Focal mechanisms of small earthquakes beneath the whole Japan Islands based on first-motion polarities picked using deep learning

Takahiko Uchide[1,1]

[1]National Institute of Advanced Industrial Science and Technology

November 30, 2022

Abstract

Knowledge of crustal stress field is essential for understanding tectonics and earthquake generation. One way to estimate the crustal stress field is based on focal mechanisms of earthquakes. This study investigated focal mechanisms of ~ 110 thousand microearthquakes in Japan Islands shallower than 20 km based on the first-motion polarities picked by a simple neural network model, which was trained using two data sets: moderate to large earthquakes all over Japan and microearthquakes in two regions in Japan. The threshold of the confidence score from the neural network model was so chosen as to maximize the overall quality of focal mechanism solutions. The P- and T-axes of the numerous focal mechanism solutions provide more detailed distributions of crustal stress field. For example, in Chugoku region, there exist slight differences in the trend of P-axes azimuths between the northern and southern areas are observed, corresponding to the geodetic observations in space.

Focal mechanisms of small earthquakes beneath the whole Japan Islands based on first-motion polarities picked using deep learning

Takahiko Uchide[1,*]

[1] Geological Survey of Japan, National Institute of Advanced Industrial Science and Technology (AIST)

* Corresponding Author: Takahiko Uchide (t.uchide@aist.go.jp)

Key Points

- Determined focal mechanisms of ~ 110,000 shallow inland earthquakes in Japan based on first-motion polarities by a neural network model.
- Chose a threshold of the confidence score of the polarities to maximize the overall quality of focal mechanism solutions.
- The numerous focal mechanism solutions indicate the crustal stress field at a fine scale.

Abstract

Knowledge of crustal stress field is essential for understanding tectonics and earthquake generation. One way to estimate the crustal stress field is based on focal mechanisms of earthquakes. This study investigated focal mechanisms of ~ 110 thousand microearthquakes in Japan Islands shallower than 20 km based on the first-motion polarities picked by a simple neural network model, which was trained using two data sets: moderate to large earthquakes all over Japan and microearthquakes in two regions in Japan. The threshold of the confidence score from the neural network model was so chosen as to maximize the overall quality of focal mechanism solutions. The P- and T-axes of the numerous focal mechanism solutions provide more detailed distributions of crustal stress field. For example, in Chugoku region, there exist slight differences in the trend of P-axes azimuths between the northern and southern areas are observed, corresponding to the geodetic observations in space.

Plain Language Summary

The Earth’s tectonic activities, such as creeping, mountain building, and earthquakes, are caused by internal forces (stress). Understanding the stress in the crust is important for understanding such tectonic activities and assessing future earthquakes. Conversely, earthquakes, especially ambient microearthquakes, can be a tool to investigate the crustal stress. We can estimate the focal mechanisms (fault geometry and slip direction of an earthquake) by considering the initial part of seismic waves goes up or down (first-motion polarities) at seismic stations. In this study, massive data from Japan were used. The polarities of ~ 2 million seismic data were measured using deep learning technique, and finally the focal mechanisms of ~ 110,000 earthquakes were determined. The result indicates the crustal stress distribution. Combination with other information such as the

ground surface deformation from geodesy, geography, and geology can enhance our understanding.

1. Introduction

Crustal stress field data are crucial to understand tectonics and seismic activity; however, measuring it at depths over a wide area is a challenge. Direct measurements at specific boreholes (e.g., Wu et al., 2007; Huffman et al., 2016; Brodsky et al., 2017; Townend et al., 2017) offer detailed information but only for one point. In contrast, seismology provides indirect measurements with more uncertainty but for a wide area. The focal mechanisms, which indicate the directions of fault plane and slip, indicate the orientation of the seismogenic stress. The World Stress Map (Heidbach et al., 2008; Heidbach et al., 2016; Heidbach et al., 2018) compiles this information all over the world.

In the past, routinely determined moment tensor solutions were used for estimating regional stress fields (Terakawa & Matsu'ura, 2010; Hardebeck, 2015). However, blank areas still exist even in seismically active area such as Japan Islands. More complete knowledge of the seismogenic stress field requires focal mechanisms of microearthquakes, especially in low seismicity areas (e.g., Imanishi et al., 2011; Imanishi et al., 2012; Matsumoto et al., 2015). Comprehensive investigations of microearthquake focal mechanisms reveal the regional stress field (e.g., Iio et al., 2018; Imanishi et al., 2019).

The focal mechanisms of moderate or larger earthquakes can be automatically determined using the full waveform from the local (e.g., Dreger & Helmberger, 1993;

Fukuyama et al., 1998) or global seismic network (e.g., Ekström et al., 2012). However, the mechanisms of small earthquakes cannot be similarly determined because of the difficulty of modeling high-frequency seismograms. We usually use the first-motion polarity: the vertical component initially goes either upward or downward. Automatic polarity-picking methods, such as one based on the sign of the first extremum after the P arrival (Nakamura, 2004; Chen & Holland, 2016), have been developed. Pugh et al. (2016a) proposed a Bayesian approach using the first extremum and a probability function of P arrival time. Recently, deep learning has enabled us to automatically pick the polarity (Ross et al., 2018; Hara et al., 2019). Thus, we are technically ready to investigate large number of microearthquakes.

The aim of this study was to obtain the focal mechanism solutions in Japan Islands, one of the most seismically active regions in the world. The first-motion polarities were picked using a neural network model and seismic data from nationwide seismic networks. Finally, the focal mechanism solutions and spatial trends in P- and T-axes were studied.

2. Training the Neural Network Model

2.1. Data

The training of the neural network model was performed in two stages. In the first stage, the Hi-net data of 18,000 earthquakes with P arrival and polarity data in the JMA catalog were used. Most of these earthquakes are larger than *M* 3 (Figure S1). The whole data were then spatially divides into the training and validation data sets (Figure 1a, Table 1).

