Microseismic Monitoring using Transfer Learning: Example from the Newberry EGS

Zi Xian Leong[1] and Tieyuan Zhu[1]

[1]Pennsylvania State University

March 07, 2024

Abstract

Enhanced geothermal systems (EGS) are promising for generating clean power by extracting heat energy from injection and extraction of water in geothermal reservoirs. The stimulation process involves hydroshearing which reactivates pre-existing cracks for creating permeability and meanwhile inducing microearthquakes. Locating these microearthquakes provide reliable feedback on the stimulation progress, but it poses a challenging nonlinear inverse problem. Current deep learning methods for locating earthquakes require extensive datasets for training, which is problematic as detected microearthquakes are often limited. To address the scarcity of training data, we propose a transfer learning workflow using probabilistic multilayer perceptron (PMLP) which predicts microearthquake locations from cross-correlation time lags in waveforms. Utilizing a 3D velocity model of Newberry site derived from ambient noise interferometry, we generate numerous synthetic microearthquakes and 3D acoustic waveforms for PMLP training. Accurate synthetic tests prompt us to apply the trained network to the 2012 and 2014 stimulation field waveforms. Predictions on the 2012 stimulation dataset show major microseismic activity at depths of 0.5–1.2 km, correlating with a known casing leakage scenario. In the 2014 dataset, the majority of predictions concentrate at 2.0–2.9 km depths, consistent with results obtained from conventional physics-based inversion, and align with the presence of natural fractures from 2.0–2.7 km. We validate our findings by comparing the synthetic and field picks, demonstrating a satisfactory match for the first arrivals. By combining the benefits of quick inference speeds and accurate location predictions, we demonstrate the feasibility of using transfer learning to locate microseismicity for EGS monitoring.

Microseismic Monitoring using Transfer Learning: Example from the Newberry EGS

Zi Xian Leong[1,†] and Tieyuan Zhu[1,2]

[1] Department of Geosciences, The Pennsylvania State University, University Park, PA, USA.

[2] EMS Energy Institute, The Pennsylvania State University, University Park, PA, USA.

†Currently at Chevron Technical Center, a division of Chevron U.S.A. Inc.

Corresponding author: Zi Xian Leong (zxnleong@gmail.com)

Key Points:

- We present a novel transfer learning workflow to predict microearthquake locations in EGS, addressing data scarcity for training
- Application to Newberry EGS reveals accurate microearthquake locations, validated against known geological features
- Employs probabilistic multilayer perceptrons that map cross-correlation time lags to microearthquake locations

Abstract

Enhanced geothermal systems (EGS) are promising for generating clean power by extracting heat energy from injection and extraction of water in geothermal reservoirs. The stimulation process involves hydroshearing which reactivates pre-existing cracks for creating permeability and meanwhile inducing microearthquakes. Locating these microearthquakes provide reliable feedback on the stimulation progress, but it poses a challenging nonlinear inverse problem. Current deep learning methods for locating earthquakes require extensive datasets for training, which is problematic as detected microearthquakes are often limited. To address the scarcity of training data, we propose a transfer learning workflow using probabilistic multilayer perceptron (PMLP) which predicts microearthquake locations from cross-correlation time lags in waveforms. Utilizing a 3D velocity model of Newberry site derived from ambient noise interferometry, we generate numerous synthetic microearthquakes and 3D acoustic waveforms for PMLP training. Accurate synthetic tests prompt us to apply the trained network to the 2012 and 2014 stimulation field waveforms. Predictions on the 2012 stimulation dataset show major microseismic activity at depths of 0.5–1.2 km, correlating with a known casing leakage scenario. In the 2014 dataset, the majority of predictions concentrate at 2.0–2.9 km depths, consistent with results obtained from conventional physics-based inversion, and align with the presence of natural fractures from 2.0–2.7 km. We validate our findings by comparing the synthetic and field picks, demonstrating a satisfactory match for the first arrivals. By combining the benefits of quick inference speeds and accurate location predictions, we demonstrate the feasibility of using transfer learning to locate microseismicity for EGS monitoring.

