JGR Solid Earth

RESEARCH ARTICLE
10.1029/2019JB017536

Key Points:
- PickNet can rapidly acquire vast high-quality seismic arrival times
- The picking accuracy by PickNet is close to that by human experts
- The picked arrival times can be used directly to study the Earth's interior structure

Supporting Information:
- Supporting Information S1
- Data Set S1
- Data Set S2
- Data Set S3

Correspondence to:
J. Wang and Z. Xiao,
jianwang@mail.iggcas.ac.cn;
xiaozhuowei@mail.iggcas.ac.cn

Citation:
Wang, J., Xiao, Z., Liu, C., Zhao, D., & Yao, Z. (2019). Deep learning for picking seismic arrival times. *Journal of Geophysical Research: Solid Earth*, *124*, 6612–6624. https://doi.org/10.1029/2019JB017536

Received 14 FEB 2019
Accepted 2 JUN 2019
Accepted article online 7 JUN 2019
Published online 3 JUL 2019

Deep Learning for Picking Seismic Arrival Times

Jian Wang[1,2,3], Zhuowei Xiao[1,4], Chang Liu[4], Dapeng Zhao[5], and Zhenxing Yao[1,2]

[1]Key Laboratory of Earth and Planetary Physics, Institute of Geology and Geophysics, Chinese Academy of Sciences, Beijing, China, [2]Institutions of Earth Science, Chinese Academy of Sciences, Beijing, China, [3]Laboratory for Marine Mineral Resources, Qingdao National Laboratory of Marine Science and Technology, Qingdao, China, [4]University of Chinese Academy of Sciences, Beijing, China, [5]Department of Geophysics, Graduate School of Science, Tohoku University, Sendai, Japan

Abstract Arrival times of seismic phases contribute substantially to the study of the inner working of the Earth. Despite great advances in seismic data collection, the usage of seismic arrival times is still insufficient because of the overload manual picking tasks for human experts. In this work we employ a deep-learning method (PickNet) to automatically pick much more *P* and *S* wave arrival times of local earthquakes with a picking accuracy close to that by human experts, which can be used directly to determine seismic tomography. A large number of high-quality seismic arrival times obtained with the deep-learning model may contribute greatly to improve our understanding of the Earth's interior structure.

Plain Language Summary Deep learning is currently attracting immense research interest in seismology due to its powerful ability to deal with huge seismic data collections. In this study we developed a deep-learning model (PickNet) that can rapidly pick a great number of first *P* and *S* wave arrival times precisely from local earthquake seismograms. The picking accuracy of the arrival times provided by our PickNet model is close to that by human experts. The data are good enough to be used directly to determine high-resolution 3-D velocity models of the Earth. Our PickNet model can deal with seismic waveforms provided by data centers of different earthquake networks. Furthermore, our PickNet model is also a potential tool for automatically picking later seismic phases accurately. A large number of high-quality seismic arrival times can be used to illuminate the Earth structure clearly. Hence, this study may greatly contribute to improve our knowledge of the Earth's interior.

1. Introduction

Arrival times of seismic phases provide key information on the Earth's structure, making them essential for studying earthquakes (Hardebeck, 2002; Yang et al., 2012) and illuminating the Earth's interior (Nelson & Grand, 2018; Tkalčić & Phm, 2018; Wang & Zhao, 2012). The significant increase in seismic data collections is creating an impossible task for human experts to manually pick all the arrival times in these enormous data volumes. To collect the observed seismic data efficiently, many researchers have proposed a variety of automatic arrival time picking algorithms, such as short-term average/long-term average algorithm (Allen, 1978; Stevenson, 1976), auto regression with Akaike Information Criterion (Sleeman & van Eck, 1999), higher-order statistics (Ross et al., 2016; Saragiotis et al., 2002), and so on. Although these classical automatic picking algorithms are helpful in seismological studies, their performances are less precise than manual picks, making them inefficient for seismic imaging (Akram & Eaton, 2016; Sharma et al., 2010). Besides, a threshold is usually required in these algorithms, making them impractical to handle the complex seismic data. Different from these traditional methods that use only a small number of manually designed features for seismograms, machine learning methods especially deep neural networks can extract abundant features automatically from numerous labeled seismic data. Previous studies have tried to solve the picking problem with shallow neural networks (Dai & MacBeth, 1997; Gentili & Michelini, 2006), but their performances were greatly limited because of their low computational speed and simple network complexity at that time. With the fast development of the graphic processing units and computational technology, recent studies have achieved remarkable picking performance on the tested seismic data via deep neural networks (Ross, Meier, & Hauksson, 2018; Ross, Meier, Hauksson, & Heaton, 2018; Zhu & Beroza, 2019). However, due to the small epicentral distances in the training seismic data set (<200 km), the usefulness of these deep-learning models is greatly limited. As a result, these deep-learning models can only be applied to

precisely pick arrival times of local earthquakes with short epicentral distances to illuminate the very shallow structure of the Earth.

The arrival-time picking studies have many common features with the edge detection problem, because the arrival time point can be regarded as the seismic phase edge in one-dimensional space, which is different from directly predicting arrival times by regression (Ross, Meier, & Hauksson, 2018) and similar to treating picking as a segmentation problem by predicting the probability distribution of arrival times (Ross, Meier, Hauksson, & Heaton, 2018; Zhu & Beroza, 2019). The major difference between treating seismic phase picking as edge detection and segmentation is that the ground truths in the former are much narrower and sharper than those in the latter. Because edge detection has been well studied in computer vision via deep-learning models (Ke et al., 2017; C. Liu et al., 2017; Xie & Tu, 2017), we adopt the ideas from these previous studies and present an end-to-end deep-learning approach (PickNet) to pick first *P* and *S* wave arrival times of local earthquakes with epicentral distances up to ~1,000 km by mapping a normalized seismic record to an impulse-like time series, where the position of the impulse peak indicates the picking time. Our PickNet is designed based on the Rich Side-output Residual Network (C. Liu et al., 2017) that has rich side outputs of multiscale and multilevel information. Every output of a branch in our model is supervised, and thus, the picking task is decomposed to multiple different levels for convolutional filters of different depths to learn, making the detection more accurate. This advantage is validated by many tasks, such as edge detection (Xie & Tu, 2017), object detection (Lin et al., 2017), skeleton detection (Ke et al., 2017), and text detection (X. Liu et al., 2018). Combining these side outputs can make the model robust on signals that possess various scales and aspect ratios (Y. Liu et al., 2018) like seismic waveforms. The residual units in PickNet can speed up the training convergence and refine the detection result hierarchically by reducing the residual between the rich side outputs and the ground truth. Moreover, PickNet is developed under the linear span framework with a solid mathematical background (C. Liu et al., 2018). By increasing independence of spanning sets and the enlarged spanned output space, it can gain good performance on the seismic phase picking problem.

We train the PickNet model with ~460,000 first *P* wave and ~280,000 first *S* wave arrival times that have been manually picked with high accuracy from three-component seismograms recorded at 782 stations of the dense High-sensitivity Seismic Network (Hi-net) deployed on the Japan Islands. We test our model using the Hi-net seismograms of 300 local earthquakes and obtain ~10 times more *P* and *S* wave arrival times than those released by the Japan Meteorological Agency (JMA). These picks are then tested by using them to perform a tomographic inversion, resulting in three-dimensional (3-D) *P* and *S* wave velocity (Vp, Vs) images that exhibit nearly the same patterns as those revealed by previous tomographic studies of the Japan subduction zone (Liu & Zhao, 2016; Wang & Zhao, 2012; Zhao et al., 2012). Note that the previous studies used much more local earthquakes as well as teleseismic events. We also verify the universality of our PickNet model using data from the China Earthquake Administration (Zheng et al., 2010), the International Seismological Centre, and the Southern California Earthquake Data Center (Ross, Meier, & Hauksson, 2018). Furthermore, our PickNet model can potentially be extended for picking other seismic phases by simply changing the training data set, because little information on a specific phase is required. This data-driven method can rapidly acquire arrival time picks that approach the accuracy of manual picking by human experts. Thus, a huge number of seismic data can be used to greatly improve our understanding of the Earth's interior structure and dynamics.

