

# Application of Deep Learning-Based Neural Networks Using Theoretical Seismograms as Training Data for Locating Earthquakes in the Hakone Volcanic Region, Japan

**Daisuke Sugiyama**

JAMSTEC: Kaiyo Kenkyu Kaihatsu Kiko

**Seiji Tsuboi** (✉ [tsuboi@jamstec.go.jp](mailto:tsuboi@jamstec.go.jp))

JAMSTEC CEIST: Kaiyo Kenkyu Kaihatsu Kiko Chikyu Joho Kiban Center <https://orcid.org/0000-0003-3422-6879>

**Yohei Yukutake**

Hot Springs Research Institute

---

## Research Article

**Keywords:** deep learning, hypocenter determination, synthetic seismograms

**Posted Date:** March 16th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-295787/v1>

**License:** © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

1 **Application of deep learning-based neural networks using theoretical seismograms as training**  
2 **data for locating earthquakes in the Hakone volcanic region, Japan**

3

4 Daisuke Sugiyama

5 Research Institute for Value-Added-Information Generation, Japan Agency for Marine-Earth Science  
6 and Technology, Kanazawa-ku, Yokohama, Kanagawa 236-0001 Japan. sugiyamad@jamstec.go.jp

7

8 Seiji Tsuboi

9 Research Institute for Value-Added-Information Generation, Japan Agency for Marine-Earth Science  
10 and Technology, Kanazawa-ku, Yokohama, Kanagawa 236-0001 Japan. [tsuboi@jamstec.go.jp](mailto:tsuboi@jamstec.go.jp)

11

12 Yohei Yukutake

13 Hot Springs Research Institute of Kanagawa Prefecture, Odawara, Kanagawa, Japan.  
14 yukutake@onken.odawara.kanagawa.jp

15

16 Corresponding author: Seiji Tsuboi; e-mail address: [tsuboi@jamstec.go.jp](mailto:tsuboi@jamstec.go.jp)

17

## 18 **Abstract**

19 In the present study, we propose a new approach for determining earthquake hypocentral parameters.  
20 This approach integrates computed theoretical seismograms and deep machine learning. The  
21 theoretical seismograms are generated through a realistic three-dimensional Earth model, and are then  
22 used to create spatial images of seismic wave propagation at the Earth's surface. These snapshots are  
23 subsequently utilized as a training dataset for a convolutional neural network. Neural networks for  
24 determining hypocentral parameters such as the epicenter, depth, occurrence time, and magnitude are  
25 established using the temporal evolution of the snapshots. These networks are applied to seismograms  
26 from the seismic observation network in the Hakone volcanic region in Japan to demonstrate the  
27 suitability of the proposed approach for locating earthquakes. We demonstrate that the determination  
28 accuracy of hypocentral parameters can be improved by including theoretical seismograms for  
29 different earthquake locations and sizes, in the learning dataset for the deep machine learning. Using  
30 the proposed method, the hypocentral parameters are automatically determined within seconds after  
31 detecting an event. This method can potentially serve in monitoring earthquake activity in active  
32 volcanic areas such as the Hakone region.

33

34 **Keywords:** deep learning, hypocenter determination, synthetic seismograms

35

## 36 **Introduction**

37 Currently, the machine learning technique is widely exploited in disciplines of science and technology  
38 including Earth sciences. Specifically, in seismology, combining numerical seismograms and pattern  
39 recognition for earthquake location was proposed by Käufl et al. (2014, 2015, 2016a, 2016b).  
40 According to these studies, the technique can be applied for determining fault parameters. In this study,  
41 we propose a different technique for earthquake location and determining its magnitude. The proposed  
42 approach is based on theoretical seismograms from a realistic Earth model and a deep learning-based  
43 convolutional neural network (CNN). The approach relies on spatial images of seismic wave  
44 propagation at the Earth's surface.

45 The spectral element method (SEM) is used for numerical modeling of seismic wave propagation  
46 in a realistic three-dimensional (3D) Earth model. The 3D SEM was initially employed in seismology  
47 to perform local and regional simulations (Faccioli et al., 1997; Komatitsch, 1997; Komatitsch and  
48 Vilotte, 1998), and later adapted for wave propagation at the Earth scale (Komatitsch and Tromp,  
49 2002a; Komatitsch and Tromp, 2002b; Komatitsch et al., 2005; Tsuboi et al., 2003; Carrington et al.,  
50 2008; Rietmann et al., 2012; Tsuboi et al., 2016). In the present study, the SEM is used to simulate a  
51 regional scale seismic wave propagation problem. The surface propagation of seismic waves at each  
52 time step is calculated for different earthquake locations and sizes, and the values are differentiated to  
53 obtain the seismic wave propagation in terms of the velocity. These seismic wave propagation and

54 temporal evolution images are then utilized as training data for a deep machine learning CNN.  
55 Considering that setting any number of hypocenters or earthquake magnitude is possible, creating a  
56 large volume of training data is easy. Therefore, deep machine learning is expected to improve the  
57 hypocenter determination accuracy.

58 A CNN is a deep learning model that has attracted attention in computer visualization because of  
59 the accuracy of its high image recognition tasks, such as in the ImageNet's large scale visual  
60 recognition challenge (Vaillant et al., 1994; Krizhevsky et al, 2012; Simonyan and Zisserman, 2014;  
61 Szegedy et al, 2015). Artificial neural networks are commonly used for classifying seismograms  
62 (Wang and Teng, 1995: 1997), and recently, CNNs have been employed for earthquake detection  
63 (Perol et al., 2018; Ross et al., 2018; Zhu and Beroza, 2019). In this study, theoretical seismograms  
64 computed for different earthquake locations and magnitudes are utilized for creating images of seismic  
65 wave propagation at the Earth's surface. These snapshots then serve as training data in the training  
66 phase of the CNN to estimate the hypocenter location, earthquake occurrence time, and earthquake  
67 magnitude.

68 The advantage of using our technique in real time hypocenter determination system is its speed.  
69 The learning procedure for creating a neural network is time-intensive, especially if a large training  
70 dataset is involved. However, after establishing the neural network, the hypocentral parameters are  
71 efficiently estimated within 0.1 s from each seismic wave propagation image using GPGPUs, even if  
72 multiple seismic stations are involved. Considering that the creation of images from seismic stations  
73 requires little time, efficient hypocenter determination is possible through the implementation of our  
74 technique in a local seismic network. In the present study, the proposed technique is tested in the  
75 Hakone volcanic region in Japan, where earthquake swarms closely linked to volcanic activity are  
76 frequent. Regarding earthquake swarms, in 2015, for example, numerous micro earthquakes occurred  
77 in the region within a short duration. Monitoring such activities using conventional automated  
78 hypocenter location software is commonly challenging. In the present study, we demonstrate the  
79 application of the proposed technique for accurate and rapid micro earthquake monitoring in the  
80 Hakone volcanic region in Japan.

81 Another advantage of our technique is the magnitude determination. The magnitude range can be  
82 set arbitrarily if the training data are generated from theoretical seismograms. Therefore, any moment  
83 tensor solution can be employed for producing the theoretical seismograms and generating the training  
84 dataset. This enables the incorporation of high magnitude events, which rarely occur. The lower limit  
85 for magnitude determination using this technique will be determined when the system is implemented  
86 in a local seismic network and the S/N ratio is obtained by monitoring seismic activity. As previously  
87 mentioned, the images used for the training dataset are generated via modeling of seismic wave  
88 propagation using a 3D Earth model and the moment tensor solutions. Therefore, the magnitude  
89 estimation is also expected to be accurate when applied to a real seismic network. Although the

90 examples are of limited scope, the approach exhibits potential for real time monitoring of seismic  
91 activity.

92

### 93 **Theoretical seismograms**

94 The SPECFEM3D program package (Komatitsch and Tromp, 2002a; Komatitsch and Tromp, 2002b;  
95 Komatitsch et al., 2005; Tsuboi et al., 2003) was used to generate theoretical seismograms for a  
96 realistic Earth model using the SEM. The theoretical seismograms were synthesized using  $48 \times 48$ ,  
97 yielding 2304 slices. Each slice was allocated to a CPU core of a parallel supercomputer, and then  
98 subdivided into  $768 \times 768$  grid points, thereby enabling generation of theoretical seismograms with  
99 an accuracy of 5.6 s. To generate a mesh for the computation, the surface topography and attenuation  
100 were included. A 3D seismic wave speed model of the Hakone region constructed by Yukutake et al.,  
101 (2015) was also considered. Three-components displacement seismograms, lasting 30 s, were  
102 synthesized at seismographic stations operated by the Hot Springs Research Institute of the Kanagawa  
103 Prefecture using a sampling interval of 0.1 s. A map showing the stations in this network is in Figure  
104 1. The theoretical seismograms were then differentiated once to create velocity seismograms.  
105 Subsequently,  $32 \times 32$  images occupying  $0.01^\circ \times 0.01^\circ$  grids were generated from the seismograms  
106 synthesized for various stations. To create the training dataset, the hypocenter was set in the Hakone  
107 region using a spatial distance of  $0.02^\circ$  to produce the seismic wave propagation images (Figure 2).  
108 The hypocentral depth was varied over a vertical distance of 2 km, and the maximum depth was set at  
109 10 km. The occurrence time and moment magnitude ( $M_w$ ) were also randomly altered between  
110 approximately 2.2 and 4.5, to create images for the training data required for deep learning of the CNN.  
111 The moment tensor of the source can be varied while creating the training data; however, this was not  
112 performed in the present study. A typical reverse fault mechanism was assumed for the region, and  
113 was not expected to affect the estimation accuracy. An example of a wave propagation image versus  
114 time is displayed in Figure 3. Among the 1,998 earthquakes involved, 80% of the events were included  
115 in the training dataset. Consequently, about 600,000 images of seismic-wave propagation at the surface  
116 were generated, with a  $32 \times 32$  and time interval of 0.1 s.

117

### 118 **Convolutional Neural Network**

119 In the present study, two types of CNNs were tested. First, the test using an ordinary CNN is described.  
120 A  $32 \times 32$  image served as the input, and hypocentral parameters including the latitude, longitude,  
121 depth, occurrence time, and magnitude, were estimated through regression. The network architecture  
122 employed was similar to that in previous studies (LeCun et al., 1999; Onishi and Sugiyama, 2017). As  
123 the convolutional layer, a set of learnable filters were applied to the input image to extract the  
124 characteristic features, while a pooling layer was used to reduce sensitivity associated with the location  
125 of the characteristic feature. The fully connected layer conducted the reasoning based on the output

126 from the convolutional and pooling layers. These layers were sequentially connected to produce a  
 127 CNN capable of associating an input image to a specific hypocenter parameter. TensorFlow was  
 128 employed as the framework to establish a LeNet-based CNN (LeCun et al., 1999).

129 The input image data created using the seismic wave propagation at the Earth's surface comprised  
 130  $32 \times 32$  pixels. The convolutional layer map and the input image were utilized to generate an output  
 131 image using convolutional filters. The dimensions of the convolutional filters used were  $3 \times 3 \times n$  or  
 132  $7 \times 7 \times n$ , and can be represented as follows:

$$133 \quad u_{i,j,k} = f \left( b_k + \sum_{i'=-2}^2 \sum_{j'=-2}^2 \sum_{k'=1}^n z_{i+i',j+j',k'} \cdot w_{i',j',k',k} \right), \quad (1)$$

134 where  $u_{i,j,k}$  and  $z_{i,j,k}$  are the  $(i, j)$ -th pixel of the  $k$ -th channel in the output and input images for the  
 135 layer, respectively. The weights  $w_{i',j',k',k}$  constitute a filter that was applied to the  $k'$ -th channel of  
 136 the input image,  $b_k$  is the bias for the  $k$ -th channel, and  $n$  is the number of channels in the input layer.  
 137 In this study, as only the vertical seismogram component was used, the number of channels ( $n$ ) is 1.  
 138 Considering  $f(x)$ , a ReLU function which is commonly used for regression analysis in neural  
 139 networks, this can be expressed as follows:

$$140 \quad f(x) = \begin{cases} 0, & x \leq 0 \\ x, & x > 0 \end{cases} \quad (2)$$

141 Zero padding was not utilized on the input image, and the output image size was reduced by four pixels  
 142 in two directions. The pooling layer was processed using max pooling with a  $2 \times 2$  filter and a stride  
 143 of 2, thereby downsampling the image size from  $h \times h \times n$  to  $h/2 \times h/2 \times n$  through the following  
 144 expression:

$$145 \quad u_{i,j,k} = \max \left\{ z_{2j+j',2i+i',k} \mid i', j' = (0, 1) \right\}. \quad (3)$$

146 In the final connected layer, to avoid overfitting, a 50% dropout was applied. Dropout is a procedure  
 147 for avoiding overfitting by creating a pattern without updating the neural network weight of a  
 148 designated layer. A modified stochastic gradient descent (Kingma and Ba, 2015) served as the  
 149 optimizer, while the learning proceeded using a batch size of 50 and epoch of 250. The calculated root  
 150 mean square error (RMSE) for the learning data was then employed as a loss function.

