MAGMA: Machine learning Automatic picker for Geothermal Microseismicity Analysis for practical procedure to reveal fine reservoir structures

Kyosuke Okamoto¹, Yusuke Mukuhira², Hiroshi Asanuma³, Hirokazu Moriya², and Häring Markus⁴

¹National Institute of Advanced Industrial Science and Technology
²Tohoku University
³National Institute of Advanced Industrial Science and Technology (AIST)
⁴Häring GeoProject

November 21, 2022

Abstract

In geothermal development, microseismic monitoring is an important technique to monitor the various phenomena in the reservoirs throughout location, activity, and magnitude of microseismicity. Picking P- and S-wave arrivals accurately from seismic data is an inevitable process for subsequent seismic analyses. However, the manual phase picking is a time- and cost-consuming process and several automatic pickers still require considerable quality checks and corrections by human analysts. Automatic pickers based on deep learning have recently been developed for natural earthquake analysis, whose accuracy has been confirmed to be comparable to that of human analysts. These phase pickers were mainly trained using natural earthquakes recorded by regional seismic networks. However, seismic networks and events in geothermal fields have features that differ from those of natural earthquakes. In such fields, seismic events with very low magnitudes occur immediately under the seismic network and are sometimes triggered by fluid activity. Therefore, the direct application of the existing deep learning phase pickers to such seismic networks may have difficulty. Here, we focus on developing a deep learning model specialized for local seismic networks in geothermal fields. We used microseismic data from four representative enhanced geothermal and hydrothermal fields and trained the model with deep learning. Based on the developed model, the hypocenter distribution was determined using continuous seismic waves in the Okuaizu geothermal field, Japan. These procedures were performed automatically without manual operations and we propose them as MAGMA: Machine learning Automatic picker for Geothermal Microseismicity Analysis. Subsurface fine structures were then revealed by relocating the hypocenters using a double-difference algorithm. The same procedures for the same data were then conducted using a deep learning model that was trained by other field data, and the equivalent structures were successfully revealed. Thus, MAGMA is applicable to new fields even when data is lacking, such as green fields.
ABSTRACT

In geothermal development, microseismic monitoring is an important technique to monitor the various phenomena in the reservoirs throughout location, activity, and magnitude of microseismicity. Picking P- and S-wave arrivals accurately from seismic data is an inevitable process for subsequent seismic analyses. However, the manual phase picking is a time- and cost-consuming process and several automatic pickers still require considerable quality checks and corrections by human analysts. Automatic pickers based on deep learning have recently been developed for natural earthquake analysis, whose accuracy has been confirmed to be comparable to that of human analysts. These phase pickers were mainly trained using natural earthquakes recorded by regional seismic networks.

However, seismic networks and events in geothermal fields have features that differ from those of natural earthquakes. In such fields, seismic events with very low magnitudes occur immediately under the seismic network and are sometimes triggered by fluid activity. Therefore, the direct application of the existing deep learning phase pickers to such seismic networks may have difficulty. Here, we focus on developing a deep learning model specialized for local seismic networks in geothermal fields. We used microseismic data from four representative enhanced geothermal and hydrothermal fields and trained the model with deep learning. Based on the developed model, the hypocenter distribution was determined using continuous seismic waves in the Okuaizu geothermal field, Japan. These procedures were performed automatically without manual operations and we propose them as MAGMA: Machine learning Automatic picker for Geothermal Microseismicity Analysis. Subsurface fine structures were then revealed by relocating the hypocenters using a double-difference algorithm. The same procedures for the
same data were then conducted using a deep learning model that trained by other field data, and the equivalent structures were successfully revealed. Thus, MAGMA is applicable to new fields even when data is lacking, such as green fields.
INTRODUCTION

Microseismicity during the development in both hydrothermal system and Enhanced Geothermal Systems (EGS) is one of the important phenomena when trying to monitoring and understanding the changing subsurface conditions associated with fluid extraction and injection. As a result, systems that monitor microseismic activity are commonly deployed in geothermal areas under development. Microseismicity is the most practical geophysical monitoring method which provides information of various subsurface phenomena in high space resolution in real time. In this context, microseismicity generally occur from shear slip on existing fractures in geothermal reservoir. The location of microseismicity generally provide the information of existing fractures which are flow paths of geothermal fluids. Shear slip on existing fractures is induced by stress changes which are caused by human activities such as fluid injection, and production, as well as natural change in reservoir characteristics. Microseismicity can track such a time variant change of the reservoir (e.g., Kwiatek et al., 2015; Martínez-Garzón et al., 2014; Okamoto et al., 2018; 2020) and contribute long term management of geothermal resources. In the case of EGS, microseismic monitoring plays very important roles to provide the extension of stimulated zone and pore pressure migration (Mukuhira et al., 2017; Shapiro et al., 2002). Furthermore, the spatiotemporal transition of microseismic events can shed light on fluid flow paths (Asanuma et a., 2008; Moriya et al., 2003; Evans et al.,2005; Mukuhira et al., 2020), and their extent can indicate active reservoir areas. Microseismicity also can be utilized as a tools to assess the problems of injection induced seismicity (Kwiatek et al., 2014; Kwiatek et al., 2019; McGarr, 2014; Shapiro et al., 2010; van der Elst et al., 2016).
All of these insights from microseismicity relies on the quality of analysis of microseismicity. Picking seismic phases (P- and S-wave arrivals) after event detection from continuous seismic waves at multiple seismic stations is the first step in determining the location of microseismic events, providing fundamental information for further microseismic analyses. For a quick and accurate evaluation of the changing subsurface conditions, picking the seismic phases in near real-time with high accuracy is indispensable. However, in exchange for a high accuracy, picking is a high-load task in terms of both time and cost when performed by human analysts. Thus, automatic phase pickers are widely used to reduce human resources, time, and costs.