2. Methods and Data

Deep convolutional neural networks (DCNNs), which are inspired by the connectivity patterns of human neurons in the visual cortex, use no preprocessing because the network can learn the convolutional kernels by itself (Agarwal et al., 2018). Therefore, the DCNNs are extremely powerful in performing vision-based tasks (Ke et al., 2017; C. Liu et al., 2017), making them well suited for picking seismic arrival times. Here we define the seismic arrival time picking problem as $f[n] = T(g[n])$, where f is the output signal that contains information on the seismic arrival time, g is the input observation of the corresponding normalized seismic waveform, n denotes the discrete time series, and T is an ideal operator. f is a unit impulse that can be expressed as $f[n] = \delta[n - N]$, where N denotes the arrival time picking point. The approximation of T is obtained using DCNNs that are powerful algorithms for approximating nonlinear functions (Adler & Öktem, 2017).

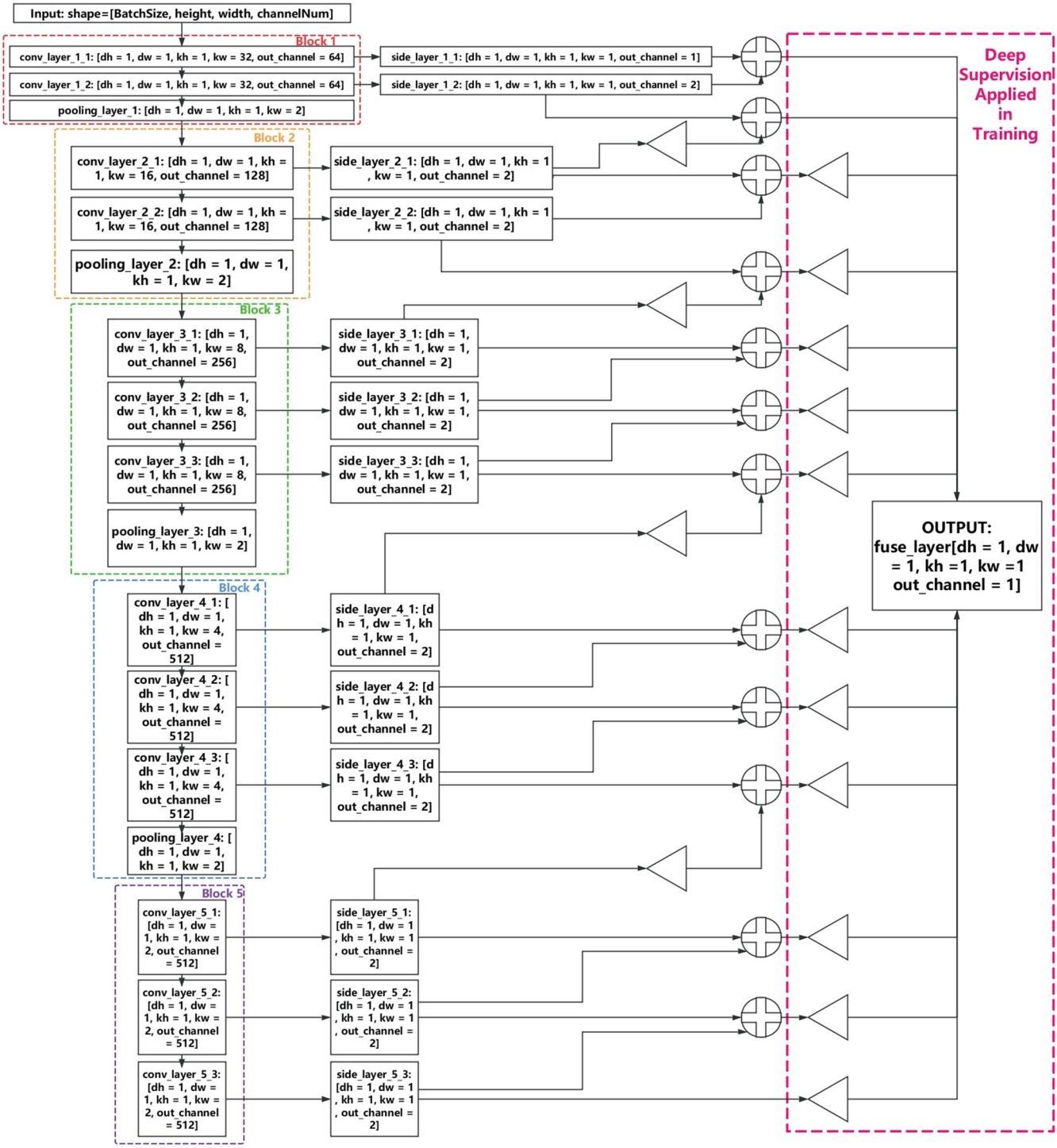

Figure 1. The structure of PickNet model in this study. "dh" and "dw" stand for stride on height and width, respectively. "kh" and "kw" stand for kernel height and width, respectively. "out_channel" stands for the number of output filters in the convolution. Triangles denote the up-samplings by deconvolution. Circles with an inside plus denote the residual units. Deep supervision is applied in training, which means that all tensors at the end of every branch are also supervised with the ground truth using the same loss function as final output before being fused.

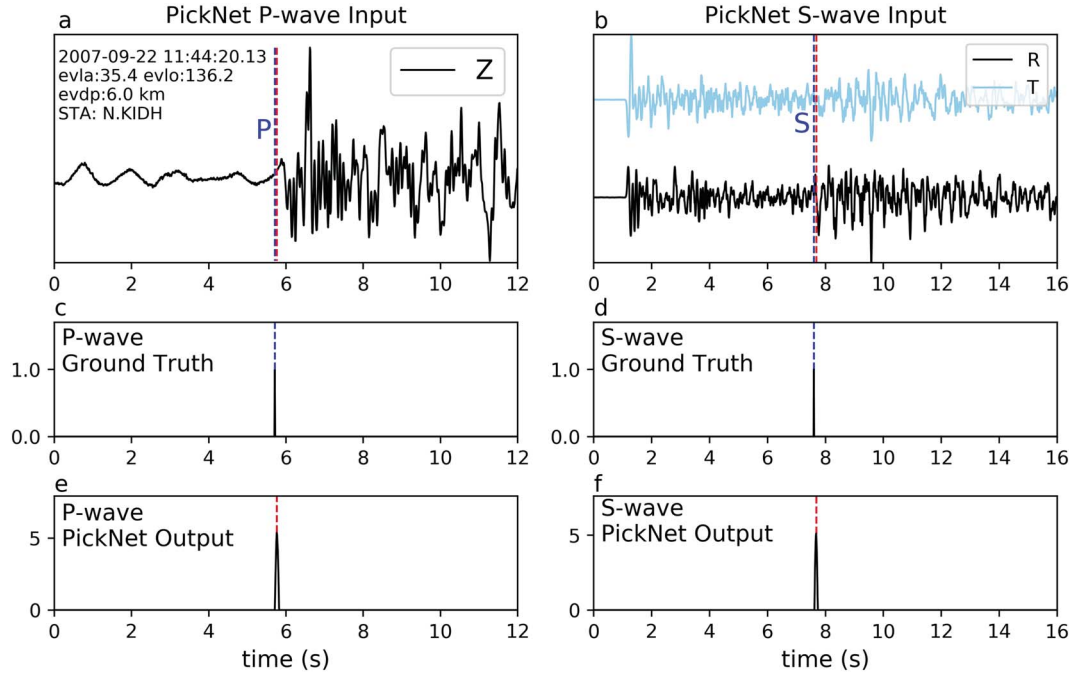

Figure 2. Input and output examples of PickNet for picking first *P* and *S* wave arrival times. The normalized waveforms (a and b) of an earthquake recorded at station N.KIDH in Japan are centered in theoretical arrival times as inputs for the PickNet model. Blue dashed lines indicate the picks, given by human experts, as ground truths, and they are expressed as unit impulse (c and d). Red dashed lines indicate the picks by the PickNet model and their positions corresponding to the maximum values indicate the outputs (e and f). Y axes in (e) and (f) represent PickNet output values calculated with the rectified linear unit (ReLU) activation function at the testing stage.