In the second stage, I used the P arrival time and polarity of microearthquakes in Kanto and Chugoku regions, manually picked by Geological Survey of Japan, National Institute of Advanced Industrial Science and Technology (AIST). The Kanto data were used by Imanishi et al. (2019). The spatial distributions of used events in Kanto and Chugoku regions are shown in Figures 1b and 1c, respectively. The number of seismogram sets and earthquakes is summarized in Table 1.

In the both stages, seismograms of three components (up-down, north-south, and east-west) were used. Each component had 256 samples: 156 samples before and 100 samples after P arrival. The samples are 2.56-s long, as the data were sampled at 100 Hz. Low-frequency noise was removed by applying a high-pass filter at 1 Hz. I emphasized the initial portion of the P-wave by clipping seismograms at a certain threshold.

Furthermore, I augmented the data four times by flipping all three components, rotating horizontal components by randomly selected angles, and time-shifting. The flipping procedure equalizes the number of positive and negative polarity data. The time shift addresses potential misalignment of data due to uncertainties in the arrival time picking.

Later, I examined various values of the clipping threshold and the time-shift range.

2.2. Design of the Neural Network Model

Figure 2 summarizes the neural network used in this study. The input of the neural network models is a three-component 256-sample long seismogram set where the 156th sample corresponds to the P-arrival time already picked either manually or automatically.

The output comprises two scores corresponding to the upward and downward polarities. Note that, in the case of Southern California, Ross et al. (2018) classified the polarity as “Up,” “Down,” and “Unknown”; however, in this study, the “Unknown” class is not set. The data set contained many seismograms with impulsive onset but no polarity information (e.g., Figure S2), and the lack of polarity information does not mean “Unknown” in this case.

I designed a simple neural network model (Figure 2) similar to the ones used in prior studies (Ross et al., 2018; Hara et al., 2019). The neural network model started with two convolution layers, followed by three units composed of convolution, batch-normalization (Ioffe & Szegedy, 2015), and pooling layers. The models ended with two fully connected layers. The kernel size of the convolutional layers was 11. For all but the final layers, theactivation function was the Rectified Linear Units (ReLU) (Nair & Hinton, 2010); SoftMax function was chosen for the final layer:

$$\mathrm{softmax}(\mathbf{z})_i = \frac{\exp(z_i)}{\sum_j \exp(z_j)}, \#(1)$$

where $\mathbf{z} = (z_1, z_2)$ is the output of the final layer corresponding to the positive and negative polarities, respectively. Then, the outputs are non-negative, and their summation is always 1. In order to address the overfitting problem, the dropout technique (Srivastava et al., 2014) was adopted: 50 % of randomly selected perceptions were muted during the training. The loss was evaluated by the negative log-likelihood function and the parameters of the neural network model were updated by back-propagating the loss (Rumelhart et al., 1986) optimized by the adaptive moment estimation (Adam) method (Kingma & Ba, 2014).

2.3. Result

Hundred cases with randomly selected clipping thresholds in the range of 10^{-6} to 10^{-4} m/s and the half-width of the time-shift ranging 0–30 samples were examined. The result was evaluated based on the loss value for the test data set. The result shows that the shorter half-width of the time-shift range, the smaller is the loss (Figure 3a). The clipping threshold has no correlation with the loss value (Figure 3b).

Hereafter, I did not apply time-shift and used 10^{-5} m/s as the clipping threshold. The neural network model was trained using these values. The precision-recall curve of the trained model is shown in Figure 3c.

3. Application to Crustal Earthquakes in Japan

I applied the trained model to event data of earthquakes that occurred in the period 2005–2019 at depths less than 20 km within the coastline, excluding the events for which polarity information is already available in the catalog (Table 1). I used seismograms from Hi-net and the JMA seismic network with P-wave arrival times in the JMA catalog. Preprocessing was done in the same way as the training. Good results were obtained for polarity picking with high scores (Figure 4), even in noisy cases.

The focal mechanisms were determined using polarity information with scores larger than a confidence threshold and the HASH code (Hardebeck & Shearer, 2002, 2008). The quality of focal mechanisms depends on the confidence threshold (Figure 3d). If the threshold is too high, the very small number of polarity picks cannot constrain focal mechanisms well. If the threshold is lower than 0.7, the fraction of A and B ranks given by the HASH code (Hardebeck & Shearer, 2008) is almost constant. I adopted a confidence threshold of 0.7.

Figure 5 shows the estimated focal mechanisms and their P- and T- axes in addition to the NIED F-net Moment Tensor solutions for reference. The focal mechanisms of 113,700 events are estimated, while those of 6830 events are undetermined because the number of stations was smaller than 8. Ranks A, B, C, and D by the HASH code (Hardebeck & Shearer, 2008) were given to 1060, 17890, 36958, and 50962 events, respectively. The focal mechanism solutions cover much more space than those in a routine catalog.

The obtained P- and T-axes (Figures 5c and 5d, respectively) are well consistent with stress regimes reported in prior studies: north-south extensional stress field in Kyushu region (Matsumoto et al., 2015; Savage et al., 2016); normal faulting earthquakes in the area of Fukushima-Hamadori and northern Ibaraki prefecture (Imanishi et al., 2012).

4. Discussion

It may be surprising that the narrower time-shift range of the data, the better is the model performance, because the time-shifting would make the model more flexible and robust to uncertainties in arrival time picking. There are two potential reasons. One is that the arrival times in the test data were accurate because of careful review by an analyst, and therefore the time-shift was not really required. Another possible reason is the shortage of training data from microearthquakes.