A commonly used method for automatically finding P-and S-wave arrivals is the short-time average to long-time average (STA/LTA) trigger (Allen, 1978; Withers et al., 1998). This trigger identifies phase arrivals by abrupt changes in amplitude as a function of time using a pair of short and long time windows sliding over continuous seismic waves. Pickers based on autoregression-Akaike’s information criterion (AR-AIC) are also commonly used (Sleeman and van Eck, 1999). However, the accuracy of these automatic pickers is lower than that of manual picking by human analysts, and the data requires subsequent manual correction to enhance its accuracy for practical use. Generally, microseismic monitoring systems employ a higher sampling rate (200 Hz or higher than 1,000 Hz) than regional seismic networks that target natural earthquakes. As a result, the manual correction of the picked phases is also a high-load task. Template matching, a more modern microseismic picking method, is widely used in natural earthquake and microseismicity analysis too (e.g., Gibbons and Ringdal, 2006; Skoumal et al., 2014; Huang and Beroza, 2015). Template matching method refers the similarity of waveforms by signal processing and detects precise P- and S-wave arrivals in comparison with template
waveforms. Obviously, these methods require large amount data accumulation of well picked seismograms and would not realistically work in real time, especially for green field.

Recently, various types of automatic pickers have been developed using deep learning, some of which have focused on picking phases in seismic waves that have been triggered beforehand. For example, P-wave arrivals (and first-motion polarities) are determined from 4-s long pre-triggered waves (Ross et al., 2018a). Alternatively, PhaseNet (Zhu and Beroza, 2019) treats longer seismic waves (30 s), from which P- and S-wave arrivals were picked, and can recognize the presence of noise. Meanwhile, there are multi-level pickers based on deep learning that can detect earthquake signals from continuous seismic waves, from which the precise locations of P- and S-wave arrivals are picked (e.g., EQTransformer (Mousavi et al., 2020) and GPD (Ross et al., 2018b)). The accuracy of these deep learning pickers approaches that of manual picking, and can at times read the phases more accurately than human analysts (Zhu and Beroza, 2019). Thus, pickers developed using deep learning have made significant progress for use as automatic phase pickers.

The existing deep learning pickers mentioned above were trained mainly using natural earthquakes recorded by regional seismic networks (epicentral distances are within several hundred kilometers). All seismic data used were sampled at (or resampled to) 100 Hz, which is a common sampling rate used in such regional networks. However, in this study, the microseismic monitoring systems that we focused on targeted local microseismic events occurring immediately under them. The epicentral distances ranged from almost zero to several kilometers at most, resulting in extremely short time intervals between P- and S-wave arrivals. Focusing on detecting the location of microseismic events with a high accuracy, the sampling rate was
sometimes higher than 100 Hz, and a sampling rate of 1,000 Hz or much higher is commonly used. The conditions mentioned above differ from those used in existing deep learning pickers. Moreover, the signal-to-noise ratio of microseismic events is relatively low because of their small moment magnitude, and the dominant frequency range is relatively higher compared to the natural earthquakes employed for training existing deep learning pickers. Many microseismic events in natural resource fields are related to subsurface fluid activities, which add non-double coupled components to faulting, unlike ordinary natural earthquakes. For these reasons, the direct application of existing deep learning pickers to microseismic monitoring systems may be difficult. Chai et al. (2020) applied PhaseNet, an existing deep-learning picker, to microseismic events recorded in a geothermal field during hydraulic fracturing. They found that the direct use of PhaseNet for microseismic events resulted in inadequate accuracy compared to human analysts, although the results were much better than those obtained by an automatic picker based on the STA/LTA and AR-AIC methods (Akazawa, 2004). They then conducted transfer learning of PhaseNet using local microseismic data, which improved the accuracy significantly. This is an example of how the precision of phase-picking approaches or exceeds that of human analysts. The learning of local microseismic data is key to continue to make improvements in accuracy.