Plain Language Summary

Enhanced geothermal systems (EGS) are an emerging technology that generates clean electricity by injecting water into underground hot rocks and pumping it back to the surface for power generation. However, this stimulation process causes tiny earthquakes, known as microearthquakes. Tracking the location of these microearthquakes is crucial for monitoring the EGS creation process. Unfortunately, finding where these microearthquakes occur is a complex task. Using deep learning methods is challenging because of the general lack of microearthquakes for training. To overcome this, we employ transfer learning, which allows computer models to train on realistic data, and eventually deploy to real-world EGS microearthquake data. We create a realistic geological model of the Newberry EGS site and generate many artificial microearthquake data for deep learning training. During the application on field data from 2012 and 2014 stimulation, the computer model successfully identifies the depth and location of MEQs. Our results match well with what we already know about the underground structure, such as the presence of natural fractures in the rock. This study shows that our approach can effectively predict microearthquake locations even when presented with limited earthquake data for training, which is promising for monitoring and improving EGS operations in the future.

1 Introduction

Geothermal energy has emerged as a valuable and sustainable resource in the global energy landscape, which harnesses the Earth's natural heat to generate electricity, providing a reliable and consistent supply, unlike intermittent resources such as solar or wind power (Muffler & Cataldi, 1978). As a low-emission energy form, geothermal power mitigates greenhouse gas emissions and reduces the nation's reliance on fossil fuels (Tester et al., 2006). According to the U.S. Energy Information Administration (EIA, 2023), geothermal energy generation in 2022 reached approximately 17 billion kWh, positioning the United States as the leading global producer of geothermal electricity. Moreover, electricity generated from geothermal plants is projected to increase to 37.2 billion kWh in 2050. Enhanced geothermal systems (EGS) technology harnesses heat energy produced from areas of young tectonism and volcanism, but contains relatively low permeability (e.g., Häring et al., 2008; Cladouhos et al., 2016; Schill et al., 2017; Lu, 2018; Tomac and Sauter, 2018). In an EGS, fluid is injected into the subsurface under carefully controlled conditions, which caused pre-existing fractures to reopen, enhancing permeability. Increased permeability allows fluids to circulate in the now-fractured rock and to transport heat to the surface where electricity can be generated.

The creation of EGS has been widely known to induce microearthquakes (MEQs) (Zang et al., 2014; Majer et al., 2007). These MEQs, serving as reservoir stimulation diagnostic indicators, can locate fluid-induced fractures and monitor EGS stimulation progress such as crack propagation, permeability evolution, and temperature changes (Izadi and Elsworth, 2013; Fang et al., 2016). However, elevated occurrence of MEQs may lead to negative public perception regarding EGS deployment, particularly felt seismicity may be perceived as an isolated annoyance. Furthermore, there is concern about the cumulative effects of recurrent events and the potential of larger earthquakes in the future (Majer et al., 2007). Consequently, there is both a scientific and societal need to locate and monitor MEQs associated with EGS reservoir stimulation.

Conventional earthquake location methods involve iteratively minimizing the difference between picked P- and/or S-wave first arrival times and predicted data at multiple seismic stations (Geiger, 1912; Tarantola & Valette, 1982; Bondar et al., 2014; Karasözen & Karasözen, 2020). While these methods have been widely employed in seismology, they exhibit certain limitations. The accuracy of earthquake location estimates can be affected by convergence issues, particularly when the initial location guess is not sufficiently close to the true hypocenter, the solution may converge at a local minimum, leading to inaccurate location estimates. Additionally, conventional methods can be computationally intensive, particularly when applied to large datasets or in regions with complex geology (local heterogeneities). As such, most location algorithms rely on one-dimensional (1D) velocity models, where the velocity changes only with depth. Furthermore, waveform-based methods that are based on time-reversal imaging principles utilize finite difference to compute time-reversed seismograms and the actual source location is determined by identifying the point of highest energy concentration (e.g., McMechan, 1982; Chang and McMechan, 1994; Gajewski and Tessmer, 2005; Zhu, 2014; Li et al., 2020). Wavefield simulation method is unsurprisingly computationally expensive, and the energy focusing can be ambiguous for noisy data and very heterogeneous models. Waldhauser and Ellsworth (2000) proposed hypoDD, a widely used location inversion method that iteratively minimizes the misfit between theoretical and observed differential travel-times for pairs of earthquakes (double-difference) at each station. Nonetheless, the system can get very large if all event pairs are used in double-difference methods and reducing the efficiency of location estimation.