The deep-learning model PickNet we employ to estimate *T* consists of fully convolutional networks constructed from a modified VGG-16 model (Simonyan & Zisserman, 2014) with Rich Side-output Residual Networks (C. Liu et al., 2017; see Figure 1). A side-output layer is added after every layer in the VGG-16 model to extract rich convolutional features at different levels. These rich side outputs are combined with residual units, and then these outputs are supervised with same ground truth as final output in order to learn the picking task hierarchically. Finally, a multiscale combination strategy with convolutional layers is used to generate the final output. More details on the PickNet structure can be found in Supporting Information S1.

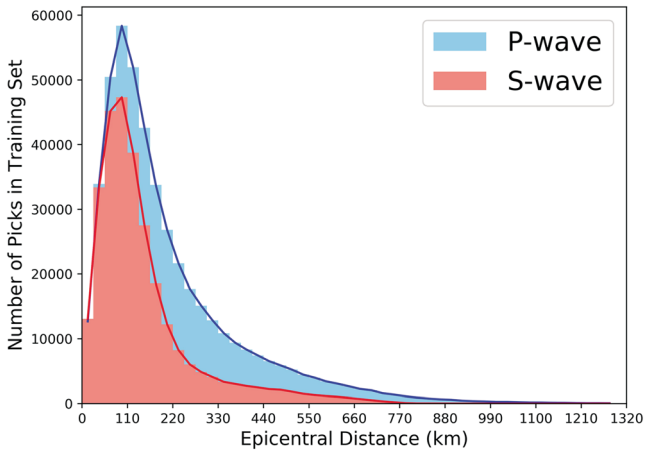

Figure 3. The number of picks versus epicentral distance in the training data set. The training data set for both the first *P* and *S* wave arrival times is imbalanced and decreases rapidly with the epicentral distances exceeding ~110 km.

The model inputs consist of 12-s-long slices from the vertical (Z) component and 16-s-long slices from the radial (R) and transverse (T) components of the seismograms for the first *P* and *S* wave arrival time picking, respectively (Figure 2). Because *S* wave phase is more complicated than *P* wave phase, longer-length slice containing more waveform information is used to determine *S* wave arrival. Two different models are trained separately. One model uses only the vertical component to estimate the *P* wave phase, the other model uses the R and T components to estimate the *S* wave phase. Each slice is centered on its theoretical arrival time, which is calculated using the TauP code (Crotwell et al., 1999) and the AK135 velocity model (Kennett et al., 1995). Both inputs of *P* and *S* waves slices are processed using zero-mean and maximum-absolute-value normalizations. We construct the training stage outputs by setting *N* to the positions of manually picked arrival times and obtain the approximation of *T* using a large number of input-output (*g*-*f*) pairs. The testing stage outputs are calculated using the previously trained model and the corresponding test seismograms. The outputs are expected to be zero at every time point

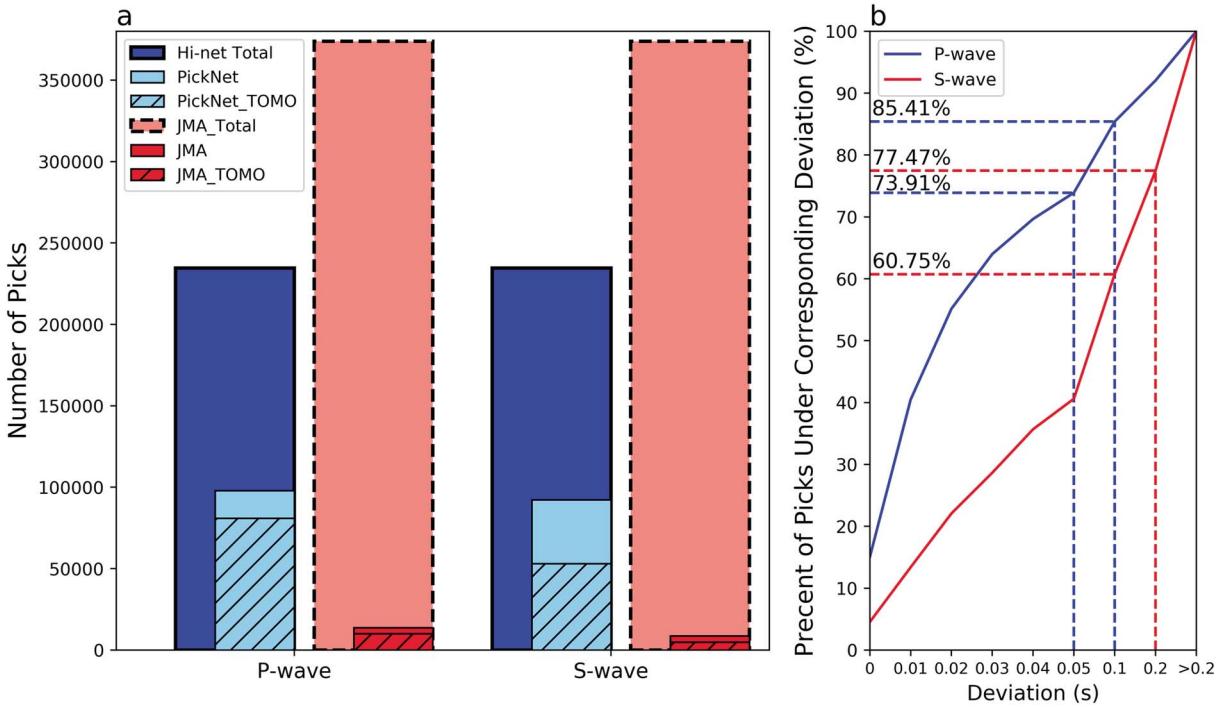

Figure 4. Comparisons of the picked arrival times provided by PickNet and Japan Meteorological Agency (JMA) in testing data set. (a) The blue columns indicate the total number of seismograms recorded at 782 Hi-net stations. The total number of seismograms is the number of earthquakes multiplying the number of seismic stations, because each earthquake could be recorded at all stations. All data in the blue columns are used as inputs, and the sky blue columns denote the picked arrival times by our PickNet model. Light-coral columns indicate the ideal total number of seismograms recorded at 1,246 JMA stations, and red bars indicate the picked arrival times provided by JMA. Slashes indicate the numbers of arrival times, which are used to determine seismic tomography. (b) The red and blue lines show the deviation distributions of the first *P* and *S* wave arrival times between the overlapped picks provided by PickNet and JMA, respectively.

except for the positions of the proposed first arrival times. The outputs exhibit sharp spike-shape patterns, with the time of the maximum value representing the picked arrival time (Figure 2), because the trained model is an estimation of *T*. As shown in Figure S1, if a given seismic arrival is either missing or difficult to distinguish in its corresponding input, the resultant output yields all zeros, such that no pick is made for that seismogram.

Our training data set contains ~460,000 first *P* wave and ~280,000 first *S* wave high-quality arrival times and the corresponding seismograms of 6,272 local earthquakes recorded at 782 Hi-net stations. These local events occurred from 2005 to 2007, and the first *P* and *S* wave arrival times were manually picked from the original Hi-net seismograms with high accuracy by the staff of Research Center for Prediction of Earthquakes and Volcanic Eruptions, Tohoku University. Although the size of our training data set for the first *P* wave arrival times in this study is smaller than that used by Ross, Meier, and Hauksson (2018), the epicentral distances of the arrival times in our data set are up to ~1,000 km (see Figure 3), whereas the maximum distance is less than 200 km in Ross, Meier, and Hauksson (2018). We oversample the inputs with epicentral distances larger than 200 km by dividing the train set into three different distance bins (0–220, 220–550, and >550 km) and randomly choosing same amount of instances from each bin for every training step to handle the negative effect induced by the imbalance in the epicentral distance distribution in the training data set. Therefore, our PickNet model can make accurate picks for inputs with epicentral distances of up to ~1,000 km in the testing stage. This is important and useful, because the deeper Earth structure can only be illuminated by the seismic rays with larger epicentral distances. Details of the training batch size, loss function, and so on can be found in Supporting Information S1.