In this study, we established a deep learning model specialized for locally distributed microseismic monitoring systems in geothermal development to pick P- and S-wave arrivals. We focused on creating a procedure to reveal fine structures from continuous seismic waves to elucidate subsurface fluid flow paths and reservoir structures. For this purpose, we employed microseismic data recorded from four different geothermal fields. First, we trained a deep learning model using microseismic data from each field independently (solo model). Subsequently, we conducted additional training using microseismic data from the remaining
field(s) (fine-tuned model). We examined the picking accuracy using solo models and fine-tuning models on one particular field data. We also examined the applicability of phase picking using a fine-tuning model to new field. For practical use, we determined the hypocenter locations of the Okuaizu geothermal field from continuous seismic waves using the self-solo model and the fine-tuning model trained with other field data, and then revealed fine structures based on a double-difference (DD) algorithm (Waldhauser and Ellsworth, 2000). Significant progress will be made in microseismic monitoring systems in geothermal development if hypocenter determination and fine imaging of subsurface structures (e.g., flow paths and fracture networks) are automatically realized based on deep learning. Thus, the success of the application of the deep learning model that trained by other field data is important, as the deep learning model has the potential to be applied even if the target field does not have its own data, such as in green fields.

TARGET GEOTHERMAL FIELDS

Microseismic data were collected from four geothermal fields: two geothermal fields in Japan (the Okuaizu geothermal field in Fukushima Prefecture and the Kakkonda geothermal field in Iwate Prefecture) and two geothermal fields in Europe (the Basel EGS field in Switzerland and the Soultz Hot Dry Rock (HDR) field in France).

The Yanaizu-Nishiyama geothermal power plant (capacity: 30 MW) is located in the Okuaizu geothermal field. This geothermal power plant started its operation with a capacity of 65 MW in 1995, which is the largest capacity of a single geothermal power unit in Japan. Since 2015, the Japan Oil, Gas, and Metals National Corporation has been conducting a series of
gravity-driven water injection tests in the reservoir to prevent the reduction of steam production (Yoshimatsu et al., 2020; Kato et al., 2021). The microseismic monitoring network there, known as the YAE network (Okamoto et al., 2021; 2020; 2018), is composed of nine seismic stations, including four borehole stations (Figure 1a). The borehole sensors were installed at depths of approximately 400 m, whereas the other sensors were buried in shallow postholes (several meters in depth). The sensors in the boreholes and postholes were F41-15.0 (International Earth Sciences IESE Ltd.) and Trillium Compact Posthole 20s (Nanometrics Inc.), respectively. The continuous observation of seismic waves with a sampling rate of 1,000 Hz has been conducted since 2015. For deep learning, we employed 8,873 seismic events with 40,973 P-wave arrivals and 15,288 S-wave arrivals, which were manually picked.

In the Kakkonda geothermal field, the first geothermal power unit (capacity: 50 MW) and the second geothermal power unit (capacity: 30 MW) were installed in 1978 and 1996, respectively. The Kakkonda geothermal power plant, which is composed of two units, has the second largest capacity among the geothermal power plants in Japan. The microseismic monitoring network there, known as the KKD network (Okamoto et al., 2022), is composed of 14 seismic stations (Figure 1b). An L-shaped array is formed by 11 seismic stations, while the other three seismic stations surround the L-shaped array. All sensors were KVS-300 (Kinkei System Corp.) installed on the surface (or manually buried if possible), and the sampling rate was 200 Hz. Although there are five seismic stations near the L-shaped array in addition to the KKD network (not shown in Figure 1b), we did not use their seismic records for deep learning because of their lower sampling rate (100 Hz). These stations were only included for hypocenter determination; detailed information is available in Okamoto et al. (2022). We manually picked 43,719 P-wave arrivals and 39,950 S-wave arrivals from 12,501 seismic events.
The EGS Project (Deep Heat Mining project) in Basel, Switzerland, was launched in 2006, and hydraulic stimulation was carried out in late 2006 to create an artificial reservoir for cogeneration systems (Häring et al., 2008). Hydraulic stimulation was monitored with a microseismic network (BSL network) consisting of six borehole stations at different depths (to 2600 m) (Dyer et al., 2008) (Figure 1c). Each three-component seismometer was installed with HS-1LT sensors (Geospace Technologies). Seismograms were monitored at 1,000 a sampling frequency and the triggered waveforms were analyzed using automatic processing software (Dyer et al., 2008). The hydraulic stimulation continued for five days starting on December 2, 2006 and the total injected volume reached 11,570 m³. Owing to the unexpected high seismic activity accompanied by a higher flow rate stage, the injection was aborted on day 5 of the stimulation; however, seismic activity continued even after stopping the injection (Mukuhira et al., 2013). In total, 13,500 events were triggered, of which approximately 2,700 were in May 2007 (Dyer et al., 2008). We employed manually picked 16,955 P-wave arrivals and 14,909 S-wave arrivals from 3,067 seismic events.