151 Batch normalization was applied (Ioffe and Szegedy, 2015) before calculating the activation function.  
 152 This process renders characteristic features uncorrelated by converting the mean and standard  
 153 deviation of any selected batch to 0 and 1, respectively.

154 The mean and standard deviation for a batch  $\mathbf{B} = \{\mathbf{x}_1 \dots \mathbf{x}_m\}$  were calculated as follows:

155

$$\mu_B = \frac{1}{m} \sum_{i=1}^m \mathbf{x}_i,$$

$$\sigma_B^2 = \frac{1}{m} \sum_{i=1}^m (\mathbf{x}_i - \mu_B)^2, \quad (4)$$

157

158 The normalized data  $\mathbf{y}_i$  can be obtained using the learned parameters  $\gamma$  and  $\beta$  from the following  
159 equations:

160

$$\mathbf{x}_i^* = \frac{\mathbf{x}_i - \mu}{\sqrt{\sigma_B^2 + \varepsilon}}, \quad (5)$$

$$\mathbf{y}_i = \gamma \mathbf{x}_i^* + \beta.$$

161

Where  $\varepsilon$  is a constant error term added for numerical stability.

162

163

164

165

166

167

168

169

170

171

172

173

174

175

176

177

Bayesian optimization (Shahriari et al., 2015) was used to search for optimized hyperparameters for the proposed learning model to avoid manual tuning via trial and error. A grid search was unsuitable because of the high number of combinations for the search target. The Bayesian optimization is a technique for efficiently searching hyperparameters. These parameters are estimated at high resolution using approximate loss distribution functions, and assuming that the loss distribution of the hyperparameters learning model is Gaussian. Following the examination of each case, a highly accurate model was produced using few parameters. The loss function for the Bayesian optimization was the RMSE estimated for the validation data. The hyperparameter search involved the following conditions: the convolution layers varied between 1–5; the fully connected layers ranged between 0–3; the convolution filter size was between 3–9; the fully connected nodes varied between 50–1000; max pooling was applied to each convolution layer; the dropout was between 0–50%; and batch normalization was applied on ReLU, PReLU, LeakyReLU, tanh, and sigmoid activations functions. The model was then constructed after operating 20 GPUs for approximately 2 d. Amplitude ratios of the short-time average (STA) computed for 0.5 sec (for 5 images) to the long-time average (LTA) computed for 5 sec (for 50 images) were determined, and hypocentral parameters estimation was started when STA/LTA exceeded 4.

178

179

180

181

182

183

184

The ConvNetQuake (Perol et al., 2018), a CNN used for classifying earthquakes, was also tested for estimating hypocentral parameters. The results from the ConvNetQuake presented in Table 1 include regression data for hypocentral parameters from the learning stage. The results reveal that although hypocentral parameters can be adequately determined, the precision values are inferior to those from the 3D CNN examined subsequently.

### 3D Convolutional Neural Network

185 Seismic wave propagation with time was tested using 6 s images corresponding to a  $32 \times 32 \times 10$   
186 event as the input. Hypocentral parameters such as the latitude, longitude, depth, occurrence time, and  
187 magnitude were then estimated through regressions. The architecture of the 3D CNN used in the  
188 present study is displayed in Figure 4, with 3D CNN reflecting an extension of 2D CNN (Du Tran et  
189 al., 2015). This involves using a time series event (e.g., 3D structure from  $x \times y \times t$  shown in Figure  
190 4) as input for learning and estimation. 3D CNN has been used for action recognition in movies and  
191 for learning time series data. The 3D CNN learning may involve pretraining using an existing dataset  
192 for human action recognition such as the Kinetics dataset (Carreira and Zisserman., 2017). However,  
193 such a dataset for human activities and materials is unsuitable for the present study. Therefore,  
194 pretraining was not conducted, but rather, a small 3D CNN was used to construct the model  
195 experimentally. A 2DConv+LSTM model (Donahue et al., 2015) combining 2D CNN and long short-  
196 term memory (LSTM) models for time series learning is available. In addition, a two-stream model  
197 (Carreira and Zisserman., 2017) that extracts motion characteristics and learns RGB simultaneously  
198 for time series learning also exists. Searching for optimal parameters using these models should be  
199 considered in future research. Hypocentral parameter estimation starts when the STA/LTA ratio is  
200 higher than 4 as in the ordinary CNN. A duration of 6 s (60 images) was employed, with estimates  
201 performed 20 times using 60 snapshots, thus, yielding hypocentral parameters 8 s after an event  
202 detection.

203 A 3D CNN model generally involves a longer learning time compared to a 2D CNN. In the present  
204 study, we realized that this was linked to the I/O random access, instead of the GPU performance. To  
205 handle this problem, the complete training dataset was loaded to the memory, thereby shortening the  
206 learning time to 1 epoch in 54 minutes which amounts to 78 GPU h for 120 epochs involving the  
207 learning of 5 parameters. Owing to the learning time for a model exceeding the resources of available  
208 computers, parallel GPUs will be required to search models in the future.

209 In addition, time series modeling using 3D CNN appears to overfit the learning data when input data,  
210 such as the duration, is increased. Therefore, overfitting should be avoided by introducing a  
211 generalization technique, and this can be considered for future studies.

212

### 213 **Validation and testing**

214 The 600,000 training data images were divided into training data (80%) and validation data (20%). In  
215 partitioning the training data, earthquakes involved in validation were excluded from the testing  
216 (Figure 2). The network parameters were optimized for estimating the hypocentral parameters by  
217 minimizing discrepancies in the training dataset. After the optimization, these were verified using the  
218 validation dataset. To estimate the earthquake occurrence time, the time for the  $i$ -th image was

219 designated  $t_i$  and the occurrence time  $t_0$  was estimated from each image. Considering that one network  
220 was used for five hypocentral parameters, some parameters such as the magnitude, demonstrated that  
221 the loss function apparently was not improved at certain stages, indicating overfitting. Therefore, the  
222 use of at least two networks; one for the longitude and latitude and another for the depth, magnitude,  
223 and occurrence time seems necessary to prevent overfitting. The network construction to suit each  
224 hypocentral parameter requires further analysis. Each seismic wave propagation image requires  
225 hypocentral parameters output, so that the evolution of the estimates can be observed as the images  
226 are sorted. The testing of each hypocentral parameter during the 3D CNN validation process is shown  
227 in Figure 5. The hypocentral parameter values obtained after 8 s duration were considered the estimates.  
228 The errors associated with the estimated values are comparable to those from the validation dataset,  
229 and the calculated magnitudes for all earthquakes used for validation are presented in Table 1.  
230 According to the results, the estimates from the 3D CNN exhibit higher precision compared to those  
231 from the ordinary CNN. This indicates that seismic wave propagation evolution images are important  
232 for estimating hypocentral parameters using neural networks.

233

#### 234 **Generalization for real data**

235 Subsequently, we attempted to generalize the developed neural network for real seismographic data.  
236 Seismograms recorded in stations operated by the Hot Springs Research Institute of the Kanagawa  
237 Prefecture, the National Research Institute for Earth Science and Disaster Resilience, and Japan  
238 Meteorological Agency were obtained. The data were examined and 159 earthquakes with magnitude  
239 of approximately 2.0, which occurred in the Hakone region, Japan, between 2015 and 2019 were  
240 selected. The data obtained from the stations were 1 min three-components velocity seismograms  
241 associated with the earthquakes. The original 100/200 Hz sampled accelerograms were converted to  
242 velocity seismograms, decimated to 10 Hz, and utilized for generating a  $32 \times 32$  image for each time  
243 step. Since the accuracy of the synthetic seismograms is 5.6 s, 10 images per second represents an  
244 appropriate number of images for estimating the hypocentral parameters. The  $32 \times 32$  images were  
245 employed as inputs for neural network to estimate the parameters. The RMSE between the estimated  
246 parameters and those reported by the Hot Springs Research Institute of the Kanagawa Prefecture for  
247 these earthquakes are presented in Table 1. According to comparisons of the results in Figures 6 and  
248 7, the proposed neural network provides good results. However, for several earthquakes, significant  
249 discrepancies emerged between the results from our method and those reported by the Hot Springs  
250 Research Institute of the Kanagawa Prefecture. These discrepancies are likely because some stations  
251 within the seismic network were impacted by noise, as discussed in the next section. The data in Table  
252 1 show that estimates of the hypocenter depth involve higher errors compared to those of the epicenter.  
253 This is probably because the learning dataset for the epicenter was larger than that for the hypocenter  
254 depth in the learning stage. Therefore, all possible parameter ranges must be included in the learning

255 stage for the accurate determination of hypocentral parameters, especially for shallow earthquakes as  
256 in the case studied. This can be achieved through enlargement of the training dataset by computing  
257 theoretical seismograms for all possible hypocentral parameters. In addition, Figure 7 reveals slight  
258 differences in the magnitudes. This is because the magnitude estimated using our method is the  
259 moment magnitude, which is based on the scalar moment  $M_0$  of the earthquake, whereas the  
260 magnitude reported by the Hot Spring Research Institute is the micro earthquake magnitude, which is  
261 associated with the maximum amplitude and the hypocentral distance of the seismic wave (Watanabe,  
262 1971). Therefore, the scale of the differences in magnitude between the two sets of results accounts  
263 for the shift. The results also suggest that hypocentral parameters can be estimated using images of  
264 seismic wave propagation evolution without considering the arrival times of the seismic waves at  
265 observatories. Comparison of the estimated errors between the observation and simulation data are  
266 shown in Figure 8, including the regional dependence. Apparently, the accuracies of the hypocentral  
267 parameters exhibit no dependence.

268

#### 269 **Effect of noise**

270 The proposed method was generalized to estimate hypocentral parameters using actual seismograms.  
271 Considering that actual seismograms involve different types of noise, examining their effects on the  
272 estimates is necessary. During the generalization of our method for actual seismograms, some stations  
273 were noted to be associated with higher noise than others. The noise linked to such stations may  
274 adversely affect the hypocentral parameters estimation and account for the significant differences  
275 between our results and those reported by the Hot Spring Research Institute (Figure 9). To address this  
276 problem, we introduced cutout (DeVries and Taylor, 2017), and created the neural network by  
277 randomly selecting at most 3 stations and setting their amplitudes as 0. This procedure is also  
278 applicable for temporarily shut stations with no results. The hypocentral parameters were then  
279 estimated using the network for the test dataset and the observed data, and the results are presented in  
280 Table 1. The results show that the hypocentral parameters estimates improved slightly. In general,  
281 adding noise as a data augmentation process during the learning process may improve the network  
282 performance. Therefore, including physical processes causing noise, such as the scatter associated  
283 with the heterogeneous structure of the crust, or the ambiguity of earthquake source parameters by  
284 using seismic wave propagation simulation to improve the generalization technique may be necessary,  
285 and this can be investigated in future.

