and B.-S. Huang (2012). Delineating complex spatiotemporal distribution of earthquake aftershocks: An improved source-scanning algorithm, *Geophys. J. Int.* **189**, no. 3, 1753–1770, doi: [10.1111/j.1365-246X.2012.05457.x](https://doi.org/10.1111/j.1365-246X.2012.05457.x).
- Löer, K., N. Riahi, and E. H. Saenger (2018). Three-component ambient noise beamforming in the Parkfield area, *Geophys. J. Int.* **213**, no. 3, 1478–1491, doi: [10.1093/gji/ggy058](https://doi.org/10.1093/gji/ggy058).
- Lomax, A., C. Satriano, and M. Vassallo (2012). Automatic picker developments and optimization: FilterPicker—a Robust, broadband picker for real-time seismic monitoring and earthquake early warning, *Seismol. Res. Lett.* **83**, no. 3, 531–540, doi: [10.1785/gssrl.83.3.531](https://doi.org/10.1785/gssrl.83.3.531).
- Mousavi, S. M., W. L. Ellsworth, W. Zhu, L. Y. Chuang, and G. C. Beroza (2020). Earthquake transformer—an attentive deep-learning model for simultaneous earthquake detection and phase picking, *Nat. Commun.* **11**, no. 1, doi: [10.1038/s41467-020-17591-w](https://doi.org/10.1038/s41467-020-17591-w).
- Mousavi, S. M., Y. Sheng, W. Zhu, and G. C. Beroza (2019). STANford Earthquake Dataset (STEAD): A global data set of seismic signals for AI, *IEEE Access* **7**, 179,464–179,476, doi: [10.1109/ACCESS.2019.2947848](https://doi.org/10.1109/ACCESS.2019.2947848).
- Mu, D., E.-J. Lee, and P. Chen (2017). Rapid earthquake detection through GPU-based template matching, *Comput. Geosci.* **109**, 305–314, doi: [10.1016/j.cageo.2017.09.009](https://doi.org/10.1016/j.cageo.2017.09.009).
- Nishida, K. (2017). Ambient seismic wave field, *Proc. Jpn Acad. Ser. B* **93**, no. 7, 423–448, doi: [10.2183/pjab.93.026](https://doi.org/10.2183/pjab.93.026).
- Peng, Z., and P. Zhao (2009). Migration of early aftershocks following the 2004 Parkfield earthquake, *Nat. Geosci.* **2**, no. 12, 877–881, doi: [10.1038/ngeo697](https://doi.org/10.1038/ngeo697).
- Richter, C. F. (1935). An instrumental earthquake magnitude scale*, *Bull. Seismol. Soc. Am.* **25**, no. 1, 1–32, doi: [10.1785/BSSA0250010001](https://doi.org/10.1785/BSSA0250010001).
- Ross, Z. E., B. Idini, Z. Jia, O. L. Stephenson, M. Zhong, X. Wang, Z. Zhan, M. Simons, E. J. Fielding, S.-H. Yun, et al. (2019). Hierarchical interlocked orthogonal faulting in the 2019 Ridgecrest earthquake sequence, *Science* **366**, no. 6463, 346–351, doi: [10.1126/science.aaz0109](https://doi.org/10.1126/science.aaz0109).
- Ross, Z. E., M.-A. Meier, E. Hauksson, and T. H. Heaton (2018). Generalized seismic phase detection with deep learning, *Bull. Seismol. Soc. Am.* **108**, no. 5A, 2894–2901, doi: [10.1785/0120180080](https://doi.org/10.1785/0120180080).
- Ross, Z. E., D. T. Trugman, E. Hauksson, and P. M. Shearer (2019). Searching for hidden earthquakes in southern California, *Science* **364**, no. 6442, 767–771, doi: [10.1126/science.aaw6888](https://doi.org/10.1126/science.aaw6888).
- Shelly, D. R., and D. P. Hill (2011). Migrating swarms of brittle-failure earthquakes in the lower crust beneath Mammoth Mountain,

- California: Lower-crustal earthquake swarms, *Geophys. Res. Lett.* **38**, no. 20, doi: [10.1029/2011GL049336](https://doi.org/10.1029/2011GL049336).
- Stein, S., and M. Wyssession (2002). *An Introduction to Seismology, Earthquakes and Earth Structure*, Wiley-Blackwell, Malden, Massachusetts.
- Tan, F., H. Kao, E. Nissen, and D. Eaton (2019). Seismicity-scanning based on navigated automatic phase-picking, *J. Geophys. Res.* **124**, no. 4, 3802–3818, doi: [10.1029/2018JB017050](https://doi.org/10.1029/2018JB017050).
- Thurber, C., and D. Eberhart-Phillips (1999). Local earthquake tomography with flexible gridding, *Comput. Geosci.* **25**, no. 7, 809–818, doi: [10.1016/S0098-3004\(99\)00007-2](https://doi.org/10.1016/S0098-3004(99)00007-2).
- Walter, J. I., P. Ogwari, A. Thiel, F. Ferrer, and I. Woelfel (2021). easyQuake: Putting machine learning to work for your regional seismic network or local earthquake study, *Seismol. Res. Lett.* **92**, no. 1, 555–563, doi: [10.1785/0220200226](https://doi.org/10.1785/0220200226).
- Weber, B., J. Becker, W. Hanka, A. Heinloo, M. Hoffmann, T. Kraft, D. Pahlke, J. Reinhardt, J. Saul, and H. Thoms (2007). SeisComP3—Automatic and interactive real time data processing, *Geophysical Research Abstracts*, Vol. 9, no. 9, General Assembly European Geosciences Union (EGU), Vienna, Austria, 2007, 219 pp.
- Xu, X., D. T. Sandwell, and B. Smith-Konter (2020). Coseismic displacements and surface fractures from Sentinel-1 InSAR: 2019 Ridgecrest earthquakes, *Seismol. Res. Lett.* **91**, no. 4, 1979–1985, doi: [10.1785/0220190275](https://doi.org/10.1785/0220190275).
- Yeck, W. L., J. M. Patton, C. E. Johnson, D. Kragness, H. M. Benz, P. S. Earle, M. R. Guy, and N. B. Ambruz (2019). GLASS3: A standalone multiscale seismic detection associator, *Bull. Seismol. Soc. Am.* **109**, no. 4, 1469–1478, doi: [10.1785/0120180308](https://doi.org/10.1785/0120180308).
- Zhou, Y., H. Yue, L. Fang, S. Zhou, L. Zhao, and A. Ghosh (2021). An earthquake detection and location architecture for continuous seismograms: Phase picking, association, location, and matched filter (PALM), *Seismol. Res. Lett.* doi: [10.1785/0220210111](https://doi.org/10.1785/0220210111).
- Zhu, W., and G. C. Beroza (2018). PhaseNet: A deep-neural-network-based seismic arrival time picking method, *Geophys. J. Int.* doi: [10.1093/gji/ggy423](https://doi.org/10.1093/gji/ggy423).
- Zhu, W., S. M. Mousavi, and G. C. Beroza (2020). Seismic signal augmentation to improve generalization of deep neural networks, *Adv. Geophys.* **61**, 151–177, doi: [10.1016/bs.agph.2020.07.003](https://doi.org/10.1016/bs.agph.2020.07.003).

Manuscript received 28 September 2021

Published online 9 March 2022