The other HDR project started at Soultz-sous-Forêts, Alsace, northeastern France in 1987 and was led by the EU, France, and Germany (Baria et al., 1995). Shallower (2,500 to 3,600 m) and deeper reservoirs (4,000 to 5,000 m) have been developed with several hydraulic stimulations (Baumgärtner et al., 1998; Baria et al., 2000; 2005). We used the microseismic events that occurred at the focused stimulation on the deep reservoir in 2003 via the GPK2 and GPK3 geothermal wells considering the same focal depth as the BSL field. The microseismic monitoring network had several updates according to the development stages, and at the stimulation in 2003, five downhole monitoring stations were in operation (Figure 1d). Tetrahedral 4-component sensors (Brüel & Kjær type 8318, accelerometer) were deployed for
4550, 4601, and OPS4, and 3-component geophones SM-45 (Calidus, velocimeter) were installed for EPS 1 and GPK1. The seismogram was recorded at a sampling frequency of 2,000. The total injected volume reached approximately 38,000 m$^3$ over 10 d of stimulation in 2003. The microseismic monitoring network (SLZ network) triggered 87,000 possible events, 8,076 of which were located (Asanuma et al., 2004). We employed 16,531 and 6,229 manually picked P- and S-wave arrivals, respectively, from 8,067 seismic events. It is worth noting that we converted 4-components seismograms into 3-components of orthogonal coordinates and did not use the data from the 4601 station because of technical issues with some components.

A summary of the microseismic events and stations in these geothermal fields is shown in Figure 1 and Table 1.

**METHODS**

**Input seismic waves and model structure for deep learning**

Deep learning was conducted using the model shown in Figure 2a. The basic structure of the model (i.e., the size of each layer, number and order of the layers, length of input seismic waves, etc.) was the same as that of Ross et al. (2018a). The model was composed of three convolutional neural network (CNN) layers and two fully connected neural network (FN) layers with rectified linear units (ReLU) as the activation function. The Adam optimization algorithm was employed as the optimizer (learning rate = 0.001). However, in contrast to their model, which employed a sampling rate of 100 Hz, seismograms were resampled to 200 Hz (if the original sampling rate was higher), and a band-pass (5 to -30 Hz) filter was applied to them. We
introduced a dropout layer (dropout rate = 0.2) that was different from their model immediately before the final output layer (softmax) to prevent overtraining. The model was trained using 4-s long seismic data adjusted for P- and S-wave arrivals. We placed the location of P-wave arrivals between 1.0 s and 3.0 s at random in the 4-s long time window for P-waves. On the other hand, the start position of the 4-s long seismic data for S-waves was adjusted to a point 0.5 to -1.5 s prior to the preceding P-wave arrival at random (Figure 1b). The datasets for P- and S-waves were obtained from the vertical and horizontal components of the seismic waves, respectively. In the dataset for S-waves, we included some cases in which the manual picking of S-wave arrivals was not performed because of obscure arrivals (e.g., low signal-to-noise ratio, effect of S-wave radiation patterns, attenuation due to fluids in seismic wave paths, etc.) Therefore, the trained model does not necessarily read S-wave arrivals as human analysts do.

**Training using data from each field and fine-tuning by the data of the remaining field(s)**

First, we trained the model using YAE, KKD, BSL, and SLZ data individually. The data were divided into training data (70%) and validation data (30%), except for in the case of YAE, in which 5,359 P-waves and 5,322 S-waves were prepared as validation data, in addition to the training data. The number of epochs for training was 500, and the seismic data were binned for every 128 data points during the training. We evaluated the quality of the trained models at each epoch using validation loss (cross-entropy loss function), and the model with the minimum validation loss was chosen as the best model. The best models of the respective fields were the YAE, KKD, BSL, and SLZ models.
Based on the best model for each field, we conducted fine-tuning using the data of the remaining fields sequentially. For fine-tuning, we optimized all parameters in the CNN and FN layers. The order of the additional training was the reverse of the amount of training data (P-wave + S-wave data) in the fields, as follows: SLZ, BSL, YAE, and KKD. Therefore, when we conducted fine-tuning of the SLZ model using the rest of the field data, the order of the additional training was BSL, YAE, and KKD data. The number of epochs for additional training was 100, and the model with the minimum validation loss (after the first epoch) was chosen as the best model. We refer to this model as the SLZ + BSL + YAE + KKD model. In terms of fine-tuning based on the SLZ model, there were five possible combinations other than the SLZ + BSL + YAE + KKD model: SLZ + BSL, SLZ + YAE, SLZ + KKD, SLZ + BSL + YAE, and SLZ + BSL + KKD models. The success of each model was evaluated based on the standard deviation and the average error between the true and predicted arrival times of the validation data.