286

#### 287 **Summary**

288 In the present study, we established neural networks for determining hypocentral parameters such as  
289 the epicenter, depth, occurrence time, and magnitude using synthetic seismograms as a training dataset.  
290 The theoretical seismograms were synthesized based on a realistic 3D Earth model, and the

291 seismograms were exploited to produce spatial images for seismic wave propagation at the Earth's  
292 surface. The generalized neural network results for obtaining actual seismic data indicated potential  
293 for estimating earthquake hypocentral parameters using CNN trained via theoretical seismograms.  
294 Although more events are required to validate generalizing the proposed neural network, this  
295 framework is suitable for determining hypocentral parameters in a local seismic network, and has  
296 potential for seismic and volcanic activity monitoring.

297

## 298 **Declarations**

299 **Ethics approval and consent to participate:** Not applicable.

300 **Consent for publication:** Not applicable.

## 301 **Availability of data and materials**

302 The seismic waveforms utilized were provided by the Hot Spring Research Institute of the Kanagawa  
303 Prefecture, the National Research Institute for Earth Science and Disaster Resilience, and Japan  
304 Meteorological Agency. The open-source program package SPECSEM3D from the Computational  
305 Infrastructure for Geodynamics (CIG; geodynamics.org) was employed for computation, and the GMT  
306 software [Wessel and Smith, 1998] was used to generate many figures presented.

307 **Competing interests:** There is no competing interests.

308 **Funding:** This work was partially supported by the JAMSTEC Innovation Award 2018 and  
309 KAKENHI 19K12011.

310 **Authors' contributions:** All authors contributed equally to this research. Daisuke Sugiyama:  
311 Conceptualization, Methodology, Data curation and visualization, Writing; Seiji Tsuboi: Methodology,  
312 Investigation, Software, Writing and editing; Yohei Yulutake; Methodology, Data curation and  
313 visualization, Writing

## 314 **Acknowledgments**

315 The theoretical seismograms were created using the Earth Simulator and supercomputers in  
316 JAMSTEC, Japan.

317

## 318 **References**

319 Carreira, J. and Zisserman, A., (2017) Quo Vadis, Action Recognition? A New Model and the Kinetics  
320 Dataset, 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 4724-  
321 4733.

322 Carrington, L., Komatitsch, D., Laurenzano, M. et al., (2008) High-frequency simulations of global  
323 seismic wave propagation using SPECSEM3D GLOBE on 62K processors, In Proc.  
324 Supercomputing 2008 Conf., DOI: 10.1109/SC.2008.5215501.

325 DeVries, T and G. W. Taylor, (2017) Improved Regularization of Convolutional Neural Networks with  
326 Cutout, <https://arxiv.org/abs/1708.04552v2>.

- 327 Donahue, J., Hendricks, L. A., Guadarrama S., Rohrbach, M., Venugopalan, S., Saenko, K., Darrell,  
328 T., (2015) Long-Term Recurrent Convolutional Networks for Visual Recognition and Description,  
329 The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2625-2634
- 330 Du Tran, Bourdev, L., Fergus, R., Torresani, L., Paluri, M. (2015) Learning Spatiotemporal Features  
331 With 3D Convolutional Networks, The IEEE International Conference on Computer Vision  
332 (ICCV), 4489-4497
- 333 Faccioli, E., Maggio, F., Paolucci, R., et al., (1997) 2D and 3D elastic wave propagation by a pseudo-  
334 spectral domain decomposition method, *J. Seismol.* 1, 237-251.
- 335 Ioffe, S., Szegedy, C., (2015) Batch Normalization: Accelerating Deep Network Training by Reducing  
336 Internal Covariate Shift. In: *Proceedings of the 32nd International Conference on Machine  
337 Learning*, Lille Grand Palais, Lille, 6-11 July 2015
- 338 Käüfl, P., Valentine, A. P., O'Toole, T. B., et al., (2014) A framework for fast probabilistic centroid-  
339 moment-tensor determination–inversion of regional static displacement measurements, *Geophys.*  
340 *J. Int.* 196, 1676-1693.
- 341 Käüfl, P., Valentine, A. P., de Wit, R., et al., (2015) Robust and fast probabilistic source parameter  
342 estimation from near-field displacement waveforms using pattern recognition, *Bull. Seismol. Soc.*  
343 *Am.* 105, 2299-2312.
- 344 Käüfl, P., Valentine, A. P., de Wit, R., et al., (2016a), Solving probabilistic inverse problems rapidly  
345 with prior samples, *Geophys. J. Int.* 205, 1710-1728
- 346 Käüfl, P., Valentine, A. P. and Trampert, J., (2016b), Probabilistic point source inversion of strong-  
347 motion data in 3-D media using pattern recognition: A case study for the 2008 Mw 5.4 Chino  
348 Hills earthquake, *Geophys. Res. Lett.* 43, 8492-8498, DOI: 10.1002/2016GL069887
- 349 Kingma, D. P., Ba, J. (2015) Adam: A Method for Stochastic Optimization. In: *Proceedings of the 3rd  
350 International Conference on Learning Representations (ICLR2015)*, Hilton San Diego Resort &  
351 Spa, San Diego, 7-9 May 2015
- 352 Komatitsch, D., (1997) Spectral and spectral-element methods for the 2D and 3D elastodynamics  
353 equations in heterogeneous media, Ph.D. thesis, Institut de Physique du Globe, Paris, France.
- 354 Komatitsch, D. and Vilotte, J. P., (1998) The spectral element method: an efficient tool to simulate the  
355 seismic response of 2D and 3D geological structures, *Bull. Seismol. Soc. Am.* 88, 368-392.
- 356 Komatitsch, D. and Tromp, J., (2002a) Spectral-element simulations of global seismic wave  
357 propagation–I. Validation, *Geophys. J. Int.* 149, 390-412.
- 358 Komatitsch, D. and Tromp, J., (2002b) Spectral-element simulations of global seismic wave  
359 propagation–II. Three-dimensional models, oceans, rotation, and self-gravitation, *Geophys. J. Int.*  
360 150, 303-318.
- 361 Komatitsch, D., Tsuboi, S., and Tromp, J., (2005) The spectral-element in seismology, in *Seismic*

- 362 Earth: Array analysis of broadband seismograms, *Geophys. Monograph* 157, 205.
- 363 Krizhevsky, A., Sutskever, I., and Hinton, G. E., (2012) Imagenet classification with deep  
364 convolutional neural networks. *Adv. NIPS*, 25, 1097-1105.
- 365 LeCun, Y., Haffner, P., Bottou, L., et al., (1999) Object recognition with gradient-based learning. In:  
366 Shape, Contour and Grouping in Computer Vision. *Lecture Notes in Computer Science*, vol 1681.  
367 Springer, Berlin, Heidelberg.
- 368 Onishi, R., and Sugiyama, D., (2017) Deep convolutional neural network for cloud coverage  
369 estimation from snapshot camera images. *SOLA*, 13, 235-239, DOI: 10.2151/sola.2017-043.
- 370 Perol, T., Gharbi, M., Denolle, M., (2018) Convolutional neural network for earthquake detection and  
371 location. *Sci. Adv.* 4, e1700578.
- 372 Rietmann, M., Messmer, P., Nissen-Meyer, T., et al., (2012) Forward and adjoint simulations of  
373 seismic wave propagation on emerging large-scale GPU architectures, In *Proceedings of the*  
374 *International Conference on High Performance Computing, Networking, Storage and Analysis*  
375 *(SC12)*, IEEE Computer Society Press, Los Alamitos, CA, USA, Article 38.
- 376 Ritsema, J., van Heijst, H. J., Woodhouse, J. H., (1999), Complex shear wave velocity structure imaged  
377 beneath Africa and Iceland, *Science* 286, 1925-1928.
- 378 Ross, Z., E., Meler, M-A., Hauksson, E., Heaton, T. H., (2018) Generalized seismic phase detection  
379 with deep learning, *Bull. Seismol. Soc. Am.* 108, DOI: 10.1785/0120180080
- 380 Shahriari, B., Swersky, K., Wang, Z., et al., (2015) Taking the human out of the loop: A review of  
381 Bayesian optimization. *Proc. IEEE*, 104, 148-175. DOI: 10.1109/JPROC.2015.2494218
- 382 Simonyan, K., and Zisserman A., (2014) Very deep convolutional networks for large-scale image  
383 recognition. *arXiv preprint arXiv: 1409.1556*
- 384 Szegedy, C., Liu W., Jia Y., et al. (2015) Going deeper with convolutions. *Proc. IEEE Conference on*  
385 *Computer Vision and Pattern Recognition (CVPR)*, 1-9
- 386 Tsuboi, S., Komatitsch, D., Ji, C., et al. (2003) Broadband modeling of the 2002 Denali fault  
387 earthquake on the Earth Simulator, *Phys. Earth Planet. Inter.* 139, 305-312.
- 388 Tsuboi, S., Ando, K., Miyoshi, T., et al., (2016) A 1.8 trillion degrees of freedom, 1.24 petaflops global  
389 seismic wave simulation on the K computer. *Int J. High Perform Comput Appl.* 30, 411-422. DOI:  
390 10.1177/1094342016632596
- 391 Vaillant, R., Monrocq C., and Cun, Y. L. (1994) Original approach for the localization of objects in  
392 images. *IEE Proceedings-Vision, Image and Signal Processing*, 141, 245-250
- 393 Wang, J., Teng T.-L., (1995) Artificial neural network-based seismic detector. *Bull. Seismol. Soc. Am.*  
394 85, 308-319.
- 395 Wang, J., Teng, T. L., (1997) Identification and picking of S phase using an artificial neural network.  
396 *Bull. Seismol. Soc. Am.* 87, 1140-1149.
- 397 Wessel, P., Smith, W. H., (1998) New, improved version of the Generic Mapping Tools released. *EOS*

- 398 Trans AGU 79, 579-579. DOI: 10.1029/90EO00426
- 399 Yukutake, Y., R. Honda, M. Harada, R. Arai, and M. Matsubara (2015), A magma-hydrothermal  
400 system beneath Hakone volcano, central Japan, revealed by highly resolved velocity structures,  
401 J. Geophys. Res. Solid Earth, 120, 3293–3308, doi:10.1002/2014JB011856.
- 402 Zhu, W., Beroza, G. C., (2019) PhaseNet: a deep-neural-network-based seismic arrival-time picking  
403 method. Geophys. J. Int, 216, 261-273.

404 Table 1 Root Mean Square Errors (RMSE) of hypocentral parameters estimated by various networks. \*ConvNetQuake (Perol,et al., 2018) is a task to classify  
 405 earthquake location into specific regions. The results show in this table for ConvNetQuake is obtained by added our task to perform regression to locate  
 406 hypocentral parameters during the learning stage.