3. Result and Discussion

We use ~234,600 seismograms of 300 local earthquakes recorded at 782 Hi-net stations as the testing data set. These events occurred from January to June 2016 in Japan, which are randomly selected and consist of 50

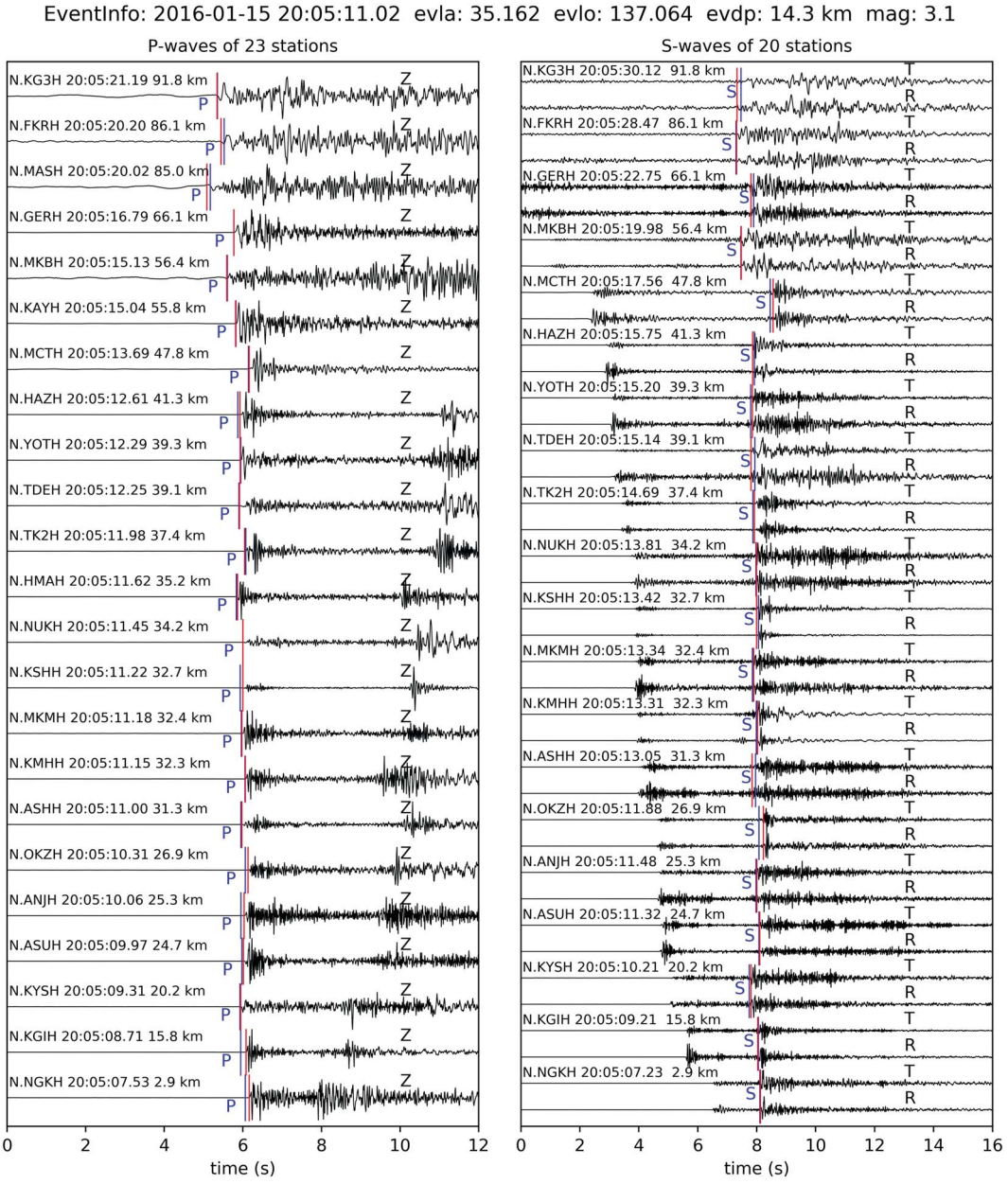

Figure 5. Comparison of picked arrival times of an earthquake recorded at Hi-net stations provided by PickNet and JMA. The station code, trace start time, and epicentral distance are shown above each seismogram. Red and blue lines indicate the first arrival times provided by PickNet and JMA, respectively. The blue line is masked by the red line when they are at the same position. Additional arrival times of this earthquake provided by PickNet are shown in Figure S9.

events with M 1.0–2.0, 100 events with M 2.0–3.0, 100 events with M 3.0–4.0, and 50 events with M >4.0. It took only ~5 min for our PickNet model to analyze these seismograms using a GEFORCE GTX1080Ti graphic processing unit.

We obtained 97,998 *P* wave and 92,229 *S* wave arrival times of the 300 events recorded at 782 Hi-net stations (see Data Sets S1 and S2). In contrast, the JMA only provided 13,765 *P* wave and 8,643 *S* wave arrival times for the same earthquakes recorded at 1,246 seismic stations (including 782 Hi-net stations and 464 stations of other local networks; Figure 4a), because JMA only picked the arrival times at stations close to the epicenter so as to locate each event quickly. Among the arrival times provided by JMA, 6,725 *P* wave and 4,656 *S* wave arrival times were recorded at the Hi-net stations. As a result, there are 6,524 *P* wave and 4,425 *S* wave arrival

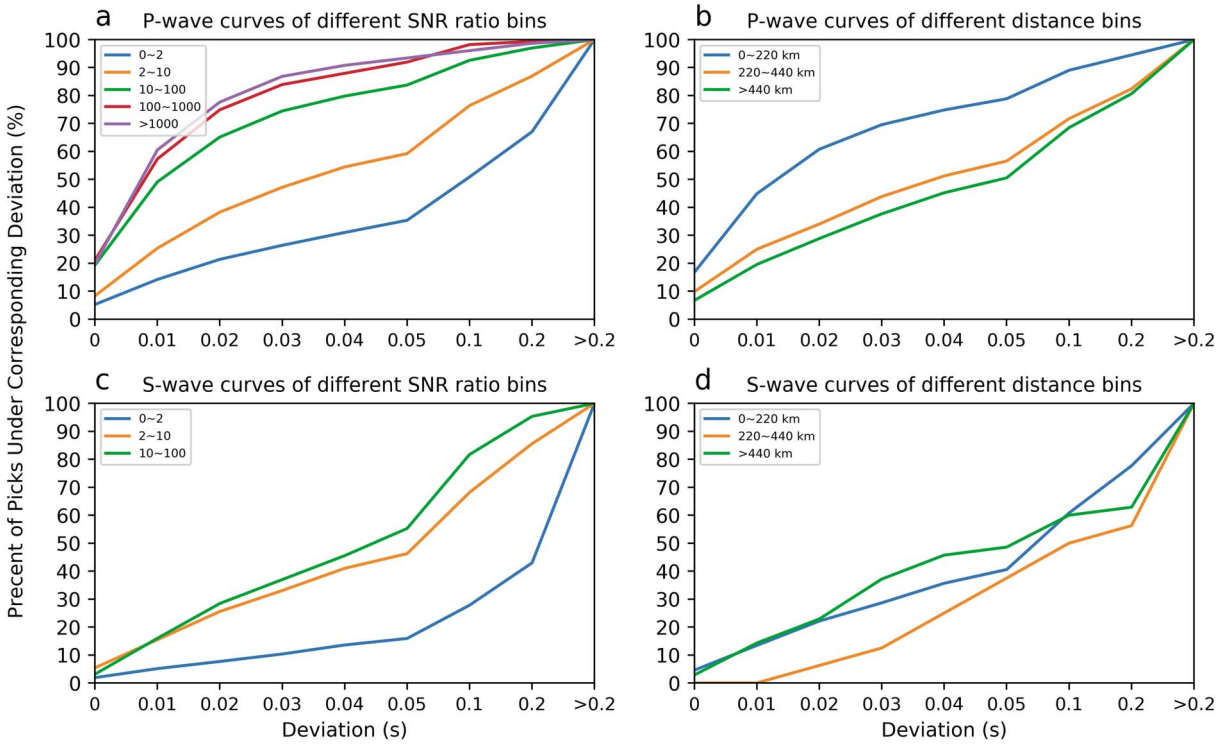

Figure 6. The PickNet performances for different SNR bins (a, c) and epicentral distance bins (b, d). SNR = signal-to-noise ratio.