Determination of focal mechanisms from the first-motion polarities of P-waves picked by the trained neural network model is also important for assessing the quality of polarity picking. In this study 50.8 % of the focal mechanism solutions are ranked D or undetermined. In a study on the determination of the focal mechanisms of earthquakes in Southern California using manually picked P-wave polarity and the amplitude ratio of P and S waves (Yang et al., 2012), the results showed that 56.6 % (101,309 out of 178,899 events) of the events were ranked D, comparable to the result of this study. Thus, this study yields a reasonable quality of the P-wave first-motion polarity picking, though the comparison is not simple because of many factors including the differences in the observational conditions such as the magnitude range and station density. Focal mechanism determination can be improved in several ways: introduction of P-wave amplitude (e.g., Matsushita & Imanishi, 2015; Pugh et al., 2016b) and the ratio of P- and S-wave amplitudes (Hardebeck & Shearer, 2003; Yang et al., 2012), as well as the advances in the P-wave polarity picking.

The quality of the focal mechanism solutions is shown by region in Figure 3e. In particular, the quality in Hokkaido region is much worse than in other regions. The reason was examined by focusing on the number of stations. First, the quality of the focal mechanism solutions is well correlated with the number of stations (Figure 3f). Next, the number of usable stations is smaller in Hokkaido than in other regions. This is probably because of the spatial density of seismic stations (Figure 5e). Hence, it is more difficult to determine the focal mechanisms in Hokkaido than elsewhere.

We see interesting features in the spatial distribution of the P- and T-axis azimuths (Figures 5c and 5d). For example, in Chugoku region, the P-axes strike in the east–west direction in the northern area (San-in area), whereas those strike in the NW–SE direction in the southern one (Sanyo area). The contrast in the P-axis azimuths in western Tottori was reported by Kawanishi et al. (2009). This study too shows a similar trend over the whole Chugoku region. This contrast geographically corresponds to the San-in shear zone (Meneses-Gutierrez & Nishimura, 2020). A combination of this study with geodetic implications will enhance our understanding of seismotectonics.

In spite of the dramatic increase in focal mechanism solutions, there are still blank areas in Japan Islands. The seismicity is quite low in such areas. Hence, this kind of study may need to be performed even for smaller earthquakes, which is a greater challenge than that tackled in the present study. Additional campaign seismic observations may improve the focal mechanism solutions of very small earthquakes. In addition, combining these observations with various observations including geological, geographical, and geodetical ones will improve our understanding of the crustal stress field and its origin.

5. Conclusions

In this study, the focal mechanisms of small to microearthquakes are estimated for better understanding of the crustal stress fields in Japan Islands. The focal mechanisms were derived using the P-wave first-motion polarities picked by a neural network model that takes three-component seismograms with P arrival times as the input. The focal mechanisms of almost all microearthquakes over the whole of Japan Islands were successfully determined. The focal mechanism solutions are generally consistent with the stress regime on a large scale. The slight but clear differences in the P-axis azimuths in the northern and southern parts of Chugoku region are consistent with the geodetic observations for this region. The results of this study will be useful for revealing the crustal stress field, and thus, for assessing the past and current tectonic activities and future earthquake generation.

Acknowledgements

I thank Kazutoshi Imanishi and Reiken Matsushita for providing phase data for microearthquakes in Kanto and Chugoku regions in Japan. I also thank the NIED, especially Takanori Matsuzawa, for helping me prepare the large seismic data set. I used seismic data from NIED Hi-net (National Research Institute for Earth Science and Disaster Resilience, 2020) and JMA available at http://www.hinet.bosai.go.jp/?LANG=en (last accessed on 25 March 2020), the phase data from JMA Unified Earthquake Catalog, available at http://www.data.jma.go.jp/svd/eqev/data/bulletin/eqdoc_e.html (last accessed on 25 March 2020) and http://www.hinet.bosai.go.jp/?LANG=en (last accessed on 25 March 2020), and the moment tensor solutions by NIED F-net project (Fukuyama et al., 1998) available at http://www.fnet.bosai.go.jp/top.php?LANG=en (last accessed on 25 March 2020). The data analyses in this study were performed using PyTorch (Paszke et al., 2019), ObsPy (Beyreuther et al., 2010; Megies et al., 2011; Krischer et al., 2015), HASH (Hardebeck & Shearer, 2002, 2008), and HASHpy (Williams, 2014). I used Generic Mapping Tools (Wessel et al., 2013) for generating Figures 1, 3, 5, and S1, and Matplotlib

(Hunter, 2007) for Figures 4 and S2. This work was supported by Mitsubishi Foundation and AIST EDGE Runners project. In this work, the computation facility of the AI Bridging Cloud Infrastructure (ABCI) maintained by AIST was employed.

References

Amante, C., & Eakins, B. W. (2009). *ETOPO1 1 Arc-Minute Global Relief Model: Procedures, Data Sources and Analysis*. NOAA Technical Memorandum NESDIS NGDC-24. National Geophysical Data Center, NOAA. https://doi.org/10.7289/V5C8276M

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

Brodsky, E. E., Saffer, D., Fulton, P., Chester, F., Conin, M., Huffman, K., et al. (2017). The postearthquake stress state on the Tohoku megathrust as constrained by reanalysis of the JFAST breakout data. *Geophysical Research Letters, 44*(16), 8294-8302. https://doi.org/10.1002/2017gl074027

Chen, C., & Holland, A. A. (2016). PhasePApy: A Robust Pure Python Package for Automatic Identification of Seismic Phases. *Seismological Research Letters, 87*(6), 1384-1396. https://doi.org/10.1785/0220160019

Dreger, D. S., & Helmberger, D. V. (1993). Determination of source parameters at regional distances with three-component sparse network data. *Journal of Geophysical Research, 98*(B5), 8107-8125. https://doi.org/10.1029/93JB00023

Ekström, G., Nettles, M., & Dziewoński, A. M. (2012). The global CMT project 2004–2010: Centroid-moment tensors for 13,017 earthquakes. *Physics of the Earth and Planetary Interiors, 200–201*, 1-9. https://doi.org/10.1016/j.pepi.2012.04.002