Deep learning (DL) techniques have been increasingly applied in earthquake seismology. For example, DL has seen significant developments in earthquake event phase detection (Ross et al., 2018; Dokht et al., 2019), phase picking (Zhu and Beroza; 2018; Mousavi et al., 2020), and phase association (Ross et al., 2019; Zhu et al., 2022). For DL-based earthquake location inversion, a large majority of studies rely heavily on training with labeled field data. Perol et al. (2018) used convolutional neural network (CNN) that trained on ~2,900 single station events near Guthrie, Oklahoma, in which the CNN accepts three-component waveforms and predicts earthquake location groups of six clusters. Later studies improved the earthquake location inversion method by employing more advanced DL algorithms and utilize multi-station three-component waveforms as input to predict three-dimensional (3D) locations. For example, Kriegerowski et al. (2019) employed deep CNN to predict easting, northing, and depth of earthquakes based on ~3,000 events from Western Bohemia, Czech Republic. Van den Ende and Ampuero (2020) used graph neural networks to predict the source latitude, longitude, depth, and magnitude based on ~1,300 events from Southern California. Shen and Shen (2021) used deep CNNs that trained on USGS Combined Cataglog earthquakes (~1,800) to predict latitude, longitude, depth, and origin time of events. Zhang et al. (2021) adapted deep CNNs to predict 3D event location probabilities based on ~1,000 events from Central Apennies, Italy. Using single-station waveforms, Mousavi and Beroza (2020) employed Bayesian neural networks to predict epicenter, depth, and origin time based on the Stanford Earthquake Data Set (~450k events).

Comparing natural earthquakes to geothermal induced MEQs reveals several distinct differences, particularly in terms of their detectability (Fang et al., 2016; Templeton et al., 2020). MEQs are generally characterized by lower magnitudes and higher scarcity compared to natural earthquakes. The lower magnitudes make MEQs more challenging to detect, as they are often masked by background noise. This results in fewer MEQ events detected in conventional catalogs. This scarcity of MEQs poses a significant challenge for DL training, as the limited amount of available data restricts the ability to build robust and accurate DL models for solving the nonlinear MEQ location inversion problem. Consequently, even though DL algorithms are strong solvers for nonlinear problems and have quick inference speeds, the data scarcity for training presents as the major challenge for using DL guided solutions to accurately locate MEQs. Moreover, the accuracy of predicted locations using conventional earthquake location methods (e.g., minimizing travel-time misfit) highly depends on the velocity model used. Simplified velocity models can result in less precise location predictions due to the lack of local heterogeneities present in the model. Using higher resolution velocity models that include more local geological features will incur higher computation costs. As such, it is pivotal to develop a practical method that combines the benefits of DL (quick inference times and strong nonlinear solving abilities), address the paucity of field training data, and integrates high-resolution realistic velocity models, to estimate induced MEQ locations for EGS monitoring.

In this study, we present a transfer learning workflow using probabilistic multilayer perceptron (PMLP) to accurately predict MEQ locations from waveform data. Transfer learning involves applying a machine learning model, initially trained on one dataset, to a different but related dataset. The knowledge transfer technique is especially beneficial in applications scenarios where collecting extensive training data is impractical or unfeasible. This approach serves as the basis of our study to locate field MEQs at the Newberry EGS site. The workflow encompasses three parts. Firstly, we use a high-resolution 3D velocity model created by Matzel et al. (2014) to simulate numerous synthetic MEQ events using 3D acoustic finite-difference modeling. From the synthetic waveforms, we extract its first arrivals. In practice, since we do not have the MEQs event origin time, we compute the cross-correlation of the first arrivals such that the first arrival of the master trace is at zero time lag. The time lags at other receivers contain the same moveout pattern as the first arrivals. Secondly, we train a PMLP that inputs cross-correlation time lags and outputs the locations (x, y, z) of MEQs. Lastly, we apply the trained PMLP onto the field dataset to obtain field MEQ location predictions. We are essentially leveraging transfer learning principles by allowing the neural network to train on realistic or *physics-informed* synthetic dataset, and then apply its knowledge learned onto field waveforms to predict the induced MEQ locations.