times, with both the PickNet and JMA picks for the same earthquake-station pairs. Therefore, the false negative rates are nearly 2.99% and 4.96% for *P* and *S* wave arrival times, respectively. We use "deviation," which is the absolute value of the picks by PickNet minus the corresponding one picked by human experts, to evaluate the accuracy of PickNet. These overlapping arrival times are analyzed to demonstrate the reliability of the PickNet results. Figure 4b shows that 73.91% (85.41%) of our first *P* wave picks have deviations of less than 0.05 s (0.10 s), and 60.75% (77.47%) of our first *S* wave picks have deviations of less than 0.10 s (0.20 s) from the JMA picks. Figure 5 shows comparisons of the picked arrival times for a local earthquake recorded at Hi-net stations provided by PickNet and JMA. Figure S2 shows additional PickNet picks for the same local earthquake, and the Data Sets S1 and S2 show all the picked arrival times of the 300 earthquakes by the PickNet model. The comparisons indicate that our PickNet model exhibits a similar performance to human experts in distinguishing the first *P* and *S* wave arrival times, with the automatic picks well within the range of picking errors. The larger deviations of the *S* wave picks than those of the *P* wave picks (Figure 4b) are likely due to the more complex patterns of the *S* wave seismograms. That is why all the seismic network data centers in the world always pick a smaller number of *S* wave arrival times with a lower picking accuracy than those for *P* wave arrival times. In addition, the number of *S* wave arrival times is approximately half that of the *P* wave arrival times in our training data set, which also reduces the PickNet picking performance for *S* wave arrival times. We have also analyzed the PickNet performance at different signal-to-noise ratio (SNR) bins and epicentral distance bins (Figure 6). The same as Ross, Meier, and Hauksson (2018), we define SNR as the ratio between the peak absolute amplitudes in the range of 0.5 s after and 0.5 s before the target pick. The PickNet performance decreases when SNR is less than 2 (Figures 6a and 6c), suggesting that a postprocessing to discard those picks with a low SNR should be done to improve the reliability of results. The PickNet performance for the long-distance picks is not as good as that for the short-distance ones (Figures 6b and 6d) because of the lack of a long-distance training data set even with oversampling applied.

As a test of the picked data, we used the *P* and *S* wave arrival times of 299 earthquakes picked by the PickNet and JMA to conduct tomographic inversions. Note that one event occurred very far from the seismic network on the Japan Islands, and so it is not used in the inversions. The computer program Tomog3d (Zhao et al.,

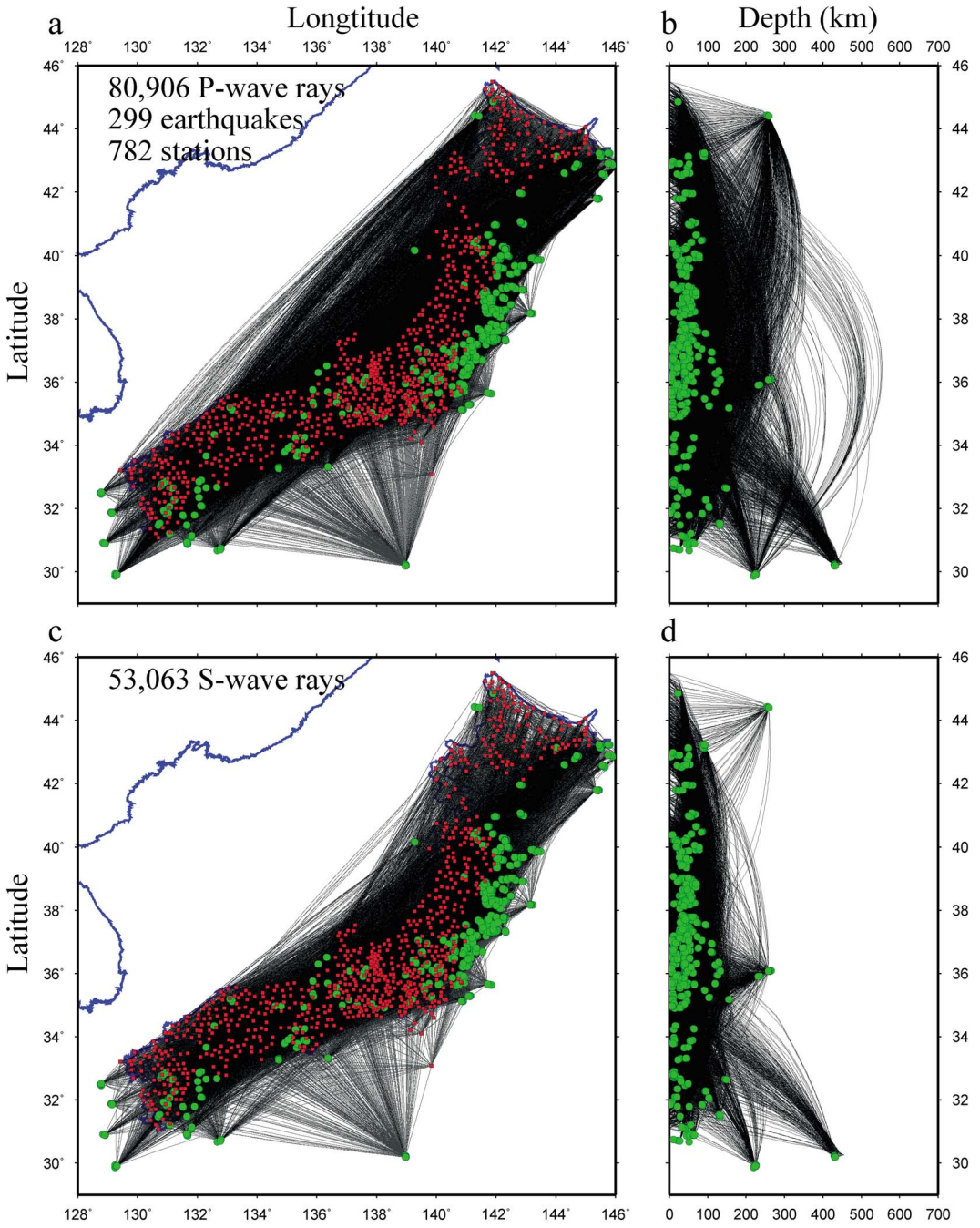

Figure 7. Distribution of (a, b) *P* and (c, d) *S* wave ray paths used to conduct the tomographic inversion. The *P* and *S* wave arrival times are obtained by our PickNet model. The numbers of the earthquakes (green dots), seismic stations (red squares), and ray paths (black lines) used are shown at the upper-left corner of (a) and (c). The earthquakes are relocated during the inversion process.

1992, 2012) is applied to simultaneously invert for both 3-D Vp and Vs variations of the Japan subduction zone. The initial velocity model is constructed from a slightly modified J-B velocity model (Jeffreys & Bullen, 1940) with lateral depth variations of the Conrad and Moho discontinuities and the upper boundary of the subducting Pacific slab (Zhao et al., 1997, 2012). The Pacific slab is assumed to be 85 km thick and with 4% faster Vp and 6% faster Vs than those of the surrounding mantle in the starting velocity model (Zhao et al., 1992). To express the 3-D velocity structure, a 3-D grid is set up in the study volume with a horizontal grid interval of 0.5° and with grid meshes at depths of 0, 20, 40, 60, 90, 120, 150, 200, 300, and 400 km. In the