Fukuyama, E., Ishida, M., Dreger, D. S., & Kawai, H. (1998). Automated seismic moment tensor determination by using on-line broadband seismic waveforms. *Zisin 2nd Series, 51*(1), 149-156. https://doi.org/10.4294/zisin1948.51.1_149 (*in Japanese with English abstract*)

Hara, S., Fukahata, Y., & Iio, Y. (2019). P-wave first-motion polarity determination of waveform data in western Japan using deep learning. *Earth, Planets and Space, 71*(1), 127. https://doi.org/10.1186/s40623-019-1111-x

Hardebeck, J. L. (2015). Stress orientations in subduction zones and the strength of subduction megathrust faults. *Science, 349*(6253), 1213-1216. https://doi.org/10.1126/science.aac5625

Hardebeck, J. L., & Shearer, P. M. (2002). A new method for determining first-motion focal mechanisms. *Bulletin of the Seismological Society of America, 92*(6), 2264-2276.

Hardebeck, J. L., & Shearer, P. M. (2003). Using S/P amplitude ratios to constrain the focal mechanisms of small earthquakes. *Bulletin of the Seismological Society of America, 93*(6), 2434-2444.

Hardebeck, J. L., & Shearer, P. M. (2008). HASH: A FORTRAN program for computing earthquake first-motion focal mechanisms – v1.2 –. Retrieved from http://earthquake.usgs.gov/research/software/HASH/hash.v1.2.tar.gz

Heidbach, O., Rajabi, M., Cui, X., Fuchs, K., Müller, B., Reinecker, J., et al. (2018). The

World Stress Map database release 2016: Crustal stress pattern across scales. *Tectonophysics, 744*, 484-498. https://doi.org/https://doi.org/10.1016/j.tecto.2018.07.007

Heidbach, O., Rajabi, M., Reiter, K., Ziegler, M., & Team, W. (2016). *World Stress Map Database Release 2016*. GFZ Data Services. https://doi.org/10.5880/WSM.2016.001

Heidbach, O., Tingay, M., Barth, A., Reinecker, J., Kurfeß, D., & Müller, B. (2008). *The World Stress Map database release 2008*. https://doi.org/10.1594/GFZ.WSM.Rel2008

Huffman, K. A., Saffer, D. M., & Dugan, B. (2016). In situ stress magnitude and rock strength in the Nankai accretionary complex: a novel approach using paired constraints from downhole data in two wells. *Earth, Planets and Space, 68*(1), 123. https://doi.org/10.1186/s40623-016-0491-4

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

Iio, Y., Kishimoto, S., Nakao, S., Miura, T., Yoneda, I., Sawada, M., & Katao, H. (2018). Extremely weak fault planes: An estimate of focal mechanisms from stationary seismic activity in the San'in district, Japan. *Tectonophysics, 723*, 136-148. https://doi.org/10.1016/j.tecto.2017.12.007

Imanishi, K., Ando, R., & Kuwahara, Y. (2012). Unusual shallow normal-faulting earthquake sequence in compressional northeast Japan activated after the 2011 off the Pacific coast of Tohoku earthquake. *Geophysical Research Letters, 39*(9), L09306. https://doi.org/10.1029/2012GL051491

Imanishi, K., Kuwahara, Y., Takeda, T., Mizuno, T., Ito, H., Ito, K., et al. (2011). Depth-dependent stress field in and around the Atotsugawa fault, central Japan, deduced from microearthquake focal mechanisms: Evidence for localized aseismic deformation in the downward extension of the fault. *Journal of Geophysical Research, 116*(B1), B01305. https://doi.org/10.1029/2010JB007900

Imanishi, K., Uchide, T., Ohtani, M., Matsushita, R., & Nakai, M. (2019). Construction of the crustal stress map in the Kanto region, central Japan. *Bulletin of the Geological Survey of Japan, 70*(3), 273-298. (*in Japanese with English abstract*)

Ioffe, S., & Szegedy, C. (2015). Batch normalization: accelerating deep network training by reducing internal covariate shift. *ArXiv e-prints*. Retrieved from https://ui.adsabs.harvard.edu/#abs/2015arXiv150203167I

Kawanishi, R., Iio, Y., Yukutake, Y., Shibutani, T., & Katao, H. (2009). Local stress concentration in the seismic belt along the Japan Sea coast inferred from precise focal mechanisms: Implications for the stress accumulation process on intraplate earthquake faults. *Journal of Geophysical Research Solid Earth*, *114*(B1). https://doi.org/10.1029/2008jb005765

Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.

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).

https://doi.org/10.1088/1749-4699/8/1/014003

Matsumoto, S., Nakao, S., Ohkura, T., Miyazaki, M., Shimizu, H., Abe, Y., et al. (2015). Spatial heterogeneities in tectonic stress in Kyushu, Japan and their relation to a major shear zone. *Earth, Planets and Space, 67*(1), 172. https://doi.org/10.1186/s40623-015-0342-8

Matsushita, R., & Imanishi, K. (2015). Stress fields in and around metropolitan Osaka, Japan, deduced from microearthquake focal mechanisms. *Tectonophysics, 642*, 46-57. https://doi.org/10.1016/j.tecto.2014.12.011

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*(1), 12. https://doi.org/10.4401/ag-4838

Meneses-Gutierrez, A., & Nishimura, T. (2020). Inelastic deformation zone in the lower crust for the San-in Shear Zone, Southwest Japan, as observed by a dense GNSS network. *Earth, Planets and Space, 72*(1), 10. https://doi.org/10.1186/s40623-020-1138-z

Nair, V., & Hinton, G. E. (2010). *Rectified linear units improve restricted boltzmann machines.* Paper presented at the 27th International Conference on Machine Learning (ICML-10), Haifa, Israel, 807-814.