Application of the trained models to continuous seismic waves with an STA/LTA trigger

Our deep learning model can be applied to 4-s long seismic waves already triggered by continuous seismic waves. However, in practical hypocenter determination, we must start from continuous seismic waves. Taking the five-month continuous seismic waves recorded by the YAE network as an example, we demonstrate a practical procedure to determine a robust hypocenter distribution using a deep learning model. First, we employed a simple method to trigger potential P-wave arrivals from continuous seismic waves (vertical component): the STA/LTA method. We employed a 0.8-s long window for the short-term average, leading to a 15.0-s long trailing window for the long-time average, which moved together with a time interval
of 1.5 s over the continuous seismic waves. The threshold for triggering using the STA/LTA method was >1.5. The STA/LTA method was applied to all stations individually. After triggering using the STA/LTA method, a precise P-wave arrival was read by the deep learning model using a 4-s long time window that starts 2.0 s prior to the STA. Then, another 4-s long time window starting 1.0 s prior to the machine-picked P-wave arrival was applied to the horizontal components of the seismic wave to pick the subsequent S-wave arrival.

Thus far in the process, numerous potential P- and S-wave arrivals were picked using deep learning, most of which were not from microseismic events in and around the geothermal reservoir but derived from noise or regional earthquakes, owing to the low threshold for the STA/LTA method. Here, we employed a seismic phase association method based on “PhaseLink”, a deep learning model invented by Ross et al. (2019). Among the numerous potential P- and S-wave arrivals at all seismic stations, PhaseLink mines the possible pair that is derived from a common seismic event. Because PhaseLink needs to be trained using synthetic data beforehand, we employed a 4 × 4 × 5 km cuboid under the YAE network, in which the P- and S-wave travel times of microseismic events were synthesized. Using these synthetic calculations, PhaseLink learns the possible pairs of P- and S-wave arrivals at seismic stations in the YAE network.

Hypocenter determinations were then conducted based on the pairs of P- and S-wave arrivals linked by PhaseLink using HYPOINVERSE-2000 (Klein, 2002). Among the determined hypocenters, reliable ones were extracted based on the following criteria: (1) a root-mean-square of the residual of travel time less than 0.1 s; (2) a better quality of hypocenter determinations defined in HYPOINVERSE-2000 than the B rating. Finally, we applied the DD method to the
extracted hypocenters to obtain a more precise hypocenter distribution. In the DD method, the original hypocenter locations were relocated to explain the time differences in phase arrivals between pairs of seismic events. The time differences were calculated based on cross-correlation (Shaff and Waldhauser, 2005) (for P-waves) and based on the event catalog (for P- and S-waves). The cross-correlation was conducted using a 2.0 s time window that began 0.1 s prior to P-wave arrivals. The time difference based on the cross-correlation was used if the value was greater than 0.65. Cross-correlation coefficients were also used to cluster relocated hypocenters. Each hypocenter was considered a node in the context of graph theory, and we employed a condition wherein a pair of nodes was tied by an edge when their cross-correlation coefficient was greater than 0.65 at least one seismic station, and the distance between the pair was shorter than 200 m. A depth-first search (Tarjan, 1972) was applied to identify clusters. The clusters can indicate fine structures, for example, fracture networks of reservoirs, geological boundaries, and paths of fluid flows. The whole procedure mentioned in this section is summarized in Figure 2c.

RESULTS

Training the neural network and application to field data

Training using field data was conducted separately for the P- and S-waves. The training and validation losses as functions of the epoch number are shown in Figure 3. In the cases of YAE and KKD, the best model (where the validation loss is the minimum) appeared at an early epoch number for both the P- and S-waves (red squares, Figures 3a-d). On the other hand, in the cases of BSL and SLZ, the best model appeared at a later epoch number (Figures 3e-h). The oscillations in the loss curves against the epoch number were partly due to the batch size employed (128). The batch size determines the number of data points used in each iteration. The
smaller the batch size, the more sensitive the model optimization process is to the data, which leads to oscillations. Additional training was conducted using the best model for each field.

Each of the best models trained using the solo-field data was inspected using the validation data of the YAE (Figure 4). For P-wave picking, the YAE model provided the best accuracy among the models in terms of the standard deviation and the average of the errors between the true and predicted P-wave arrival times (Figure 4a). The KKD model yielded the second-best accuracy (Figure 4c). The difference in the standard deviation and average between the KKD and YAE models was minor, whereas the BSL and SLZ models yielded poorer results than the other models (Figures 4e and g). This was partly due to differences in hypocentral distances. The hypocentral distances of the YAE and KKD cases were similar, whereas those of the BSL and SLZ cases were much greater (deeper) than the YAE case. If the fields have different settings (geometry of stations and distance to the focal region) that leads different S-P time (time difference between S- and P-wave arrivals) and frequency content, the performance of their deep learning models were not satisfactory. The KKD model was the most accurate for S-wave picking (Figures 4b, d, f, and h). On the other hand, the S-wave arrivals predicted by the BSL and SLZ models tended to deviate from the true ones (Figures 4f and h).

The results of the application of the fine-tuning models to the validation data of the YEA are shown in Figure 5 (except for the results of the fine-tuning by the YAE data itself). Fine-tuning of the BSL and SLZ models using the KKD data greatly improved the accuracy for both P- and S-wave picking (Figures 5a-d and Figures 5g and h). The standard deviation of each case was less than half that of the model before fine-tuning. However, fine-tuning the SLZ model using BSL data did not improve the accuracy of P- and S-wave picking well (Figures 5e and f).
This suggests combination two deeper fields data still did not cover the features of YEA validation data presumably due to its large S-P time.