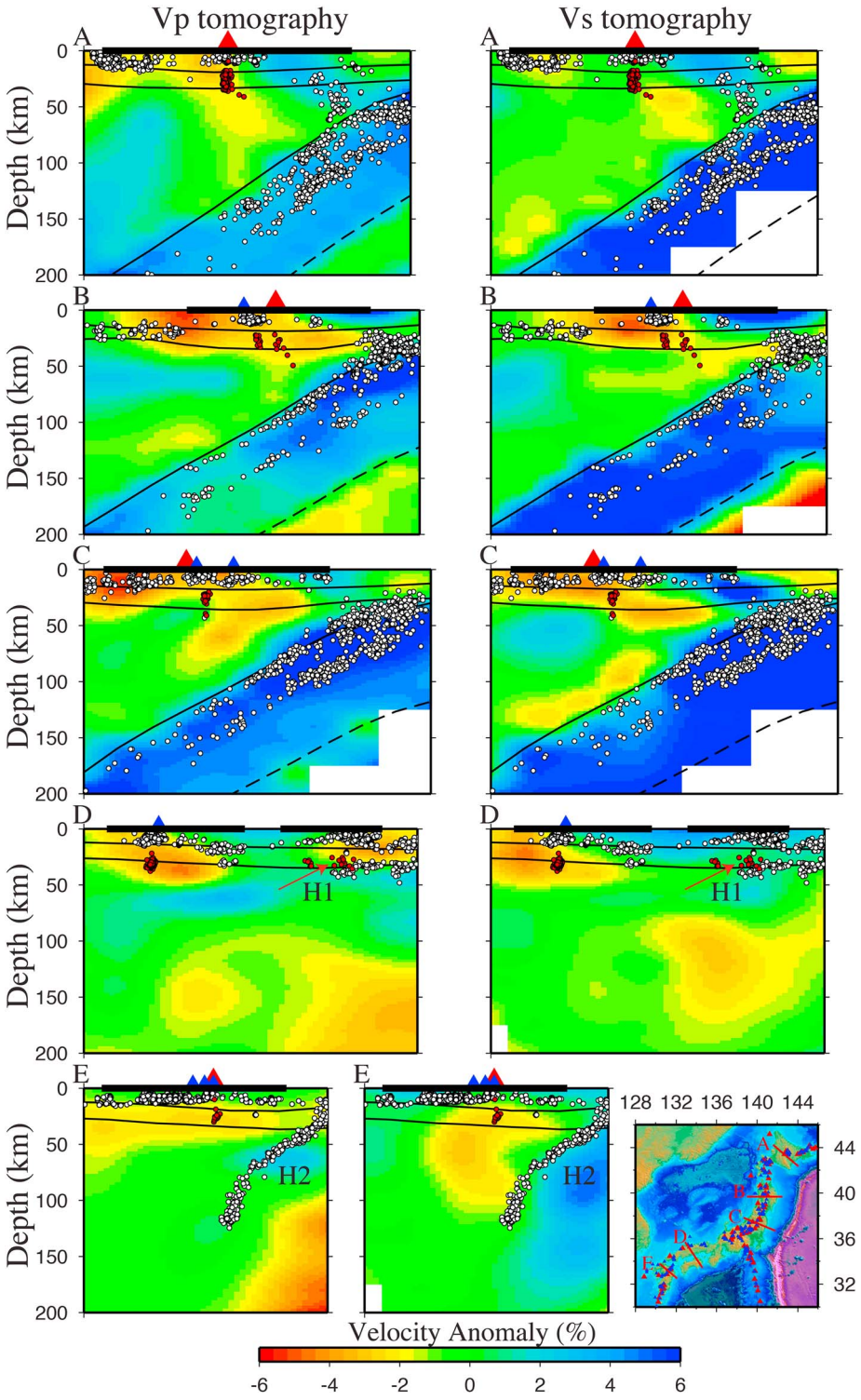

Figure 8. East-west vertical cross sections of *P* and *S* wave tomographic images. The arrival times used to determine the tomography are obtained by our PickNet model. Locations of the cross sections are shown in the inset map. The red and blue colors denote low and high velocity anomalies, respectively, whose scale is shown at the bottom. The curved lines denote the Conrad and Moho discontinuities and the upper boundary of the subducting Pacific slab. The dashed lines denote the estimated lower boundary of the subducting Pacific slab whose thickness is assumed to be 85 km in the inversion. The red and blue triangles denote the active and quaternary volcanoes, respectively. The white and red dots denote normal earthquakes and low-frequency microearthquakes (M < 2.5), respectively, which occurred during 2002 to 2007 within a 5-km width of each cross section.

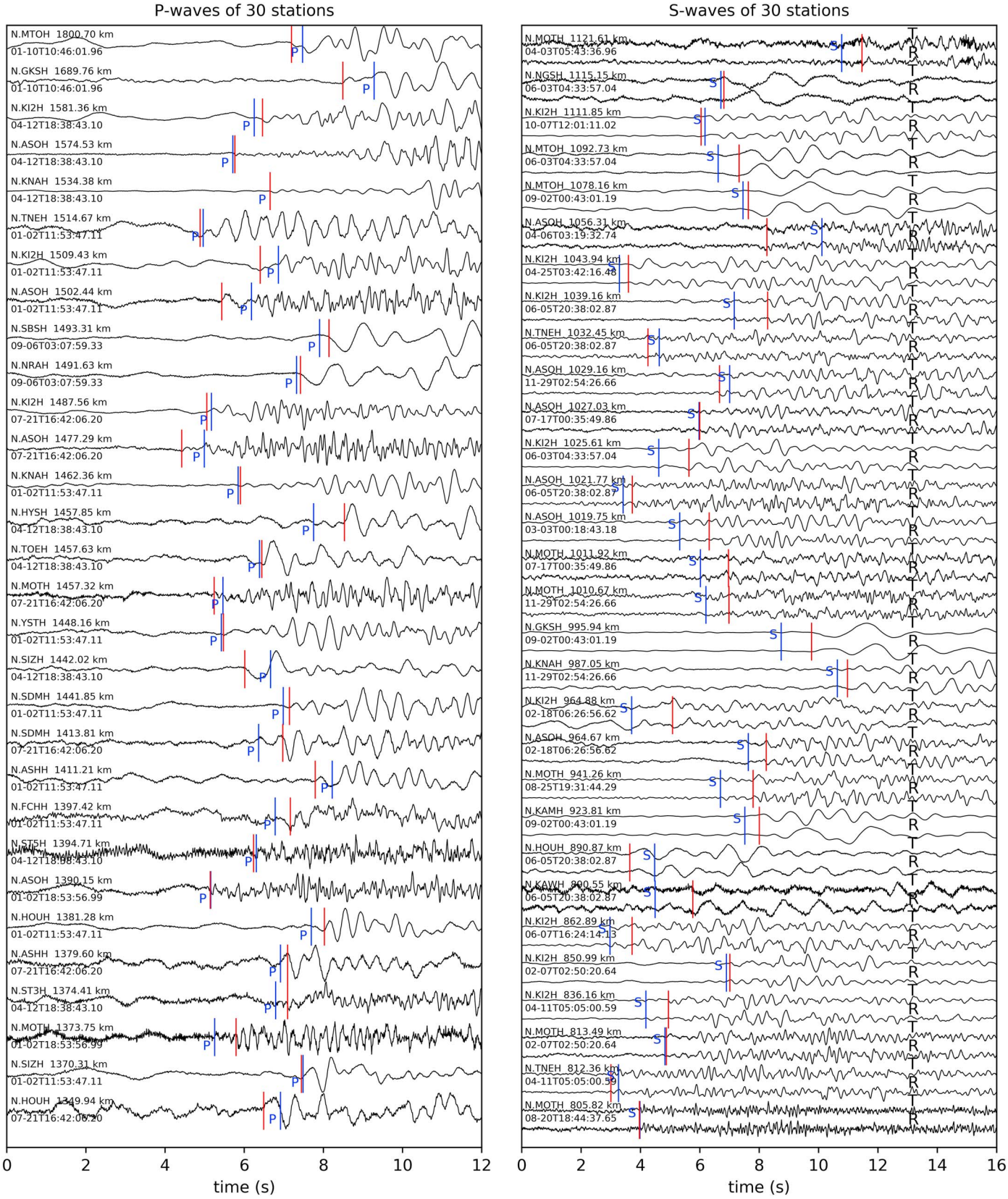

Figure 9. Comparison of the performances between PickNet and JMA on picking long-distance arrival times of regional earthquakes in 2018. The labeling is the same as that of Figure 5.

tomographic inversions, the *P* and *S* wave arrival times with travel-time residuals (i.e., difference between the observed and theoretical travel times) greater than 1.8 s are excluded. The *P* and *S* wave ray coverages obtained by PickNet (Figure 7) are much denser than those of the JMA picks (Figure S3). The tomographic results from the PickNet data (Figure 8) and the JMA data (Figure S4) are obtained after four iterations of the inversion. The local earthquakes are relocated (Engdahl & Lee, 1976) using the *P* and *S* wave arrival times and the resultant 3-D velocity models.