Nakamura, M. (2004). Automatic determination of focal mechanism solutions using initial motion polarities of P and S waves. *Physics of the Earth and Planetary Interiors, 146*(3), 531-549. https://doi.org/10.1016/j.pepi.2004.05.009

National Research Institute for Earth Science and Disaster Resilience. (2020). *NIED Hi-net*.

https://doi.org/10.17598/NIED.0003

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., et al. (2019). PyTorch: An imperative style, high-performance deep learning library. In *Advances in Neural Information Processing Systems* (Vol. 32, pp. 8024-8035) Curran Associates, Inc.

Pugh, D. J., White, R. S., & Christie, P. A. F. (2016a). Automatic Bayesian polarity determination. *Geophysical Journal International, 206*(1), 275-291. https://doi.org/10.1093/gji/ggw146

Pugh, D. J., White, R. S., & Christie, P. A. F. (2016b). A Bayesian method for microseismic source inversion. *Geophysical Journal International*. https://doi.org/10.1093/gji/ggw186

Ross, Z. E., Meier, M.-A., & 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

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature, 323*(6088), 533-536. https://doi.org/10.1038/323533a0

Savage, M. K., Aoki, Y., Unglert, K., Ohkura, T., Umakoshi, K., Shimizu, H., et al. (2016). Stress, strain rate and anisotropy in Kyushu, Japan. *Earth and Planetary Science Letters, 439*, 129-142. https://doi.org/10.1016/j.epsl.2016.01.005

Shearer, P. M., Prieto, G. A., & Hauksson, E. (2006). Comprehensive analysis of earthquake source spectra in Southern California. *Journal of Geophysical Research, 111*(B6), B06303. https://doi.org/10.1029/2005JB003979

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: a simple way to prevent neural networks from overfitting. *The Journal of Machine Learning Research, 15*(1), 1929-1958.

Terakawa, T., & Matsu'ura, M. (2010). The 3-D tectonic stress fields in and around Japan inverted from centroid moment tensor data of seismic events. *Tectonics*, *29*(6), TC6008. https://doi.org/10.1029/2009TC002626

Townend, J., Sutherland, R., Toy, V. G., Doan, M.-L., Célérier, B., Massiot, C., et al. (2017). Petrophysical, Geochemical, and Hydrological Evidence for Extensive Fracture-Mediated Fluid and Heat Transport in the Alpine Fault's Hanging-Wall Damage Zone. *Geochemistry, Geophysics, Geosystems, 18*(12), 4709-4732. https://doi.org/10.1002/2017GC007202

Wessel, P., Smith, W. H. F., Scharroo, R., Luis, J., & Wobbe, F. (2013). Generic Mapping Tools: Improved Version Released. *Eos, Transactions American Geophysical Union, 94*(45), 409-410. https://doi.org/10.1002/2013eo450001

Williams, M. C. (2014). HASHpy. Retrieved from https://doi.org/10.5281/zenodo.9808

Wu, H.-Y., Ma, K.-F., Zoback, M., Boness, N., Ito, H., Hung, J.-H., & Hickman, S. (2007). Stress orientations of Taiwan Chelungpu-Fault Drilling Project (TCDP) hole-A as observed from geophysical logs. *Geophysical Research Letters, 34*(1). https://doi.org/10.1029/2006gl028050

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

Tables

Table 1

Numbers of seismograms and earthquakes contained in data sets.

Region	Type of Data Set	Seismogram Sets	Earthquakes
All Japan	Training	279,483	17,402
	Validation	7,666	598
Kanto	Training	12,814	1,262
	Validation	784	56
	Test	1,483	113
Chugoku	Training	63,359	2,259
	Validation	7,674	322
	Test	12,838	595
All Japan	Application	1,930,132	113,700