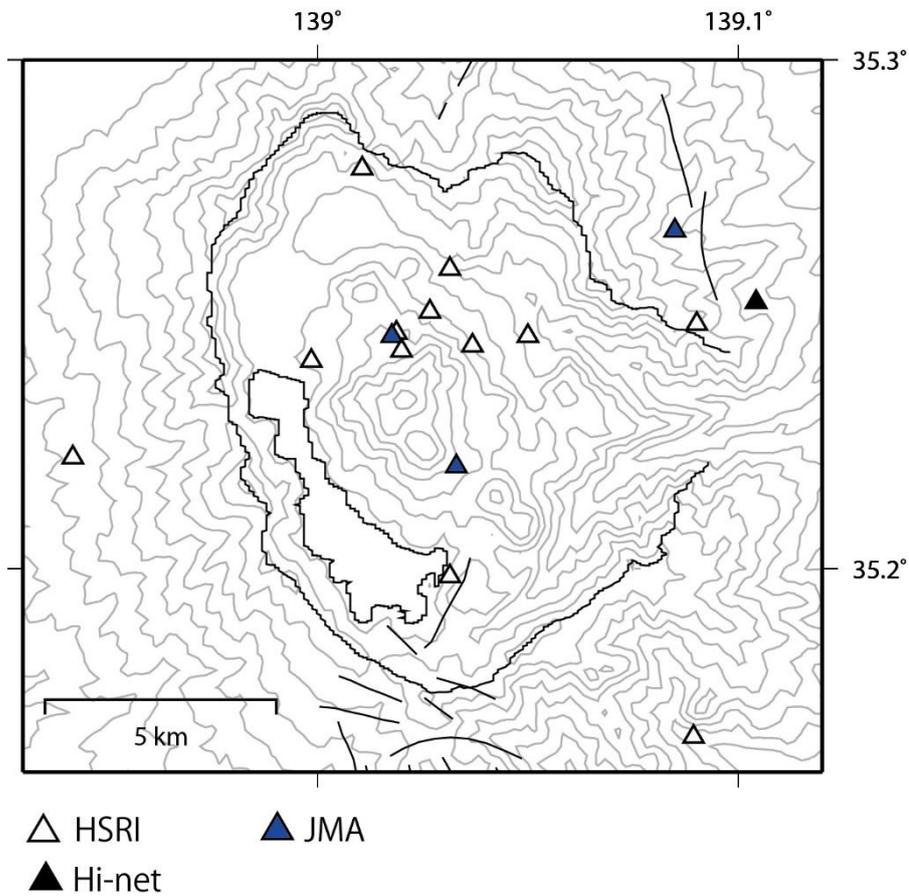
407  
 408  
 409

Method	Training	Estimation	Root Mean Square Error (RMSE)				
			Longitude (degree)	Latitude (degree)	Depth (km)	Time (sec)	Magnitude
1DCNN (ConvNetQuake※)	Simulation data (80%)	Simulation data (20%)	0.0406	0.0365	1.679	-	0.0930
2DCNN	Simulation data (80%)	Simulation data (20%)	0.0213	0.0184	1.3267	0.5002	0.1018
3DCNN	Simulation data (80%)	Simulation data (20%)	0.0100	0.0078	0.6208	0.2095	0.0371
		Observed data	0.0301	0.0291	1.3062	1.3816	0.5830
3DCNN +CUTOUT	Simulation data (80%)	Simulation data (20%)	0.0075	0.0091	0.5298	0.2444	0.0311
		<b>Observed data</b>	<b>0.0167</b>	<b>0.0313</b>	<b>1.7504</b>	<b>0.9427</b>	<b>0.6063</b>

410

411 Figure captions

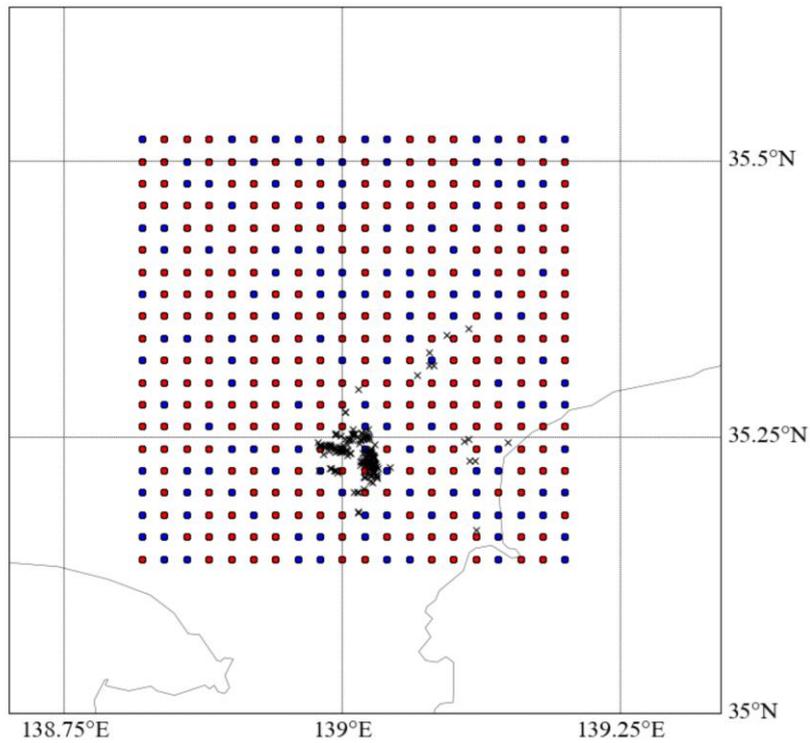
412



413

414 Figure 1. Map of station location o used to create learning dataset. Open triangle is the stations  
 415 operated by Hot Spring Research Institute. Closed trialgles are those operated by JMA and NIED.

416

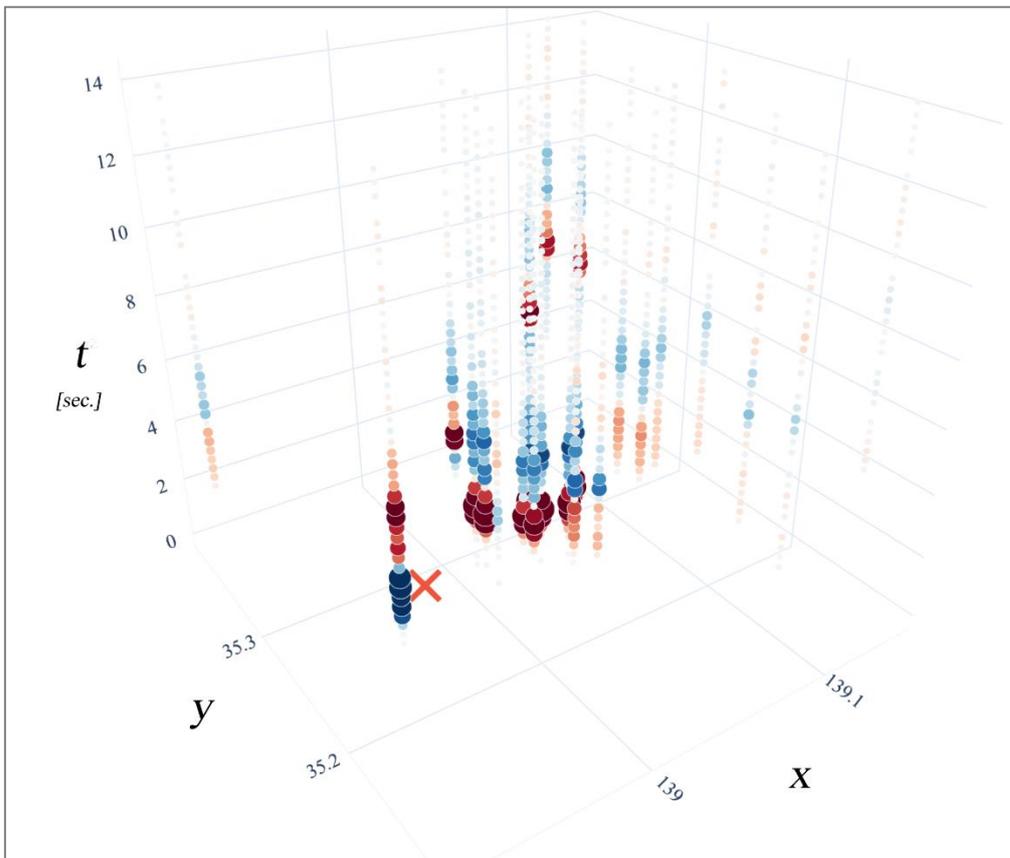


417

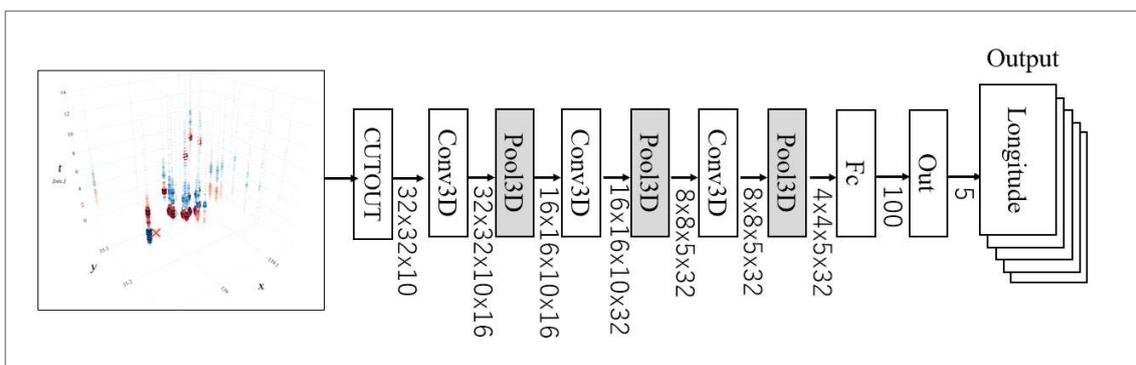
418

419 Figure 2. Red dots are the earthquakes used for learning and blue dots are the earthquakes used for  
420 testing. The size of the dots is proportional to the magnitude. Crosses are epicenters of observed  
421 earthquakes. The total number of earthquakes used is 2632 and we used 2236 events as the training  
422 dataset. Some of the earthquakes are removed from the dataset due to apparent numerical problems  
423 during the simulation.

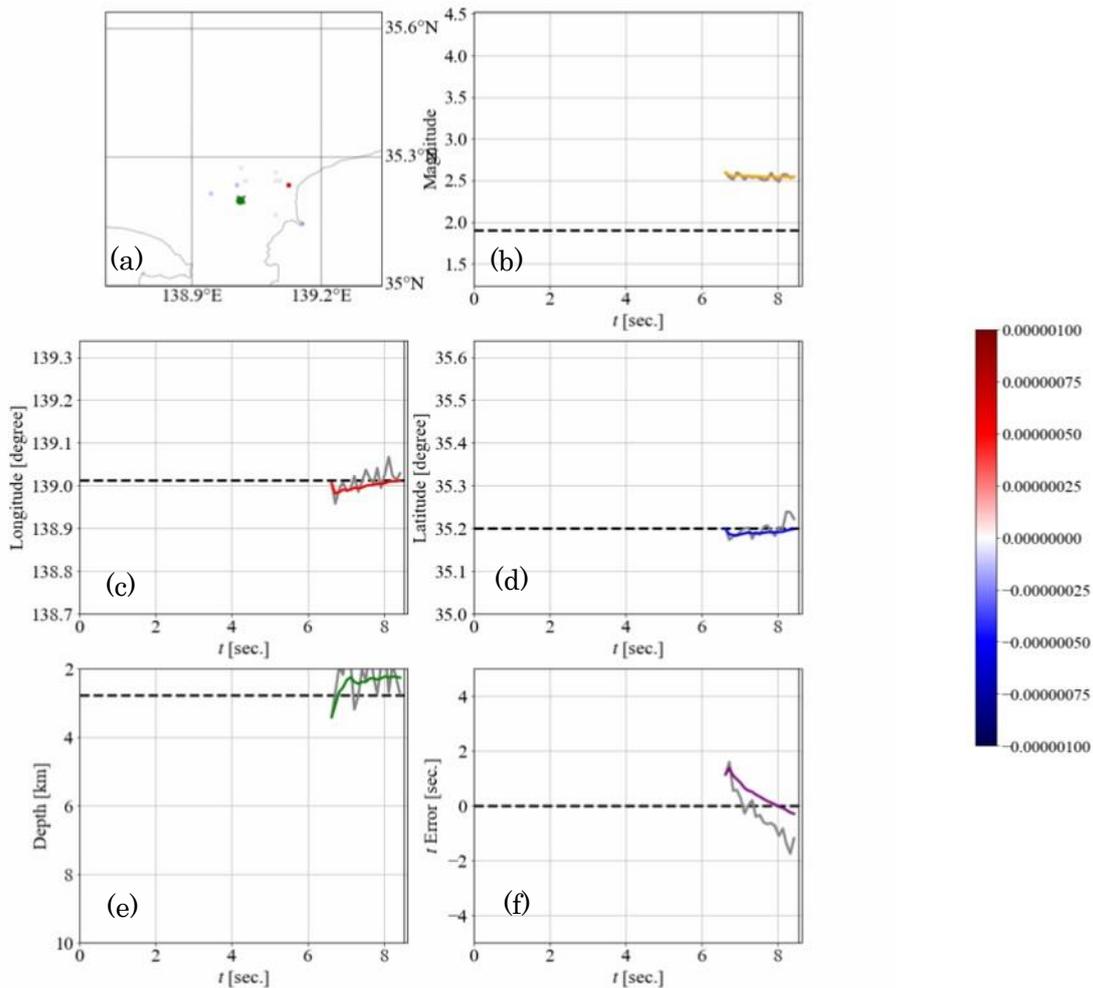
424



425  
 426 Figure 3. Map of example propagation image used for learning process. Colors indicate vertical  
 427 displacement from  $-1.4 \times 10^{-6}$  m to  $1.4 \times 10^{-6}$  m as they change from blue to red.  
 428



429  
 430 Figure 4. Network architecture used for the 3D convolutional neural network.  
 431  
 432



433

434 Figure 5. Examples of hypocenter estimation in the verification phase for one earthquake.

435 (a) Seismic wave propagation images starting at approximately 6 s after the origin time with an interval  
 436 of 1 s from left to right. The scale of the gray color ( $-1.0e-5$  to  $1.0e-5$  m/s) is shown at the left. Below  
 437 the images, the time evolution of the estimated hypocenter parameters is shown. The horizontal axis  
 438 of each panel represents the time in seconds, and the estimation starts at approximately 6 s after the  
 439 origin time when the absolute amplitude exceeds  $1.0e-7$  m/s. The estimation terminates 50 s after the  
 440 origin time, before the surface waves are generated. (b) magnitude, (c) longitude, (d) latitude, (e) depth,  
 441 and (f) origin time. The dotted line in each figure represents the parameter to be estimated by each  
 442 snapshot. The gray line in each figure shows the value estimated using each snapshot. The colored line  
 443 shows the average of the estimated value up to each time instant. It should be noted that in the  
 444 verification of the latitude, shown in (d), the estimated value is in agreement with the actual value, and  
 445 the overlap of the three lines indicates that no differences exist.