The tomographic images (Figure 8) obtained with the PickNet data reveal similar velocity features to those of the previous tomographic models obtained with much more data of local earthquakes as well as teleseismic events (Liu & Zhao, 2016; Wang & Zhao, 2012; Zhao et al., 2012). Low-velocity anomalies are revealed in the mantle wedge beneath active arc volcanoes and back-arc areas, and low-frequency microearthquakes ($M < 2.5$) occur in the lower crust and uppermost mantle beneath the active arc volcanoes, which are closely associated with the magmatic and volcanic activity (Hasegawa & Yamamoto, 1994; Niu et al., 2018). Some nonvolcanic low-frequency events (denoted by the red arrow in section D of Figure 8) are generated by slips on the gently subducting Philippine Sea slab interface (Shelly et al., 2007), and the subducting Philippine Sea slab is imaged as high-velocity anomalies with considerable variations in dip angle and subduction depth (H1 in section D and H2 in section E of Figure 8). These velocity features are also revealed by a tomographic inversion with a starting model without the Pacific slab (Figure S5), though the subducting Pacific slab is not well recovered because only 299 earthquakes selected randomly are used in the tomographic inversion. However, these velocity features are poorly imaged or absent in the tomographic images obtained with the JMA data of the 299 events (Figures S4 and S6), because the JMA data set contains a smaller number of *P* and *S* wave arrival times than that of our PickNet data set. The direct application to seismic tomography indicates that false positive rates of *P* and *S* wave arrival times provided by PickNet are relatively low and confirms the reliability and practicality of our PickNet method.

We have also investigated the PickNet performance on long-distance seismograms. We collected all the *P* and *S* wave arrival times recorded at the Hi-net stations with epicentral distances greater than 440 km provided by JMA during the year 2018, including 554 *P* wave and 129 *S* wave arrival times. We compare the JMA data with the PickNet picks (see Figures 9 and S7 and Data Set S3), and the result shows that the PickNet performs well on the long-distance seismograms, though the picking accuracy is lower than that on the short-distance seismograms. The SNR of seismogram decreases as distance increases, which causes the lower picking accuracy on the long-distance seismograms by both human experts and the PickNet model. The PickNet performance on long-distance seismograms would be improved by adding more high-quality long-distance picks into the training data set.

We have tested our PickNet model on local earthquakes recorded at stations in China and United States and obtained ~6 times more *P* and *S* wave arrival times with a similar quality to those provided by China Earthquake Administration (Zheng et al., 2010; Figures S8 and S9) and International Seismological Centre (Figures S10 and S11). We have also trained and tested our deep-learning method using a large number of *P* wave arrival times provided by Southern California Earthquake Data Center (Ross, Meier, & Hauksson, 2018) and achieved good performance with 78% of the automatic picks having deviations of less than 0.03 s from the manual picks (Figure S12). These tests verify the generalization of our PickNet model, although our model is only trained with a seismic data set recorded at the Hi-net stations. Hence, the PickNet model can be widely used to pick high-quality *P* and *S* wave arrival times from seismograms recorded by different seismic networks.

4. Conclusions

In this work we have developed the PickNet model that can be used to pick high-quality *P* and *S* wave arrival times of local earthquakes. The PickNet model can be further improved to automatically pick other seismic phases (e.g., PmP, SmS, and PKiKP) accurately when a large high-quality training data set is provided because little information on a given seismic phase is required. Our study highlights how a deep-learning approach can efficiently mine arrival time information from vast seismic datasets and provide a powerful tool for enhancing our knowledge of the Earth's interior structure.

Acknowledgments

High-quality arrival-time data sets for this study were provided by the data centers of Tohoku University (http://www.aob.tohoku.ac.jp), the JMA Unified Earthquake Catalogue (https://hinetwww11.bosai.go.jp/auth/JMA/), the China Earthquake Administration (https://www.cea.gov.cn/), and the International Seismological Centre (http://www.isc.ac.uk). Waveform data for this study were provided by the Hi-net (http://www.hinet.bosai.go.jp), IRIS (https://www.iris.edu), the International Federation of Digital Seismograph Networks (https://www.fdsn.org), the Southern California Earthquake Data Center (http://scedc.caltech.edu/), and Data Management Centre of the China National Seismic Network at Institute of Geophysics, China Earthquake Administration (http://www.seisdmc.ac.cn). The TensorFlow libraries (Abadi et al., 2016; https://www.tensorflow.org) were used to train and build our deep-learning model. The Obspy (Beyreuther et al., 2010; Krischer et al., 2015; Megies et al., 2011) was used for reading, writing, and slicing the seismograms. Figures were generated using the free and open software GMT 4.5.3 (Wessel & Smith, 1998) and Matplotlib (Hunter, 2007). This study was financially supported by the National Key R&D Program of China (grant 2017YFC0601206), the National Natural Science Foundation of China (grants 41474043 and 41274089), and the Youth Innovation Promotion Association of CAS (2014058). Yehuda Ben-Zion (editor), Zachary Ross, and Qingkai Kong provided constructive review comments and suggestions that have improved the manuscript. The PickNet codes and the trained models associated with this paper will be available online (https://www.researchgate.net/profile/Jian_Wang78/publications).

References

Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Devin, M., et al. (2016). TensorFlow: A system for large-scale machine learning. Proceedings of the 12th USENIX Symposium on Operating Systems Design and Implementation, 265–283. https://www.usenix.org/system/files/conference/osdi16/osdi16-abadi.pdf

Adler, J., & Öktem, O. (2017). Solving ill-posed inverse problems using iterative deep neural networks. *Inverse Problems*, *33*(12), 124007. https://doi.org/10.1088/1361-6420/aa9581

Agarwal, S., Terrail, J. O. D., & Jurie, F. (2018). Recent advances in object detection in the age of deep convolutional neural networks. *ArXiv:1809.03193 [Cs]*. http://arxiv.org/abs/1809.03193

Akram, J., & Eaton, D. W. (2016). A review and appraisal of arrival-time picking methods for downhole microseismic data. *Geophysics*, *81*(2), KS71–KS91. https://doi.org/10.1190/geo2014-0500.1

Allen, R. V. (1978). Automatic earthquake recognition and timing from single traces. *Bulletin of the Seismological Society of America*, *68*(5), 1521–1532.

Beyreuther, M., Barsch, R., Krischer, L., Megies, T., Behr, Y., & Wassermann, J. (2010). ObsPy: A Python toolbox for seismology. *Seismological Research Letters*, *81*(3), 530–533. https://doi.org/10.1785/gssrl.81.3.530

Crotwell, H. P., Owens, T. J., & Ritsema, J. (1999). The TauP toolkit: Flexible seismic travel-time and ray-path utilities. *Seismological Research Letters*, *70*(2), 154–160. https://doi.org/10.1785/gssrl.70.2.154

Dai, H., & MacBeth, C. (1997). The application of back-propagation neural network to automatic picking seismic arrivals from single-component recordings. *Journal of Geophysical Research*, *102*(B7), 15,105–15,113. https://doi.org/10.1029/97JB00625

Engdahl, E. R., & Lee, W. H. K. (1976). Relocation of local earthquakes by seismic ray tracing. *Journal of Geophysical Research*, *81*(23), 4400–4406. https://doi.org/10.1029/JB081i023p04400

Gentili, S., & Michelini, A. (2006). Automatic picking of P and S phases using a neural tree. *Journal of Seismology*, *10*(1), 39–63. https://doi.org/10.1007/s10950-006-2296-6

Hardebeck, J. L. (2002). A new method for determining first-motion focal mechanisms. *Bulletin of the Seismological Society of America*, *92*(6), 2264–2276. https://doi.org/10.1785/0120010200

Hasegawa, A., & Yamamoto, A. (1994). Deep, low-frequency microearthquakes in or around seismic low-velocity zones beneath active volcanoes in Northeastern Japan. *Tectonophysics*, *233*(3-4), 233–252. https://doi.org/10.1016/0040-1951(94)90243-7

Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. *Computing in Science & Engineering*, *9*(3), 90–95. https://doi.org/10.1109/MCSE.2007.55

Jeffreys, H., & Bullen, K. (1940). Seismological tables. *Abstract Presented at the Annual Meeting of the Seismological Society of Japan (2000), British Association for the Advancement of Science*, London, pp. 50.