Figures

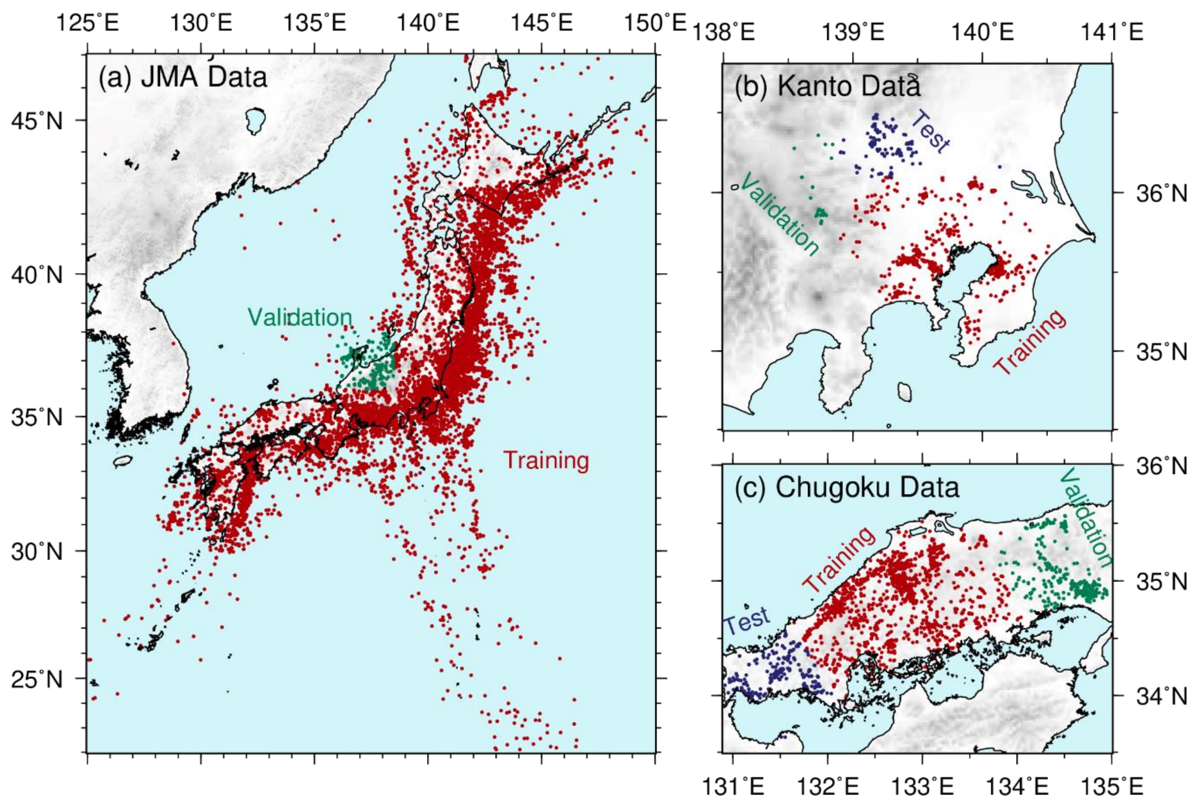


Figure 1

Distribution of the epicenters of the earthquakes used for training (red), validation (green), and testing (blue) of the neural network model. Topography is from ETOPO1 (Amante & Eakins, 2009).

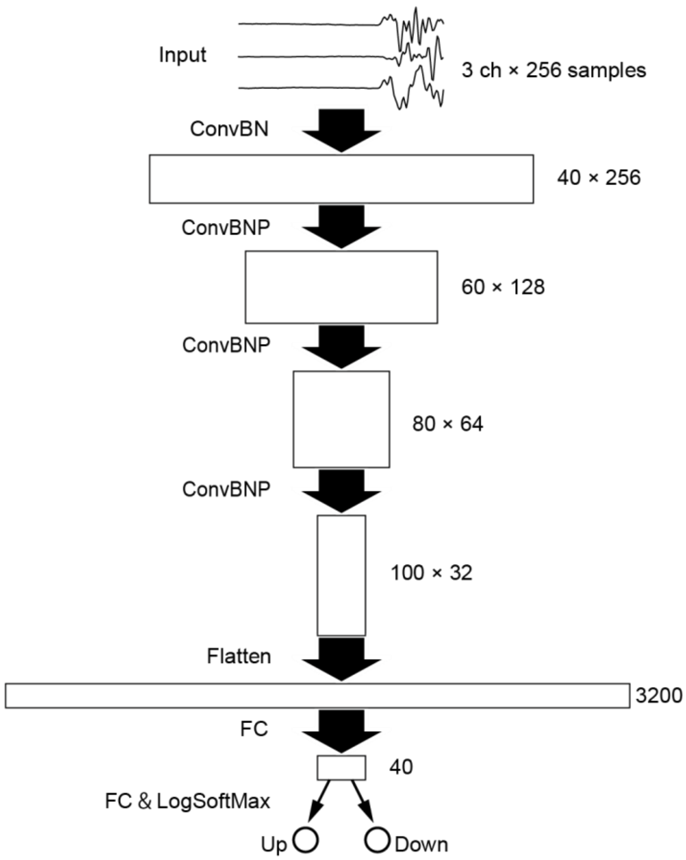
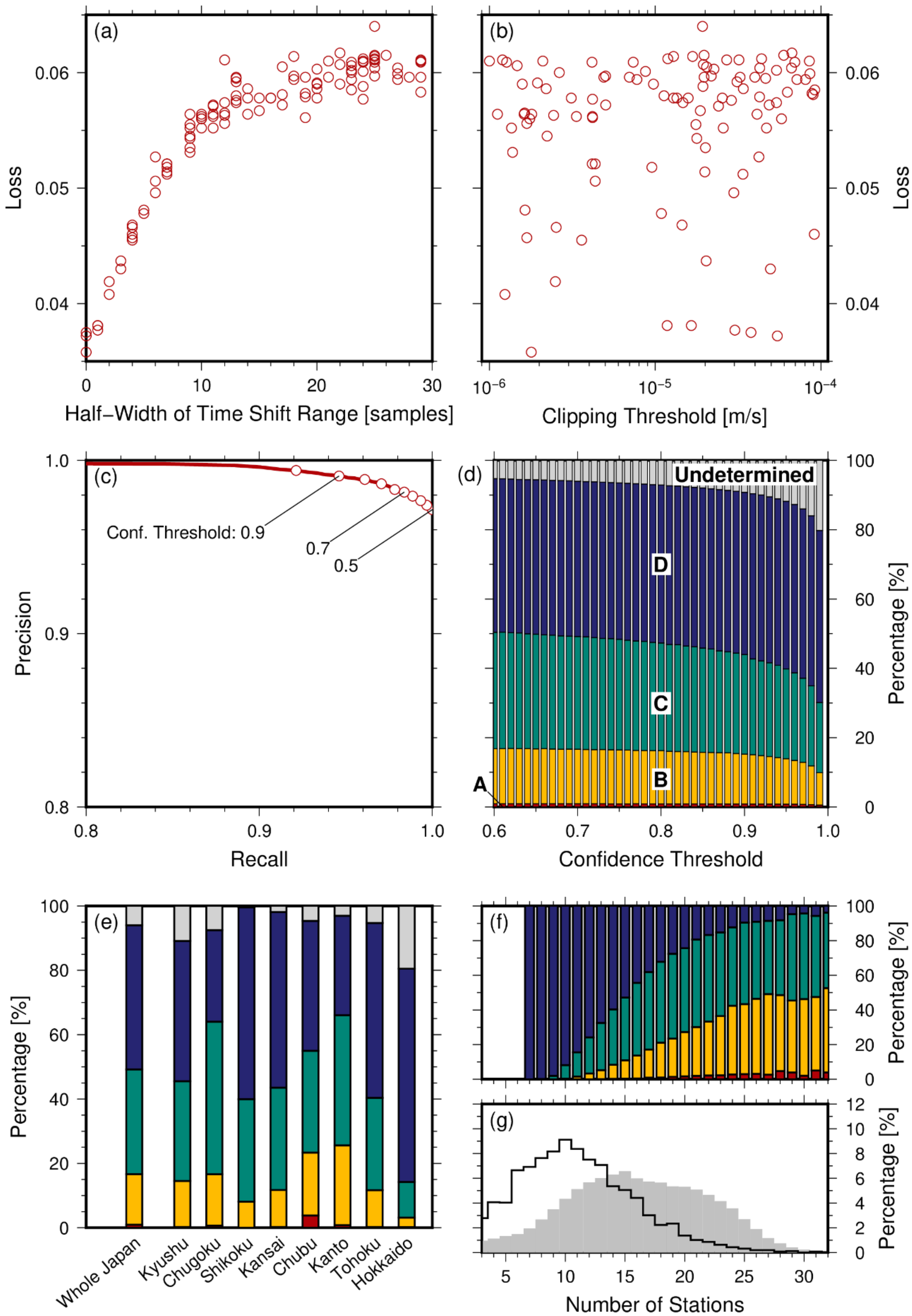