This manuscript is organized as follows. Firstly, we provide some background on the Newberry EGS and its field collected dataset. Secondly, we introduce our methodology, including the Newberry 3D velocity model, synthetic training dataset generation, and PMLP. Lastly, we discuss and interpret our results, and showcase our potential improvements to the previous understanding of Newberry EGS microseismicity.

2 Newberry EGS

Newberry Volcano is a shield volcano located in central Oregon, about 20 mi (35 km) south of the city of Bend and approximately 40 mi (65 km) east of the crest of the Cascade Range. The Newberry EGS was operated by AltaRock Energy and Davenport Newberry to test and demonstrate the EGS technology. After an extensive study of the state of the stress for the area (Cladouhos et al., 2011a; Davatzes and Hickman, 2011), this location was selected due to a very low permeability rate as well as a large conductive thermal anomaly that yields high-temperatures (Cladouhos et al., 2011b), making it ideal to test the creation of an EGS. Borehole logs reveal natural fractures extending from approximately depths of 2,000 m to 2,700 m. At these depths, the interpreted lithology consists of tuffs, basalt-andesite, and granodiorite. The EGS demonstration was stimulated two times, first in 2012 and later in 2014, to induce hydroshearing in the reservoir and enhance the movement of fluids through the system (Cladouhos et al., 2016).

In the 2012 fluid stimulation, there was a suspected casing leakage which caused induced MEQs at shallower than the intended depths. In the fall of 2014, casing repairs and re-stimulations were made. In the drilling well, the perforated liner is used to create multiple pathways for fluid injection into the rock formation, enabling efficient fracturing and increased heat exchange between the injected fluid and the surrounding hot rocks. The perforated liner starts at 1,912 m (6,272 ft) to 3,045 m (9,990 ft), along with a blank liner extending from 2,289 m (7,509 ft) to 2,493 m (8,177 ft). The depths at which the perforated liner is installed (1,912 m - 3,045 m) is considered the targeted depth for EGS stimulations. The experiment had a monitoring array of seven surface seismic stations and eight borehole stations. Figure 1 shows the general vicinity of Newberry EGS site. For the purposes of our study, we only show the borehole stations because the recorded waveforms are frequently missing at the surface stations. As such, we only work on data traces from the eight borehole stations throughout our study.