446

447



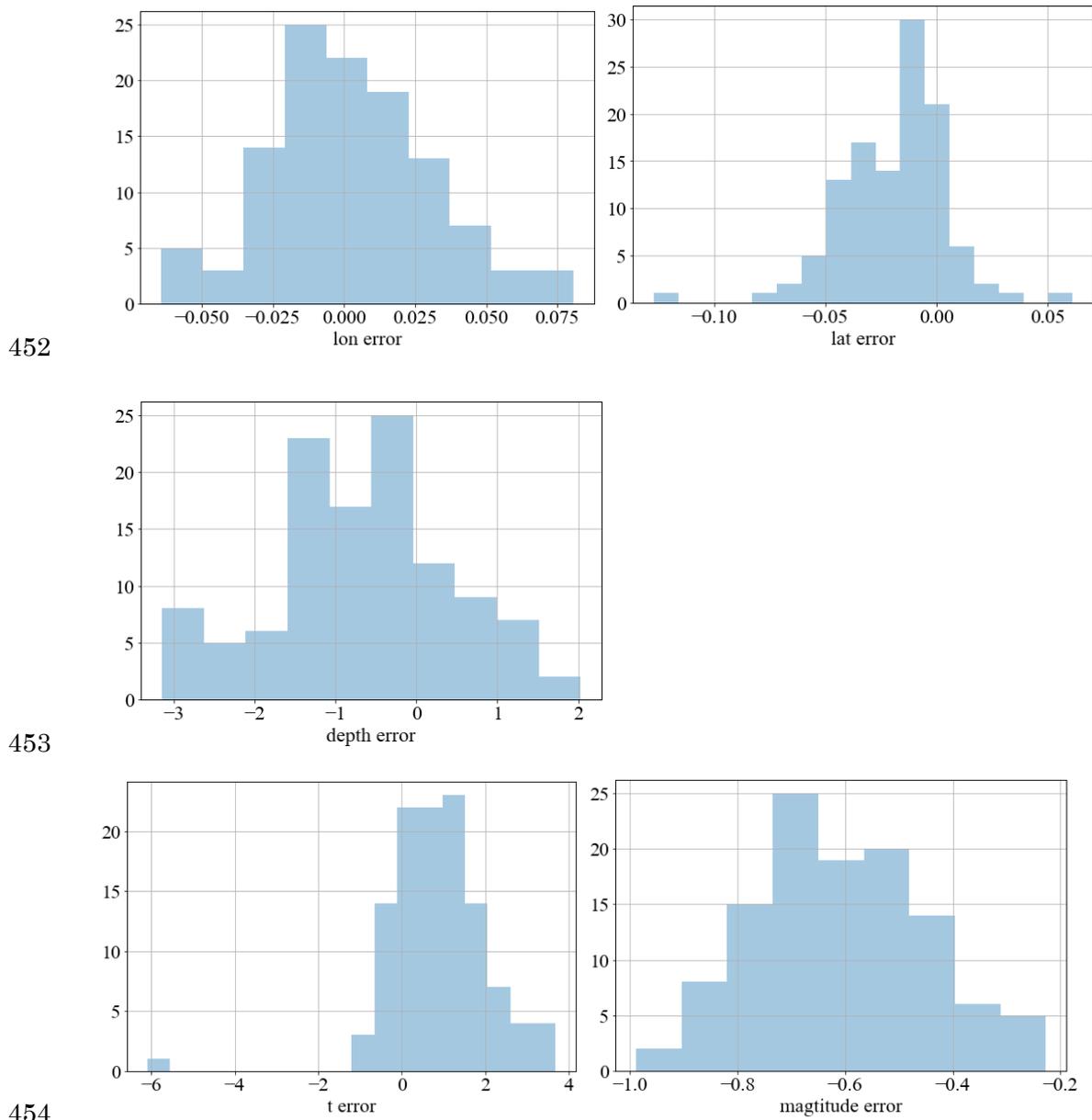
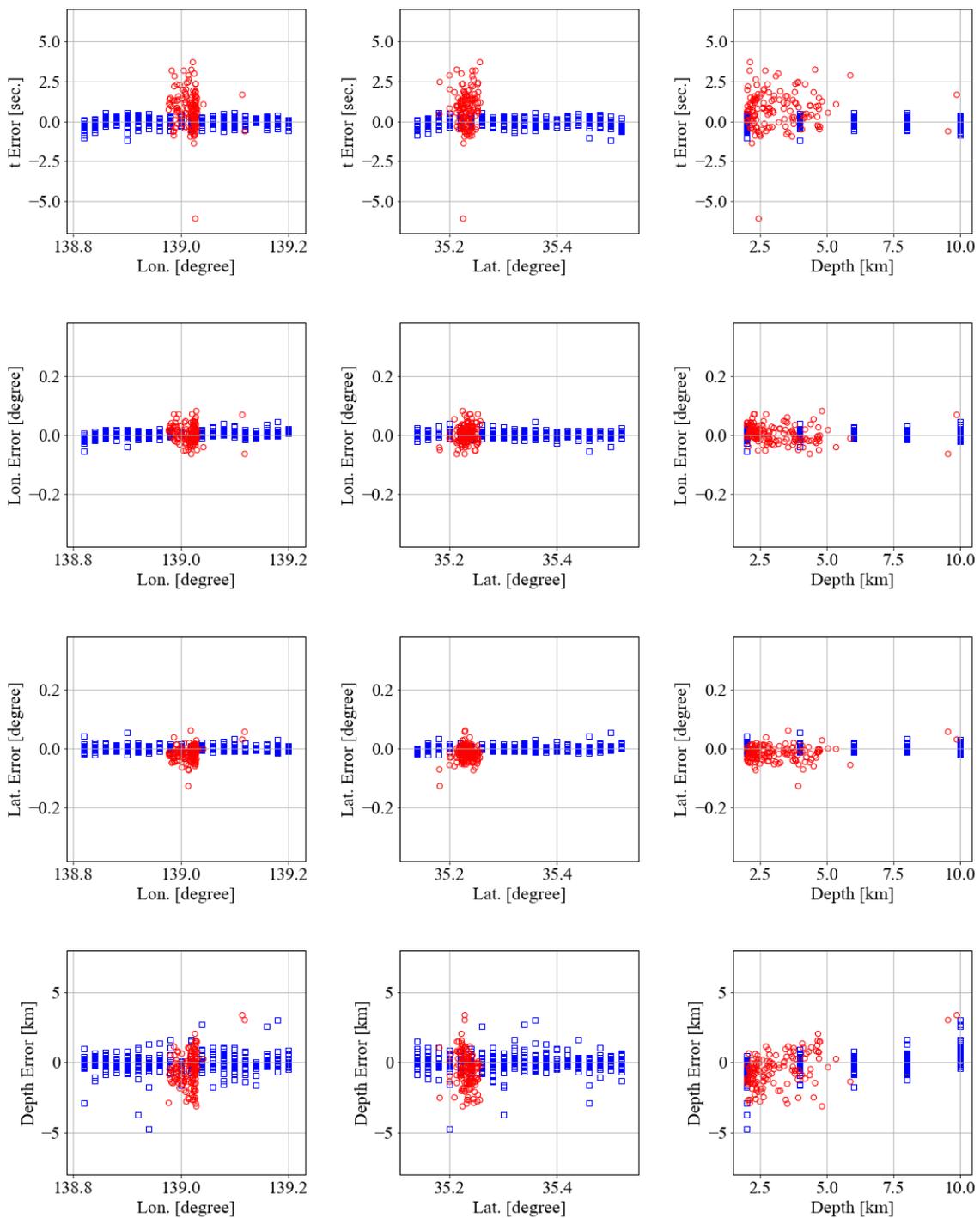


Figure 7. Histograms of errors for hypocentral parameters determined by the proposed technique.



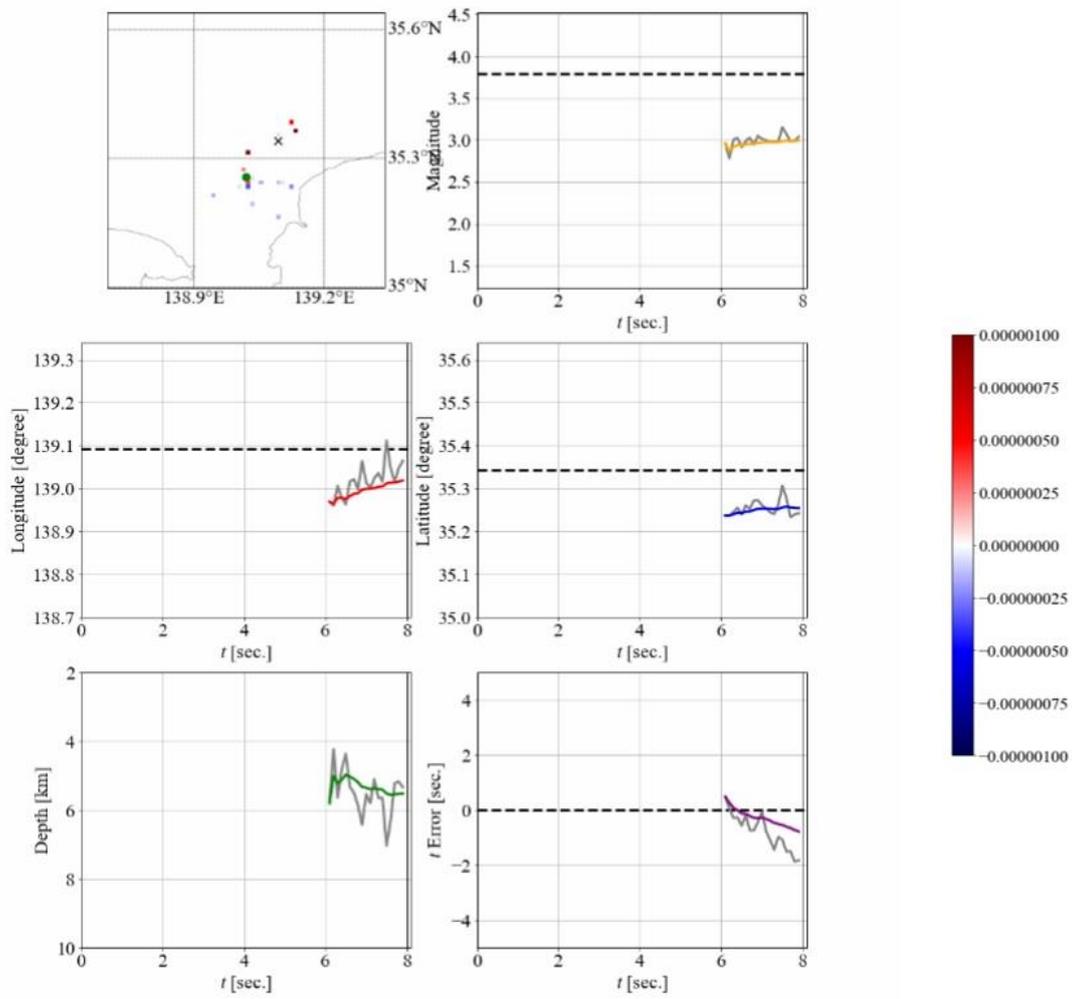
458

459 Figure 8. Distribution of errors for hypocentral parameters determined by the proposed technique.

460 Each column shows geographical distribution of errors for hypocentral parameters. Each row shows

461 errors of origin time, longitude, latitude and depth, from top to bottom. Red symbols show observation

462 and blue symbols show simulation.