Ke, W., Chen, J., Jiao, J., Zhao, G., & Ye, Q. (2017). SRN: Side-output residual network for object symmetry detection in the wild. *2017 IEEE Conference on Computer Vision and Pattern Recognition,* 302–310. https://doi.org/10.1109/CVPR.2017.40

Kennett, B. L. N., Engdahl, E. R., & Buland, R. (1995). Constraints on seismic velocities in the Earth from traveltimes. *Geophysical Journal International*, *122*(1), 108–124. https://doi.org/10.1111/j.1365-246X.1995.tb03540.x

Krischer, L., Megies, T., Barsch, R., Beyreuther, M., Lecocq, T., Caudron, C., & Wassermann, J. (2015). ObsPy: A bridge for seismology into the scientific Python ecosystem. *Computational Science & Discovery*, *8*(1), 014003. https://doi.org/10.1088/1749-4699/8/1/014003

Lin, P. Dollár, R. Girshick, K. He, B. Hariharan, & S. Belongie. (2017). Feature pyramid networks for object detection. *2017 IEEE Conference on Computer Vision and Pattern Recognition*, 936–944. https://doi.org/10.1109/CVPR.2017.106

Liu, C., Ke, W., Jiao, J., & Ye, Q. (2017). RSRN: Rich side-output residual network for medial axis detection. *2017 IEEE International Conference on Computer Vision Workshops*, 1739–1743. https://doi.org/10.1109/ICCVW.2017.204

Liu, C., Ke, W., Qin, F., & Ye, Q. (2018). Linear span network for object skeleton detection. *The European Conference on Computer Vision 2018*, 136–151. https://doi.org/10.1007/978-3-030-01216-8_9

Liu, X., D. Liang, S. Yan, D. Chen, Y. Qiao, & J. Yan. (2018). FOTS: Fast oriented text spotting with a unified network. *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 5676–5685. https://doi.org/10.1109/CVPR.2018.00595

Liu, X., & Zhao, D. (2016). P and S wave tomography of Japan subduction zone from joint inversions of local and teleseismic travel times and surface-wave data. *Physics of the Earth and Planetary Interiors*, *252*, 1–22. https://doi.org/10.1016/j.pepi.2016.01.002

Liu, Y., M. Cheng, X. Hu, J. Bian, L. Zhang, X. Bai, & J. Tang. (2018). Richer convolutional features for edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1–1. https://doi.org/10.1109/TPAMI.2018.2878849

Megies, T., Beyreuther, M., Barsch, R., Krischer, L., & Wassermann, J. (2011). ObsPy—What can it do for data centers and observatories? *Annals of Geophysics*, *54*, 47–58. https://doi.org/10.4401/ag-4838

Nelson, P. L., & Grand, S. P. (2018). Lower-mantle plume beneath the Yellowstone hotspot revealed by core waves. *Nature Geoscience*, *11*(4), 280–284. https://doi.org/10.1038/s41561-018-0075-y

Niu, X., Zhao, D., & Li, J. (2018). Precise relocation of low-frequency earthquakes in Northeast Japan: New insight into arc magma and fluids. *Geophysical Journal International*, *212*(2), 1183–1200. https://doi.org/10.1093/gji/ggx445

Ross, Z. E., Meier, M., & Hauksson, E. (2018). *P* wave arrival picking and first-motion polarity determination with deep learning. *Journal of Geophysical Research: Solid Earth*, *123*(6), 5120–5129. https://doi.org/10.1029/2017JB015251

Ross, Z. E., Meier, M., Hauksson, E., & Heaton, T. H. (2018). Generalized seismic phase detection with deep learning. *Bulletin of the Seismological Society of America*, *108*(5A), 2894–2901. https://doi.org/10.1785/0120180080

Ross, Z. E., White, M. C., Vernon, F. L., & Ben-Zion, Y. (2016). An improved algorithm for real-time S-wave picking with application to the (augmented) ANZA network in Southern California improved algorithm for real-time S-wave picking with application to ANZA. *Bulletin of the Seismological Society of America*, *106*(5), 2013–2022. https://doi.org/10.1785/0120150230

Saragiotis, C., Hadjileontiadis, L., & Panas, S. (2002). PAI-S/K: A robust automatic seismic P phase arrival identification scheme. *IEEE Transactions on Geoscience and Remote Sensing*, *40*(6), 1395–1404. https://doi.org/10.1109/TGRS.2002.800438

Sharma, B. K., Kumar, A., & Murthy, V. M. (2010). Evaluation of seismic events detection algorithms. *Journal of the Geological Society of India*, *75*(3), 533–538. https://doi.org/10.1007/s12594-010-0042-8

Shelly, D. R., Beroza, G. C., & Ide, S. (2007). Non-volcanic tremor and low-frequency earthquake swarms. *Nature*, *446*(7133), 305–307. https://doi.org/10.1038/nature05666

Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *ArXiv:1409.1556 [Cs]*. http://arxiv.org/abs/1409.1556

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

Stevenson, P. R. (1976). Microearthquakes at Flathead Lake, Montana: A study using automatic earthquake processing. *Bulletin of the Seismological Society of America*, *66*(1), 61–80.

Tkalčić, H., & Phm, T.-S. (2018). Shear properties of Earth's inner core constrained by a detection of J waves in global correlation wavefield. *Science*, *362*(6412), 329–332. https://doi.org/10.1126/science.aau7649

Wang, J., & Zhao, D. (2012). P wave anisotropic tomography of the Nankai subduction zone in Southwest Japan. *Geochemistry, Geophysics, Geosystems*, *13*, Q05017. https://doi.org/10.1029/2012GC004081

Wessel, P., & Smith, W. H. F. (1998). New, improved version of generic mapping tools released. *Eos, Transactions American Geophysical Union*, *79*(47), 579–579. https://doi.org/10.1029/98EO00426

Xie, S., & Tu, Z. (2017). Holistically-nested edge detection. *International Journal of Computer Vision*, *125*(1-3), 3–18. https://doi.org/10.1007/s11263-017-1004-z

Yang, W., Hauksson, E., & Shearer, P. M. (2012). Computing a large refined catalog of focal mechanisms for Southern California (1981-2010): Temporal stability of the style of faulting. *Bulletin of the Seismological Society of America*, *102*(3), 1179–1194. https://doi.org/10.1785/0120110311

Zhao, D., Hasegawa, A., & Horiuchi, S. (1992). Tomographic imaging of *P* and *S* wave velocity structure beneath northeastern Japan. *Journal of Geophysical Research*, *97*(B13), 19,909–19,928. https://doi.org/10.1029/92JB00603

Zhao, D., Matsuzawa, T., & Hasegawa, A. (1997). Morphology of the subducting slab boundary in the northeastern Japan arc. *Physics of the Earth and Planetary Interiors*, *102*(1-2), 89–104. https://doi.org/10.1016/S0031-9201(96)03258-X

Zhao, D., Yanada, T., Hasegawa, A., Umino, N., & Wei, W. (2012). Imaging the subducting slabs and mantle upwelling under the Japan Islands: Subducting slabs and mantle upwelling. *Geophysical Journal International*, *190*(2), 816–828. https://doi.org/10.1111/j.1365-246X.2012.05550.x

Zheng, X., Yao, Z., Liang, J., & Zheng, J. (2010). The role played and opportunities provided by IGP DMC of China national seismic network in Wenchuan earthquake disaster relief and researches. *Bulletin of the Seismological Society of America*, *100*(5B), 2866–2872. https://doi.org/10.1785/0120090257

Zhu, W., & Beroza, G. C. (2019). PhaseNet: A deep-neural-network-based seismic arrival time picking method. *Geophysical Journal International*, *216*, 261–273. https://doi.org/10.1093/gji/ggy423

Reference From the Supporting Information

Kingma, D. P., & Lei, J. (2015). Adam: A method for stochastic optimization. *The 3rd International Conference for Learning Representations*, San Diego. https://arxiv.org/abs/1412.6980