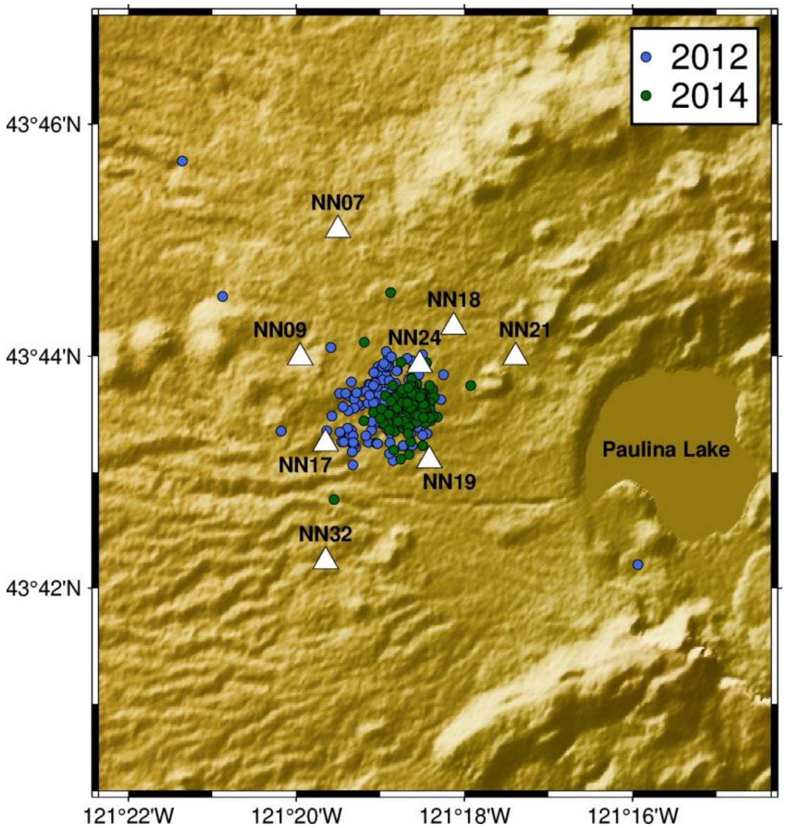

Figure 1: Aerial view of Newberry EGS site. The eight NN stations are borehole seismic stations. Events in blue are from the initial location catalog from the 2012 stimulation. Events in green are the corresponding locations of 2014 stimulation.

2.1 Microseismicity of the 2012 and 2014 EGS Stimulation

The 2012 stimulation lasted from Sept. 1, 2012, to Dec. 31, 2012. About 40,000 m^3 of water were injected with about 90% of the events were above the casing shoe (depths less than 1,830m (6,000 ft)), suggesting that injected fluid had leaked out of the casing to stimulate relatively shallow and cool rock. In the summer of 2013, caliper and video logs confirmed that there was both a horizontal crack in the casing at 683 m (2,240 ft) depth and a leak in the parasitic aeration string (AltaRock, 2014). In 2014, casing repairs were made, and second stimulation was conducted on Sept. 22, 2014, until Nov. 30, 2014. As for microseismicity, the seismic acquisition software automatically identified events, generated preliminary P- and S-wave picks and locations.

During the 2012 stimulation, about 175 events were located with magnitudes between M 0.0 and M 2.3. As for the 2014 stimulation, about 398 events were located with magnitudes between M 0.0 and M 2.2 (Cladouhos et al., 2016).

As for the data availability (http://fracture.lbl.gov/Newberry/Location.txt – assessed October 2019), there were only 149 datasets comprising waveforms and locations for the 2012 stimulation. For the 2014 stimulation, only 334 datasets are available.

3 Methodology

The main objective of this study is to develop DL algorithms to predict the locations of MEQs induced in the Newberry EGS, using waveform features, specifically cross-correlation time lags. The workflow is summarized in Figure 2. The workflow methodology can be divided into four parts. Firstly, we obtain a realistic seismic velocity model that is derived from field observations. Secondly, we simulate numerous synthetic MEQs, and their corresponding waveforms based on the field-informed velocity model. Thirdly, we use a neural network (PMLP in this study) to map the relationship from cross-correlation time lags (derived from waveforms) to MEQ locations (x,y,z). Lastly, we apply the trained PMLP onto the field waveforms to obtain Newberry MEQ location predictions.

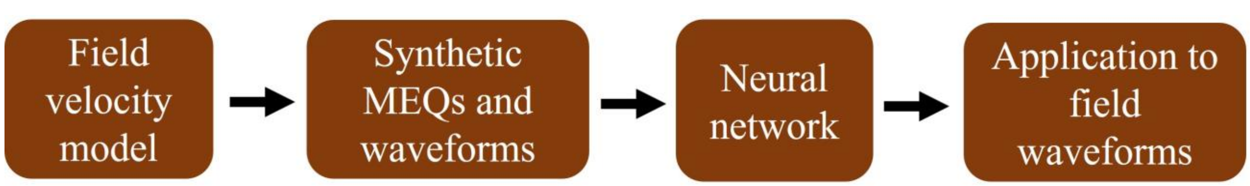

Figure 2: The workflow of this study begins with using a realistic velocity model derived from field measurements to generate numerous random MEQs. Next, we simulate the corresponding MEQ waveforms using 3D acoustic forward modeling. Following this, we extract cross-correlation time lags from these waveforms. These time lags are then utilized as inputs for our neural network, with the MEQ locations serving as outputs. After the neural network is adequately trained, we implement *transfer learning*, by applying this trained neural network to the actual field waveforms to obtain accurate location predictions.