Figure 2

Design of the neural network model. Numbers on the right indicate the number of channels and samples. “ConvBN,” “Conv BNP,” and “FC” denote convolution and batch normalization layers; convolution, batch normalization, and pooling layers; and fully connected layers, respectively.


(a)
Loss
0.06
0.05
0.04
0
10
20
30
Half−Width of Time Shift Range [samples]
(b)
Loss
0.06
0.05
0.04
10^{-6}
10^{-5}
10^{-4}
Clipping Threshold [m/s]
(c)
1.0
0.9
0.8
Precision
Conf. Threshold: 0.9
0.7
0.5
0.8
0.9
1.0
Recall
(d)
Undetermined
D
C
B
A
100
80
60
40
20
0
Percentage [%]
0.6
0.7
0.8
0.9
1.0
Confidence Threshold
(e)
100
80
60
40
20
0
Percentage [%]
Whole Japan
Kyushu
Chugoku
Shikoku
Kansai
Chubu
Kanto
Tohoku
Hokkaido
(f)
100
80
60
40
20
0
Percentage [%]
(g)
12
10
8
6
4
2
0
Percentage [%]
5
10
15
20
25
30
Number of Stations

394

Figure 3

Summary of the results. Here, (a) and (b) show the loss function values as functions of the time-shift range and the clipping threshold, respectively. (c) Precision-recall curve of the trained model for the test data set. Circles correspond to every 0.05 units of the confidence thresholds. (d) Bar graphs of the rank of focal mechanism solutions as a function of the confidence threshold. (e) Bar graphs of the rank of focal mechanism solutions for the whole of Japan and eight regions. Here a model with a confidence threshold of 0.7 was used. (f) Bar graphs as a function of the number of stations. (g) Histograms of events as a function of the number of stations. The black line and gray shaded region indicate the values for Hokkaido and other regions, respectively.