463

464 Figure 9. Same as Figure 5 but for noisy case. For details of figures see caption to Figure 5.

# Figures

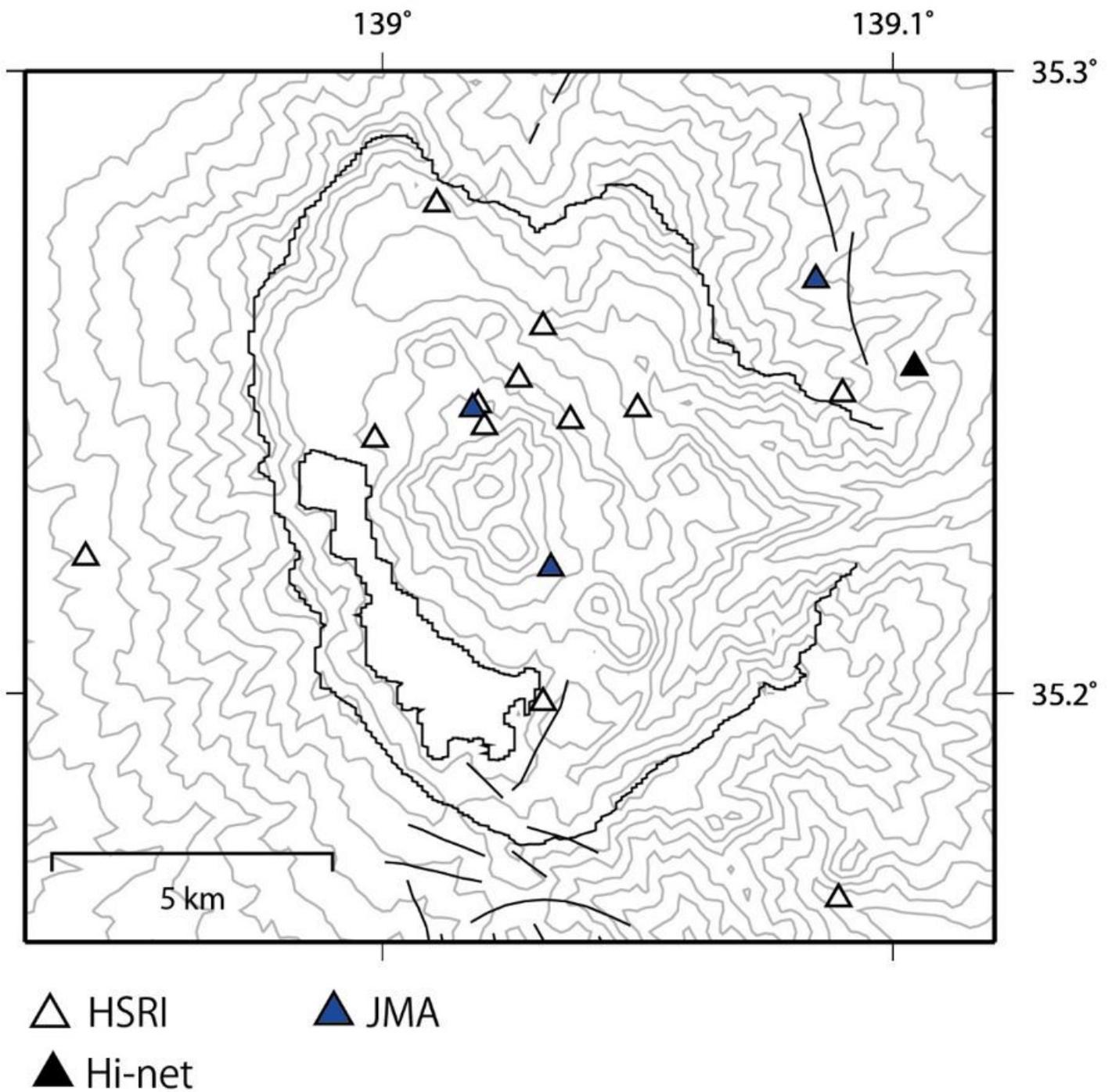
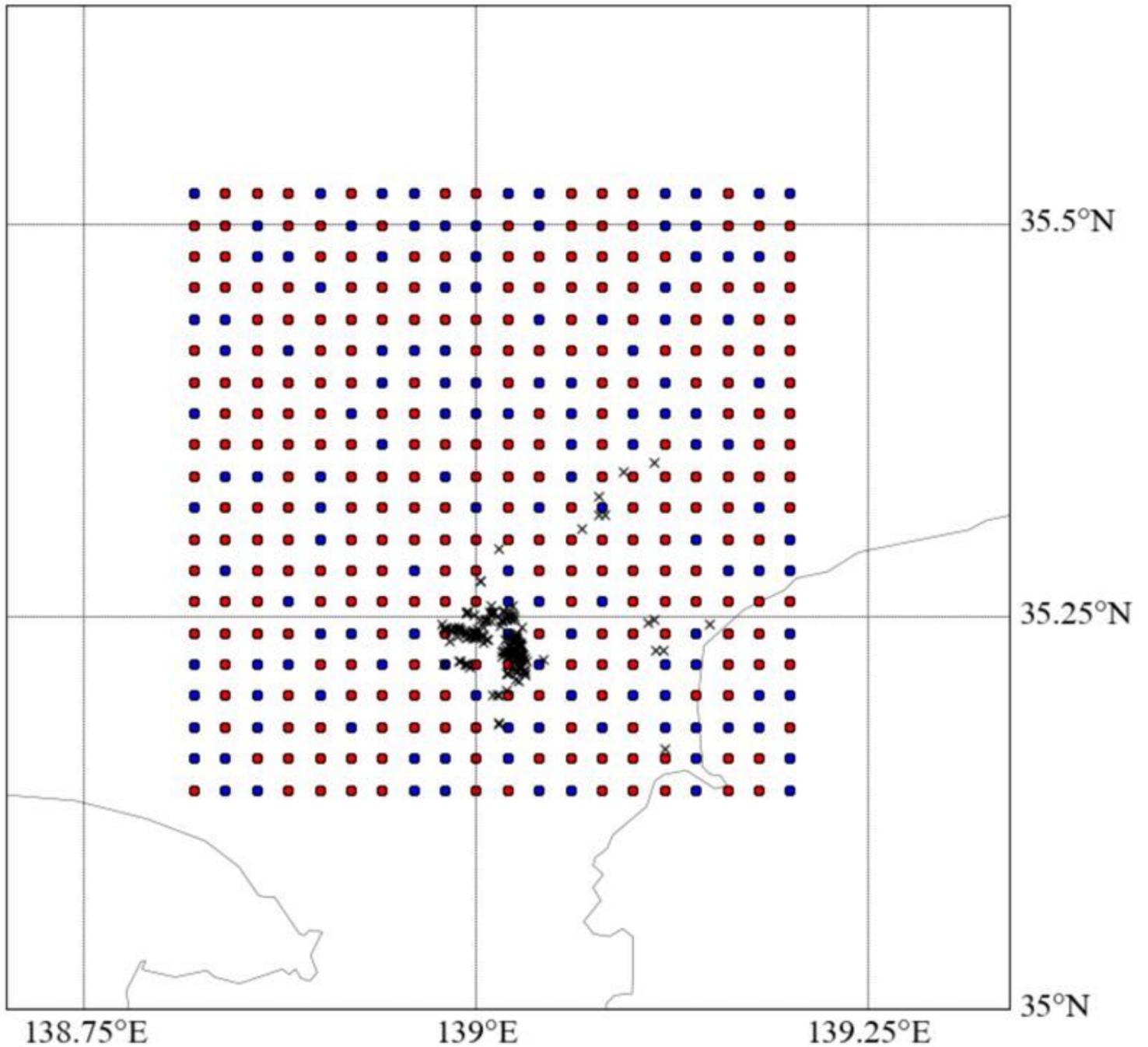


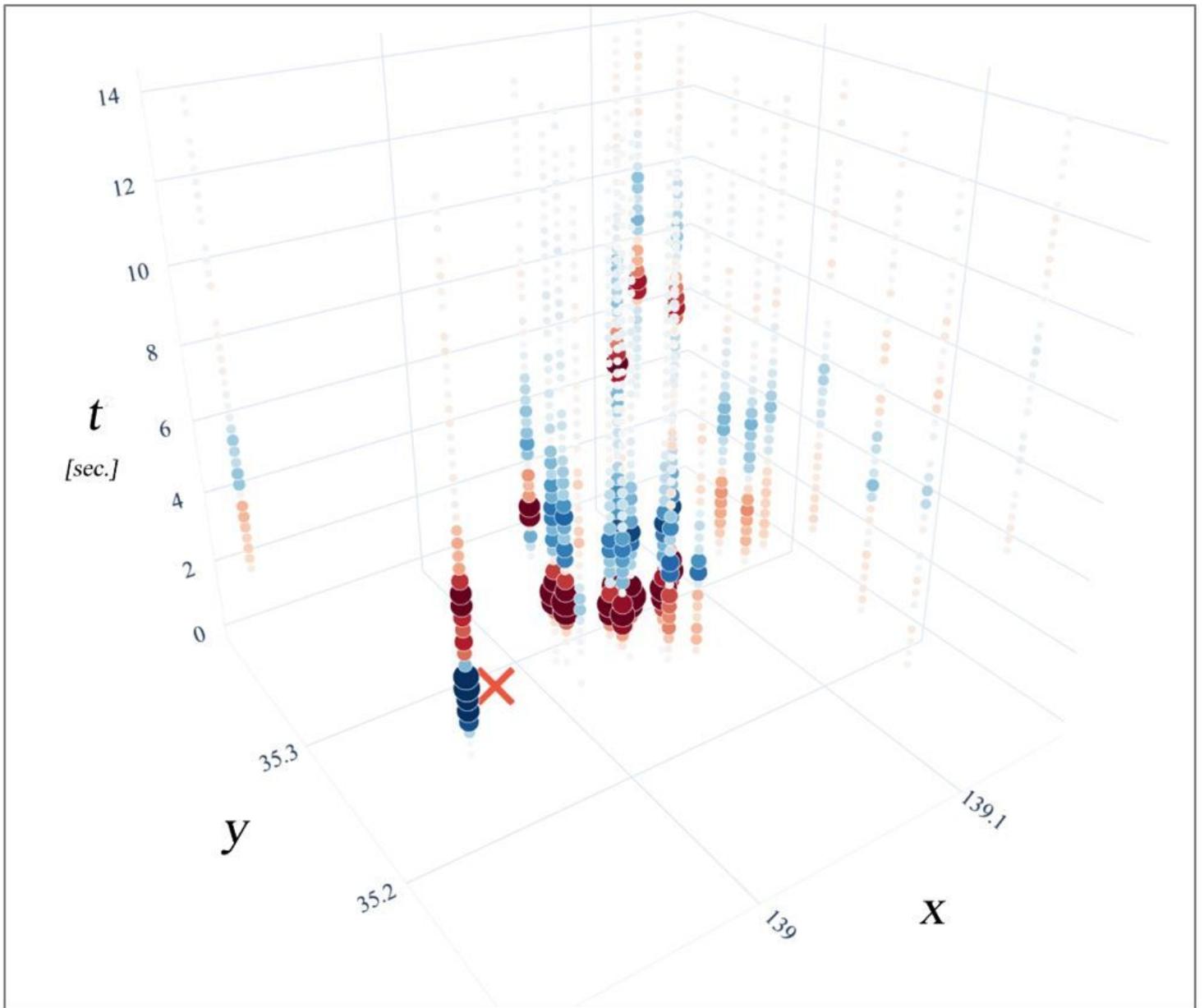
Figure 1

Map of station location o used to create learning dataset. Open triangle is the stations operated by Hot Spring Research Institute. Closed trialgles are those operated by JMA and NIED.