3.1 Newberry Seismic Velocity Model

Matzel et al. (2014) computed ambient noise correlations from 22 seismic stations in the Newberry network, together with 12 additional stations from the nearby CC (Cascade Chain), UO (University of Oregon), and UW (University of Washington) seismic networks. The Green's functions that emerged from the cross-correlation waveforms were treated as seismic record and inverted for the best fitting 1D model along each path, resulting in Vp, Vs, and Qs models. For this study, we use the Vp model as a basis for our study.

The original format of velocity model is in latitude, longitude, and altitude (elevation above sea level). As such, we apply these preprocessing steps to convert the location to appropriate scales:

1) We first convert the latitude and longitude to easting and northing coordinates using the open-source software UTM (https://github.com/Turbo87/utm).
2) Next, we convert the altitude to depth below ground by subtracting altitude from the local topography.
3) Due to the significantly larger easting and northing values compared to depth, we normalize these values by subtracting them from the easting and northing coordinates

of centroid of the 15 stations. This ensures the new coordinates system is centered around the seismic stations.

4) Finally, we upsample the original velocity model from spatial sampling (dx, dy, dz) of 500 m to 25 m to satisfy seismic modeling numerical stability requirements (Igel, 2017).

Similarly, we also preprocess the locations of the field MEQ events. We overlay the velocity model with 2012 and 2014 stimulation initially located MEQ events in Figure 3. The 2012 stimulation MEQs are scattered as far as ~2 km away from the well bore, with the majority of events lying at depths of 2.0 – 3 km. These initial location estimates are incorrect (see Figures 3a and 3b) as there was a casing leak and most of the MEQs were later relocated to much shallower depths (0.6 – 1.3 km). As for the 2014 stimulation MEQs, the initial locations are noticeable at the wrong depths (Figures 3a and 3b) as the fluid injection was correctly stimulated at intended depths of ~ 1.9 – 3.0 km (Cladouhos et al., 2016). Moreover, we note that the velocity model completely covers the spatial extent of all the MEQs. This allows us to generate synthetic MEQs anywhere within the velocity model and simulate their corresponding waveforms.