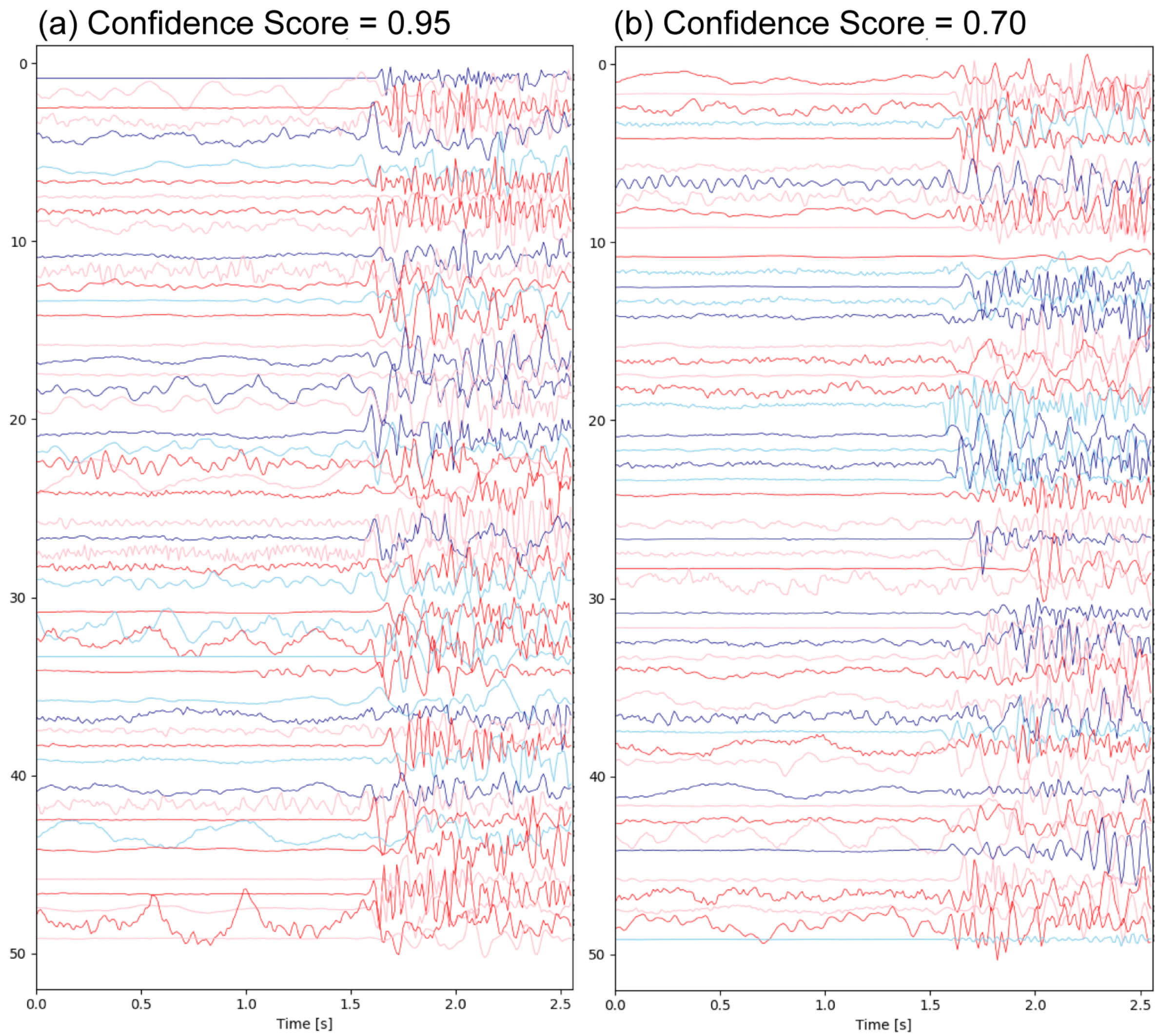
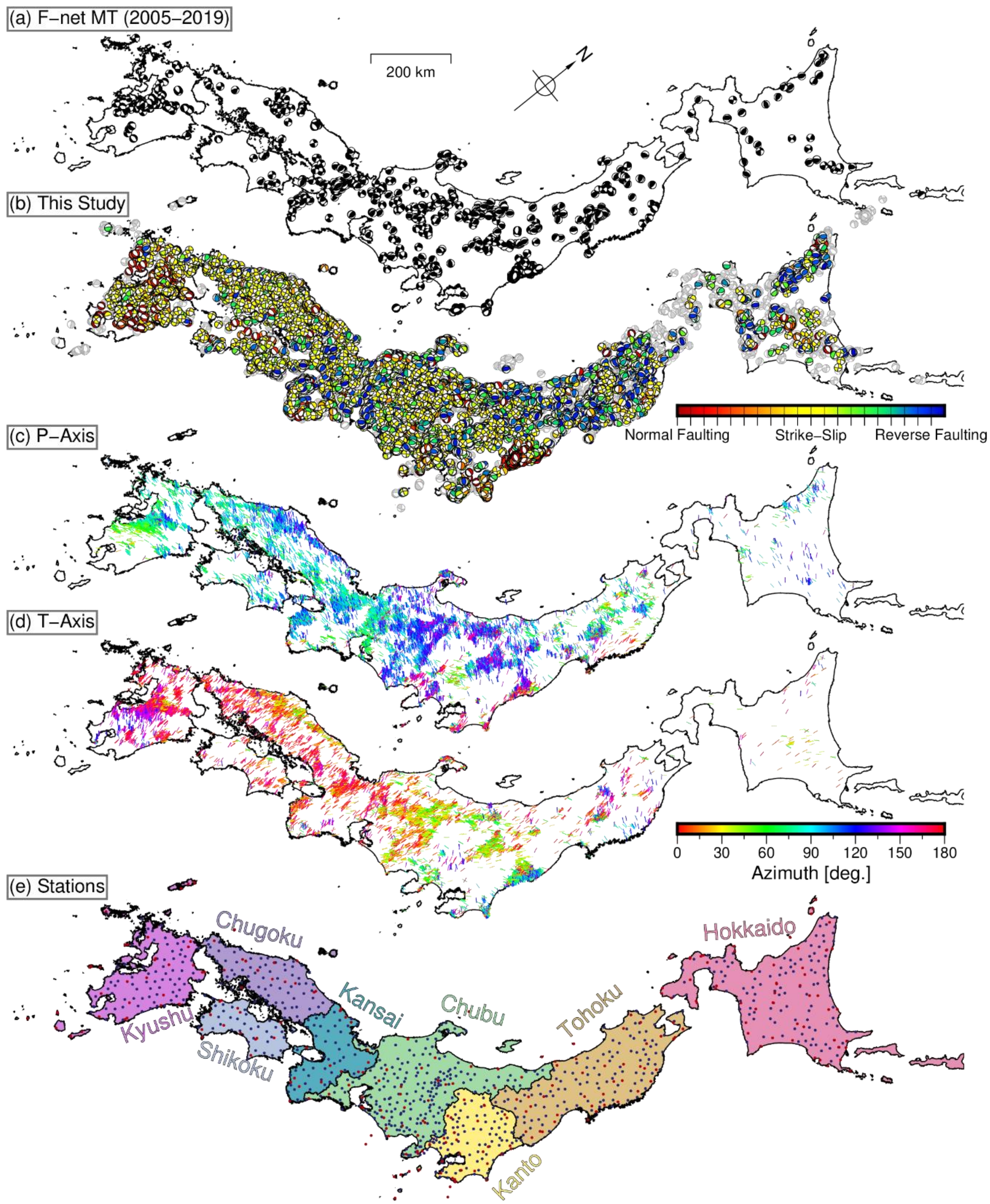


Figure 4

Examples of polarities picked by the neural network model with confidence scores of (a) 0.95 and (b) 0.70. Seismograms with negative polarities are flipped. If correctly picked, the first motion looks positive in this figure. Light and dark colors are alternatively for convenience.



411

412 Figure 5

413 (a) Moment tensor solutions of earthquakes in 2006-2019 by NIED F-net Project

(Fukuyama et al., 1998), for reference. (b) Focal mechanism solutions in this study. Solutions ranked A–C (Hardebeck & Shearer, 2008) are colored according to focal mechanism types (Shearer et al., 2006). Solution ranked D are shown by gray beach balls. (c) Azimuths of the P-axes of the estimated focal mechanism solutions ranked A – C and with less than 30° of plunge. Colors indicate the azimuths. (d) Azimuths of the T-axes. (e) Station distribution.