The results when the solo- and fine-tuning models were applied to the validation data of KKD, BSL, and SLZ, as well as YAE, are shown in Figure 6; however, we did not apply the fine-tuning models by the self-data except for the SLZ + BSL + YAE + KKD model. In Figure 6, the closer the standard deviation and average approach zero, the more accurate the result. In the results for the YAE and KKD data (Figures 6a-d), the YAE and KKD models yielded similar accuracies for both P- and S-wave picking. There was a case that the KKD model was slightly better than the YAE model itself (S-wave picking for the YAE data, Figure 6b). On the other hand, the BSL, SLZ, and SLZ + BSL models exhibited relatively lower accuracy than the YAE and KKD models for both P- and S-wave picking. Fine-tuning these three models using KKD and/or YAE data improved the accuracy. In particular, in the S-wave picking for the YAE data (Figure 6b), the fine-tuning models using the KKD data (BSL + KKD, SLZ + KKD, and SLZ + BSL + KKD models) showed a better accuracy than the YAE model itself. Regarding the BSL data and SLZ data (Figures 6e-h), almost all the models (including the self-solo models; the BSL model and SLZ model) showed similar accuracy for both P- and S-waves. Only the case in which the BSL model was applied to P-waves of the SLZ data gave a relatively lower accuracy compared to the other models (the standard deviation and average approached 1 s and –1 s, respectively). The SLZ + BSL + KKD + YAE model, which is the fine-tuning model for all data, showed comparable accuracy with the solo models of the respective fields.
Automatic hypocenter determination from continuous seismic waves and fine structures inferred by the DD method

Approximately 10,000 events per day were triggered by the continuous seismic waves recorded by the YAE after applying the STA/LTA method. The YAE model and the SLZ + BSL + KKD model were applied to the triggered events to pick P- and S-wave arrivals. The SLZ + BSL + KKD model is important because it does not learn YAE data. All of the picked P- and S-wave arrivals were linked by PhaseLink, and finally, 4,230 events based on the YAE model and 4,157 events based on the SLZ + BSL + KKD model were selected as plausible microseismic events located in and around the geothermal reservoir. Their hypocenters were determined using HYPOINVERSESE-2000 and selected according to the criteria described in the Methods section (Figure 7). The number of hypocenters determined by the YAE and SLZ + BSL + KKD models were 712 and 633, respectively, while that determined by manual picking was 1,631. Although the number of hypocenters determined by the deep learning models was less than that determined by manual picking, their distributions roughly coincided with each other. Both the YAE model and SLZ + BSL + KKD model indicated that microseismic events were divided into two major clouds, which were located around the center of the microseismic station network (one had a north-south strike and the other had a northwest-southeast strike), as indicated by manual picking. This suggests that the SLZ + BSL + KKD model, which trained by field data other than the YAE data, could predict the hypocenter distribution equivalent to that of the self-solo model (YAE model).

In the DD method, each event was linked to its neighbors through commonly observed phases with a maximum hypocentral separation between the linked events (1.0 km). After
applying the DD method, 191 hypocenters divided into 14 clusters were detected using the YAE model (Figure 8a) and 118 hypocenters were divided into six clusters using the SLZ + BSL + KKD model (Figure 8b). The relocated hypocenters, both by the YAE model and the SLZ + BSL + KKD model, formed a cluster around the injection well, which seemed to be derived from the water injection. In particular, in the result of the YAE model (Figure 8a), several seismic clusters were confirmed along the northmost existing fault (northeast-southwest strike) (dashed lines, Figure 8) and its extension line. The corresponding clusters were confirmed in the results of the SLZ + BSL + KKD model, although fewer seismic events were present (Figure 8b). These seismic clusters illuminated the existing fault and likely indicated that the existing fault extended to a greater distance. In addition to the clusters mentioned above, several minor clusters were observed in the southern area of the injection well.