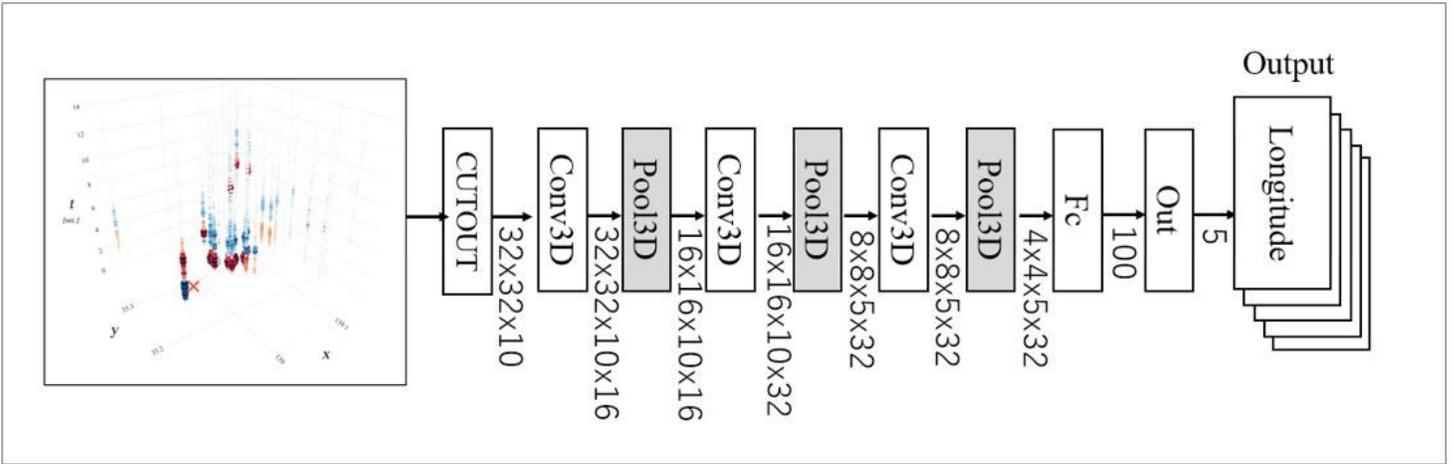
**Figure 2**

Red dots are the earthquakes used for learning and blue dots are the earthquakes used for testing. The size of the dots is proportional to the magnitude. Crosses are epicenters of observed earthquakes. The total number of earthquakes used is 2632 and we used 2236 events as the training dataset. Some of the earthquakes are removed from the dataset due to apparent numerical problems during the simulation.



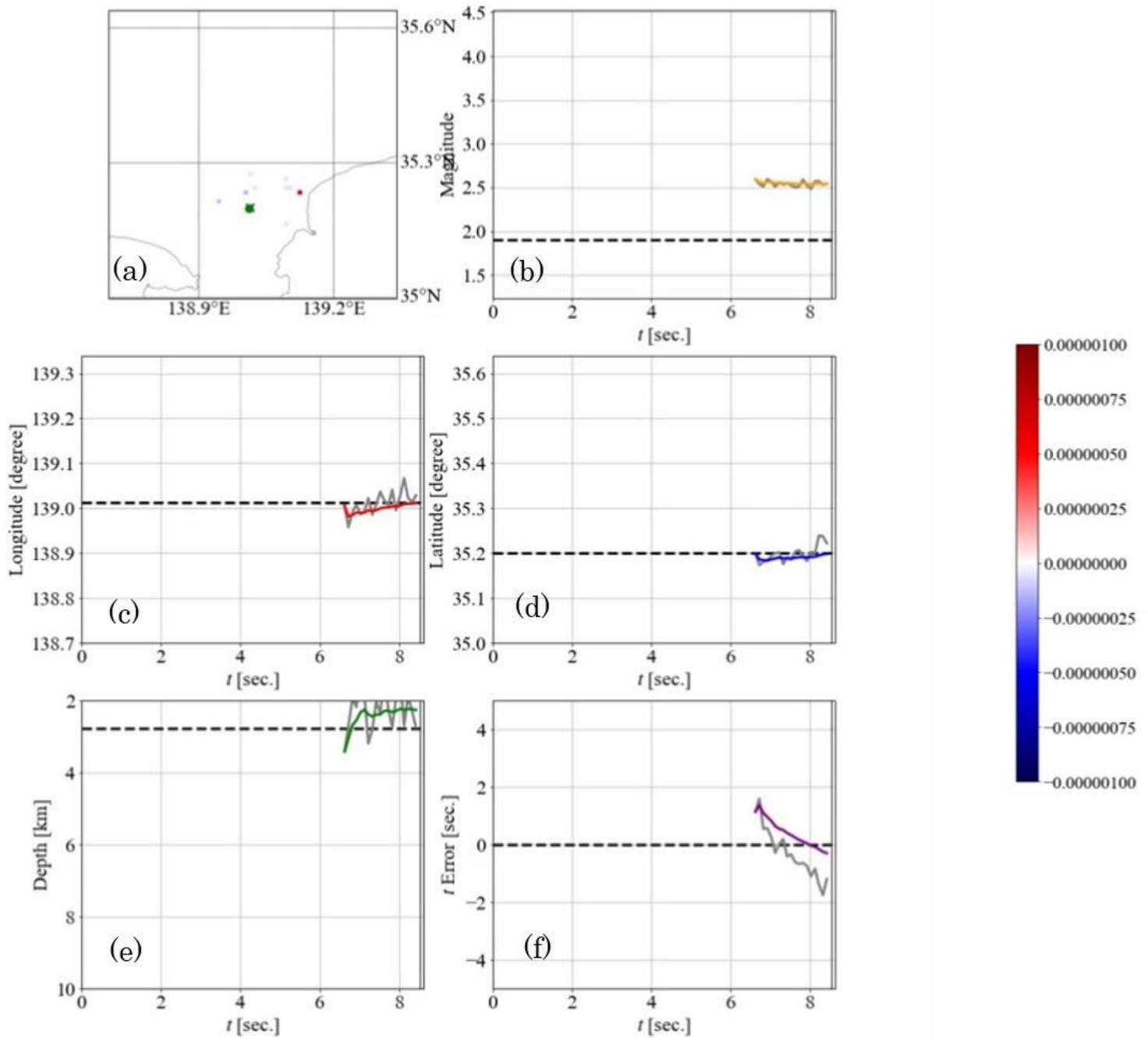
**Figure 3**

Map of example propagation image used for learning process. Colors indicate vertical displacement from  $-1.4 \times 10^{-6}$  m to  $1.4 \times 10^{-6}$  m as they change from blue to red.



**Figure 4**

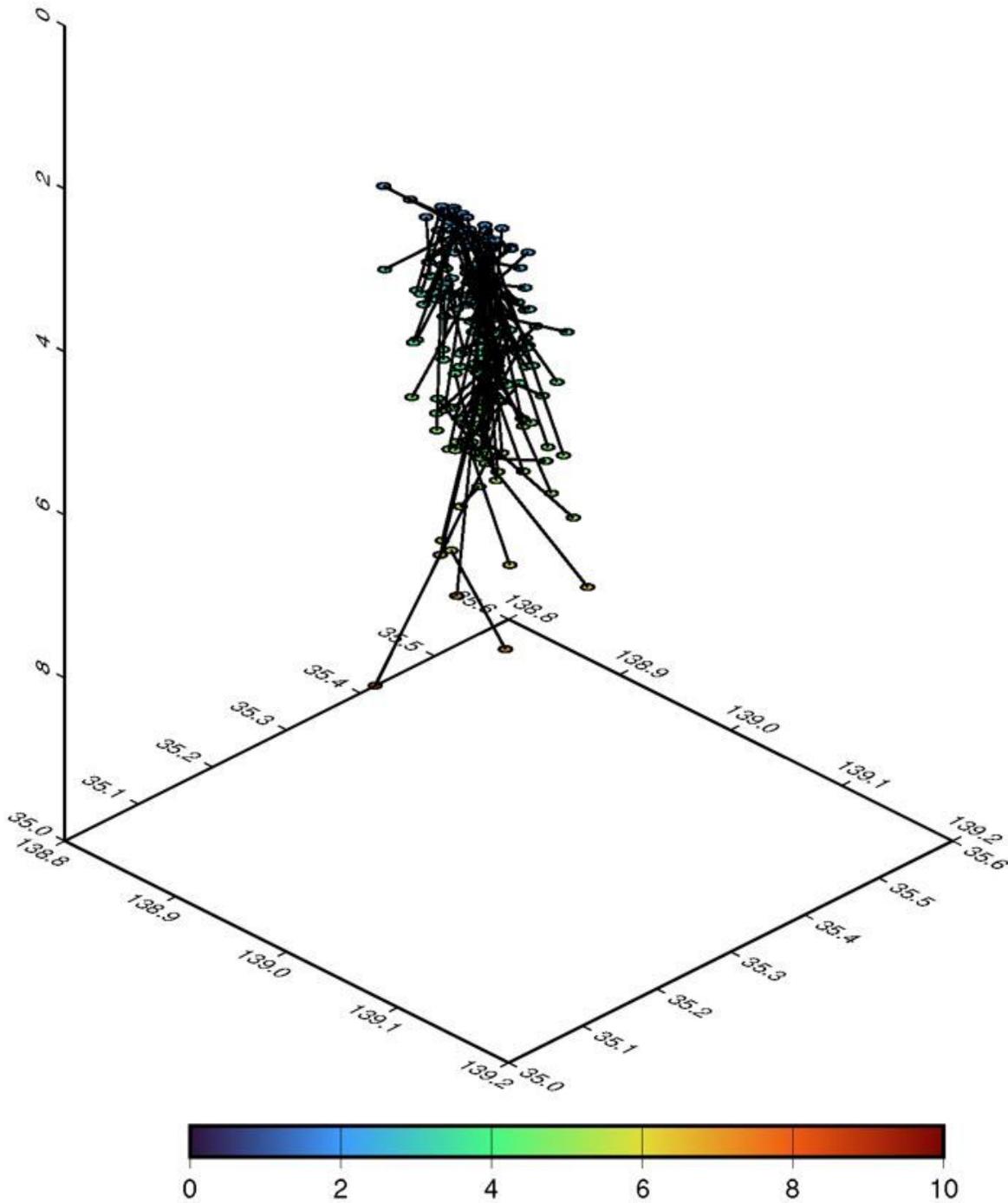
Network architecture used for the 3D convolutional neural network.