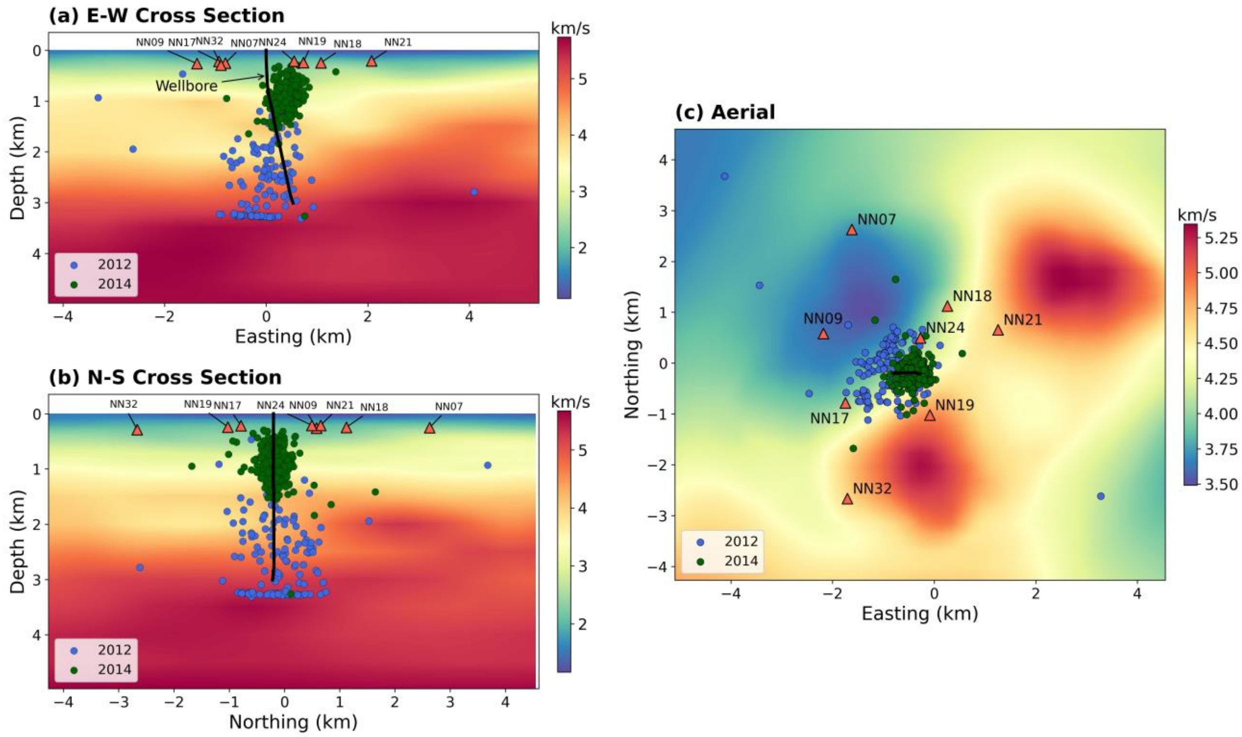

Figure 3: 3D P-wave velocity model of Newberry EGS site generated by Matzel et al. (2014). (a) represents the East-West cross section, (b) is the North-South cross section, and (c) is the aerial view of the location. Blue dots are initially located events from 2012 stimulation and green dots are from 2014 stimulation. In the downloaded raw dataset, there are 149 events for 2012 stimulation and 344 events for 2014 stimulation.

3.2 Synthetic MEQs, 3D Acoustic Waveforms, and Cross-correlation Time Lags

From the velocity model, we generate 10,000 artificial events across the entire extent of velocity model, and another 10,000 events to focus on the regions below the seismic stations which is also the injection zone (Figure 4). We note that the artificial MEQ events concentrate at the regions with field events. Next, we perform acoustic wave seismic modeling using the open-source Madagascar software (https://www.reproducibility.org/wiki/Main_Page) to generate the synthetic waveforms.