DISCUSSION

Improving the accuracy of machine picking and the applicability to a new field

When the deep learning models developed in this study (i.e., solo models and fine-tuning models) were applied to the YAE data, the YAE model itself, the KKD model, and fine-tuning models using the KKD data showed more accurate results than the BSL, SLZ, and SLZ + BSL models, which learned neither the YAE data nor the KKD data. Moreover, the KKD model and fine-tuning models using KKD data picked S-wave arrivals more accurately than the YAE model itself. One of the reasons for the relatively lower accuracy of the BSL, SLZ, and SLZ + BSL models is the number of training data; the numbers in the YAE and KKD cases were larger than
those in the BSL and SLZ cases (Table 1). This could also be attributed to the fact that the S-P times in the BSL and SLZ data had different features from those in the YAE and KKD data because of the hypocentral distances. The S-P times for each field are shown in Figure 9. The volume zone of the S-P times of the KKD data was approximately 200-300 ms, while the main volume zone of the YAE data was also 200-300 ms. Although this volume zone seemed to extend to around 500 ms in the YAE data, which was mainly derived from the sensors on the surface, the main volume zone of the S-P times was similar between the YAE and KKD data. On the other hand, the main volume zone of S-P times of the BSL data (around 900 ms) was largely different from that of the YAE data. Regarding the SLZ data, there were several volume zones around 100-200 ms and 400 ms, which were slightly shifted from those of the YAE data. Thus, the features of the distribution of the S-P times of the BSL and SLZ data were different from those of the YAE data. For the training of P-waves, we employed the 4-s long time window, in which P-wave arrivals were located between 1.0 s and 3.0 s. Therefore, S-wave arrivals were also within the 4-s long time window in most cases. The location of S-wave arrivals in the 4-s time window, that is, the length of the S-P times, likely affected the training of P-waves. It is possible that the deep learning model learned the location of P-wave arrivals in connection with the location of S-waves. Therefore, there was an accuracy gap between the YAE and KKD models and the SLZ and BSL models when applied to the YAE data.

As mentioned above, one of the factors that the fine-tuning models based on the KKD data worked well for picking the YAE data seemed to be the similarity in S-P times between the KKD and YAE data, which was derived from the similarity in their geometries (location of stations and hypocenters). Meanwhile, the amount of training data in the KKD model was greater than that in the YAE model for S-waves, which could lead to more accurate S-wave picking by
the KKD model than that by the YAE model. Additionally, manual picking of the YAE data was performed by several human analysts, while that of the KKD data was performed by a solo human analyst, which could provide consistency in P- and S-wave picking over the KKD data, leading to a more accurate deep learning model.

Thus far, it appears that the use of data for learning from fields with similar geometry conditions to those of the target field (or data of the target field itself if available) is indispensable when applying an existing deep learning model to a new field. This is true for the well-known P-wave picking model established by Ross et al. (2018a). Their model was trained using earthquakes in a wider region than in the YAE case. The maximum epicentral distance used in their model was ~120 km, whereas it was ~1 km in the case of YAE. Therefore, the prediction of P-wave arrivals for the YAE data by their model was inaccurate compared with the YAE model (Figure 10). Their model tended to predict P-wave arrivals earlier than true ones. The standard deviation and average of the errors between true and predicted arrivals by their model were 0.20 and 0.03 s, respectively, while those by the YAE model were 0.10 and 0.00 s. It is worth noting that since the model by Ross et al (2018a) supposed that P-wave arrivals were within 0.15-0.25 s in the 4-s long time window in the training (while our model supposed that they were within 0.05 to 0.35 s), the standard deviation and average when the YAE model was applied were also calculated for the same range.

Estimation of fine reservoir structures of developing fields and green fields
It is important to note that the SLZ + BSL + KKD model, which trained by the data other than the YAE data, could indicate equivalent fine structures in the Okuaizu geothermal field as the YAE model itself did after applying the DD and clustering methods. Because relocation by the DD method and clustering were performed automatically based on parameters determined beforehand, the fine structures could be obtained automatically from continuous seismic waves. The fact that the model that trained by other field data work well, as well as the fact that fine structures could be obtained automatically, indicated that the model and procedure developed in this study could reveal fine structures in local geothermal fields and other similar fields (e.g., shale gas field and carbon dioxide capture and storage field) without manual processes. We propose them as MAGMA: Machine learning Automatic picker for Geothermal Microseismicity Analysis. Especially, we call the SLZ + BSL + YAE + KKD model as MAGMA model. The main requirement for the application of the MAGMA model to a target field was simply that the field had a similar geometry to the YAE, KKD, SLZ, or BSL case. As a result, the application of our model and procedures could lead to saving on the resources (e.g., budget, human resources, and time) required for the manual picking of P- and S-wave arrivals and further analyses to reveal fine structures. Here, we should note that the similarity in geometry is not a sufficient condition for accurate picking. If the target field has its own data, fine-tuning is recommended, which can surely improve the picking accuracy. Meanwhile, our model and procedure can be applied even if the target field does not have its own data, such as in green fields.

CONCLUSION
In this study, we established a deep learning model specialized for microseismic monitoring systems for the picking of P- and S-wave arrivals using microseismic data from four different geothermal fields. Using the resulting deep learning model, we established an automatic procedure to reveal subsurface fine structures from continuous seismic waves recorded in the Okuaizu geothermal field. This provided insights into subsurface fluid flow paths and fault systems that comprise the geothermal reservoir.

The deep learning model and procedures proposed in this study (we called MAGMA) can be used to reduce the human resources, time, and cost needed for the monitoring of microseismicity and further processing to evaluate fine structures, which represents a significant advancement in microseismic monitoring and real time analysis systems. A highlight of our findings is that the deep learning model that trained by other field data was able to reveal equivalent structures to the model that learned data from the Okuaizu geothermal field.