**Figure 5**

Examples of hypocenter estimation in the verification phase for one earthquake. (a) Seismic wave propagation images starting at approximately 6 s after the origin time with an interval of 1 s from left to right. The scale of the gray color ( $-1.0e-5$  to  $1.0e-5$  m/s) is shown at the left. Below the images, the time evolution of the estimated hypocenter parameters is shown. The horizontal axis of each panel represents the time in seconds, and the estimation starts at approximately 6 s after the origin time when the absolute amplitude exceeds  $1.0e-7$  m/s. The estimation terminates 50 s after the origin time, before the surface waves are generated. (b) magnitude, (c) longitude, (d) latitude, (e) depth, and (f) origin time. The dotted line in each figure represents the parameter to be estimated by each snapshot. The gray line in each figure shows the value estimated using each snapshot. The colored line shows the average of the

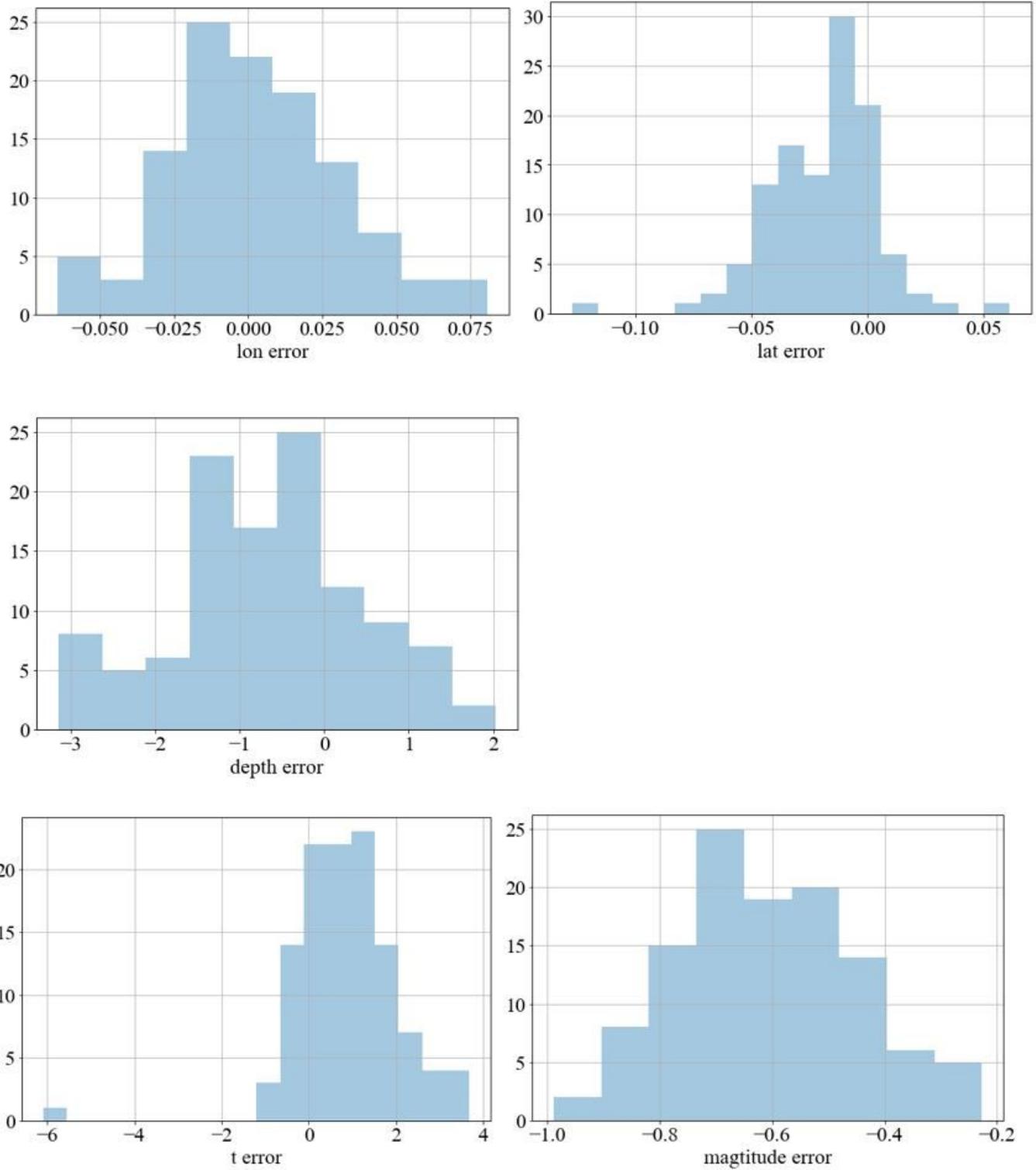
estimated value up to each time instant. It should be noted that in the verification of the latitude, shown in (d), the estimated value is in agreement with the actual value, and the overlap of the three lines indicates that no differences exist.



**Figure 6**

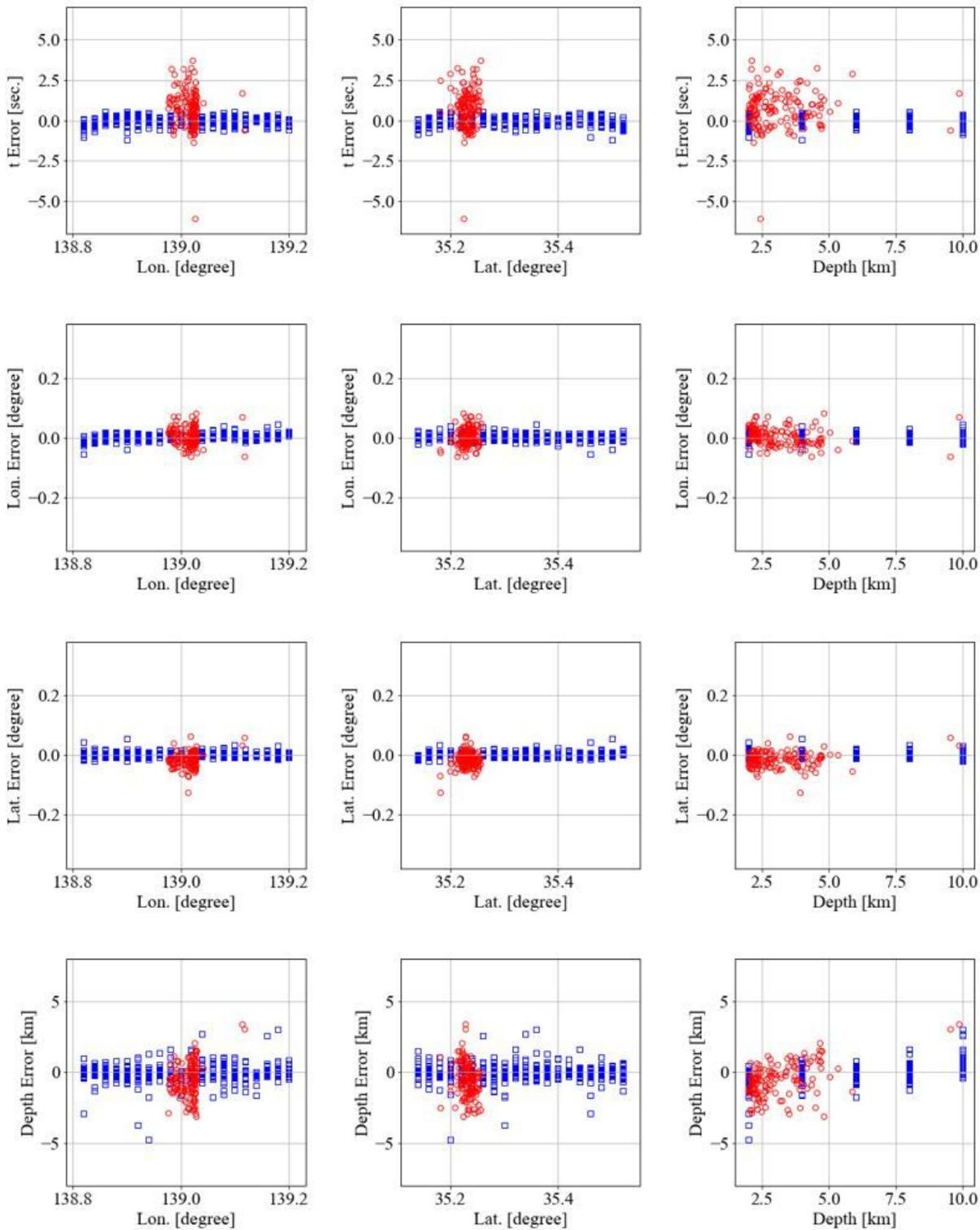
Three dimensional plot for differences of hypocentral parameters between those determined by the Hot Spring Research Institute and determined by the proposed technique. Colors of symbols show

hypocentral depth.



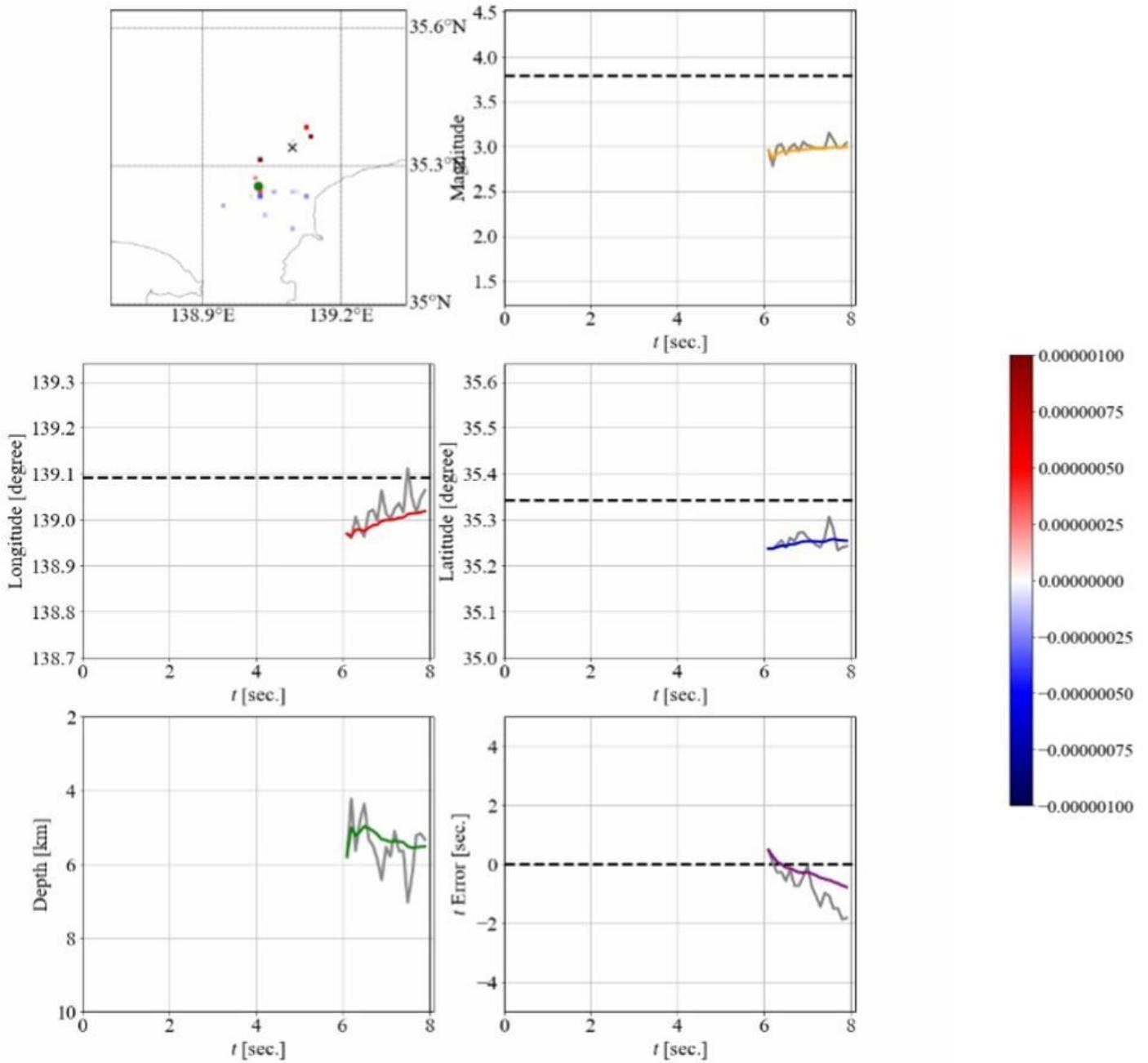
**Figure 7**

Histograms of errors for hypocentral parameters determined by the proposed technique.



**Figure 8**

Distribution of errors for hypocentral parameters determined by the proposed technique. Each column shows geographical distribution of errors for hypocentral parameters. Each row shows errors of origin time, longitude, latitude and depth, from top to bottom. Red symbols show observation and blue symbols show simulation.



**Figure 9**

Same as Figure 5 but for noisy case. For details of figures see caption to Figure 5.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [graphicalAbstract.jpg](#)