Notably, the key to the successful application of the deep learning model trained using other field data to that a target field is a similarity in the geometry conditions (e.g., locations of stations and hypocenters). As a result, the deep learning model developed in this study (MAGMA model) has the potential to be applicable even if data for the target field is not available, such as in the case of green fields.
REFERENCES


Okamoto, K., K. Imanishi, and H. Asanuma, 2022, Structures and fluid flows inferred from the microseismic events around a low-resistivity anomaly in the Kakkonda geothermal field, Northeast Japan, Geothermics, 100, 102320.


van der Elst, N. J., M. T. Page, D. A. Weiser, T. H. W. Goebel, and S. M. Hosseini, 2016, Induced earthquake magnitudes are as large as (statistically) expected, Journal of Geophysical Research Solid Earth, 121, 4575–4590.


Figure 1: Geometry of seismic stations and hypocenters. (a) Okuaizu geothermal field, (b) Kakkonda geothermal field, (c) Basel geothermal field, and (d) Soultz geothermal field.
Figure 2: Deep learning model employed in this study. (a) Detailed information for each layer. (b) Time windows used for P- and S-waves. (c) Whole procedure to reveal fine reservoir structures using continuous waves.
Figure 3: Training and validation losses against epoch. Results for P- and S-waves of (a-b) YAE, (c-d) KKD, (e-f) BSL, and (g-h) SLZ data. Red squares indicate the minimum validation loss.
Figure 4: Histograms of the errors between true and predicted arrivals when applying the solo models to the validation data of YAE. Results for P- and S-wave arrivals obtained by (a-b) the YAE model, (c-d) the KKD model, (e-f) the BSL model, and (g-h) the SLZ model.
Figure 5: Histograms of the errors between true and predicted arrivals when applying the fine-tuning models to the validation data of YAE. Results for P- and S-wave arrivals obtained by (a-
b) the BSL + KKD model, (c-d) the SLZ + KKD model, (e-f) the SLZ + BSL model, and (g-h) the SLZ + BSL + KKD model.
Figure 6: Distributions of the standard deviation and average of the errors between true and predicted arrival times when applying the deep learning models to each field data. Results for P- and S-waves of (a-b) YAE, (c-d) KKD, (e-f) BSL, and (g-h) SLZ data.
Figure 7: Hypocenter distributions in Okuaizu geothermal field determined by manual picking and machine picking. Results obtained by (a) the YAE model and by (b) the SLZ + BSL + KKD model.
Figure 8: Distributions of seismic clusters in Okuaizu geothermal field. Gray dots indicate the original hypocenters determined by machine picking and colored circles indicate relocated hypocenters by the DD method. Each seismic cluster is denoted by a different color. Black circles indicate hypocenters that do not belong to any clusters. Results obtained by (a) the YAE model and by (b) the SLZ + BSL + KKD model. Dashed lines indicate known existing faults.
Figure 9: Distribution ratio of S-P times of each field.
Figure 10: Distribution of the errors between true and predicted arrivals when applying the YAE model (blue) and the model by Ross et al. (2018a) (red) to the YAE data.
## LIST OF TABLES

Table 1: Summary of the number of training data and spec of microseismic station networks

<table>
<thead>
<tr>
<th>Site</th>
<th>Number of training data</th>
<th>Product name</th>
<th>Type</th>
<th>Feature</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yanaizu (YAE)</td>
<td>40,973 (46,332)</td>
<td>Trillium Compact PH 20s</td>
<td>In posthole,</td>
<td>−3 dB points at 20 s and 108 Hz</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(Nanometrics Inc.)</td>
<td>velocimeter</td>
<td>Natural frequency</td>
</tr>
<tr>
<td></td>
<td></td>
<td>F41-15.0</td>
<td>In borehole,</td>
<td>Natural frequency</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(International Earth Sciences IESE Ltd.)</td>
<td>velocimeter</td>
<td>of 15 Hz</td>
</tr>
<tr>
<td>Kakkonda (KKD)</td>
<td>30,603 (43,719)</td>
<td>KVS-300</td>
<td>On the surface,</td>
<td>Natural frequency</td>
</tr>
<tr>
<td>Basel (BSL)</td>
<td>11,868 (16,955)</td>
<td>HS-1LT</td>
<td>In borehole,</td>
<td>Natural frequency</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(Geospace Technologies)</td>
<td>velocimeter</td>
<td>of 4.5 Hz</td>
</tr>
<tr>
<td>Soultz (SLZ)</td>
<td>11,571 (16,531)</td>
<td>Type 8318</td>
<td>In borehole,</td>
<td>−3 dB points at 10-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(Brüel &amp; Kjær)</td>
<td>accelerometer</td>
<td>1,000 Hz</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SM-45</td>
<td>In borehole,</td>
<td>Natural frequency</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(Calidus)</td>
<td>velocimeter</td>
<td>of 15 Hz</td>
</tr>
</tbody>
</table>