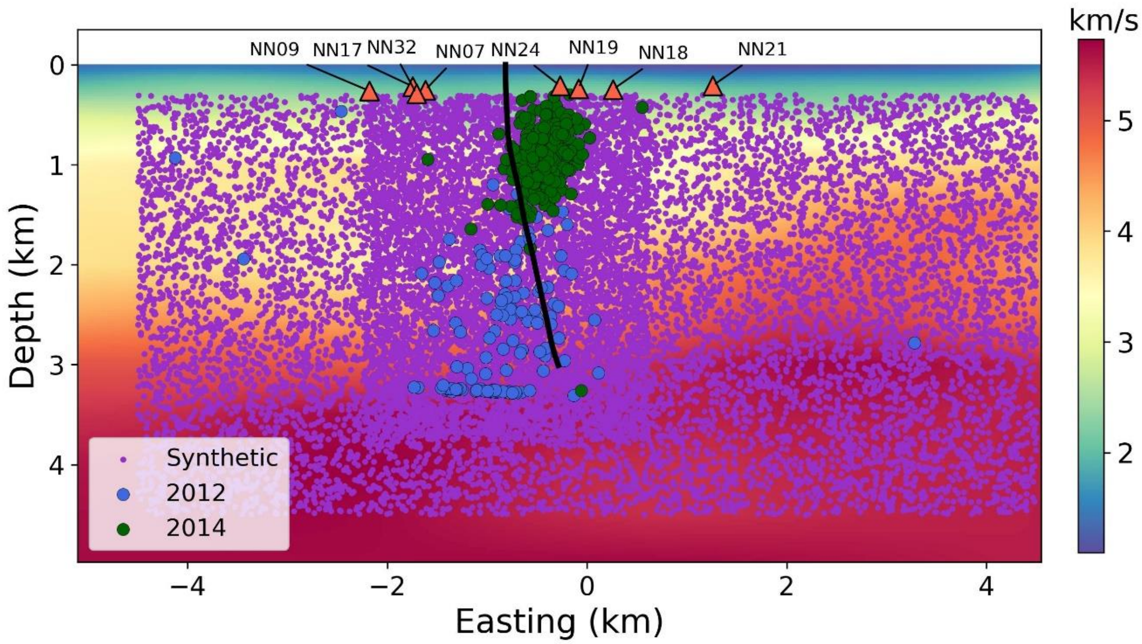

Figure 4: 10,000 synthetic events (purple) covering almost the entire spatial extent of velocity model. There are an additional 10,000 events covering the regions below the seismic stations. Blue dots are initially located events from 2012 stimulation and green dots are from 2014 stimulation.

Figure 5 shows an example snapshot of P-wave arriving at the receivers. The P-wave is analogous to the first arrivals emanated from induced MEQs during fluid stimulation. Figure 6a shows an example of waveforms generated (in black) from seismic modeling. It is important to highlight that the moveout pattern is caused by the relative MEQ location to receivers. For different MEQs at other locations, the time taken for first arrivals to arrive at the receivers cause different moveout patterns. We pick the first arrivals from the waveforms and create corresponding delta functions (red spikes in Figure 6a). Next, we use the trace at NN17 as the master trace to cross-correlate with all traces within a seismic gather. The cross-correlations aims to preserve the moveout information such that the time lag at the master trace (NN17) is zero, while the time lags at other traces correspond to the moveout pattern. Figure 6b shows the resulting cross-correlations with labeled time lags. The time lags are directly indicative of the moveout caused by the relative location of MEQ and receiver locations. The time lags are treated as the input of the neural network whereas, the location information (easting, northing, depth) is treated as the output.

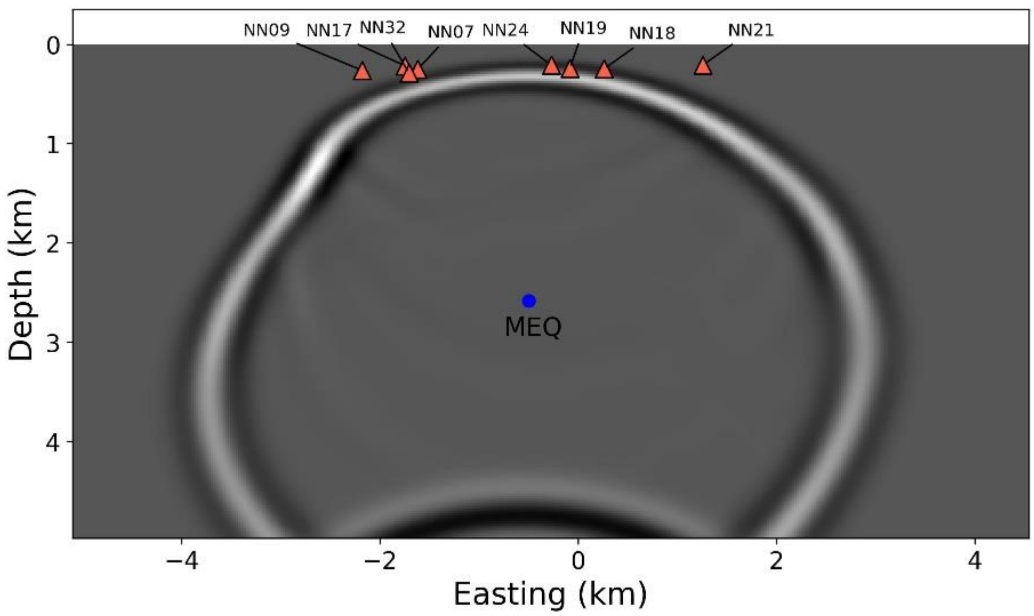

Figure 5: Example snapshot of pressure wave arriving at receivers.

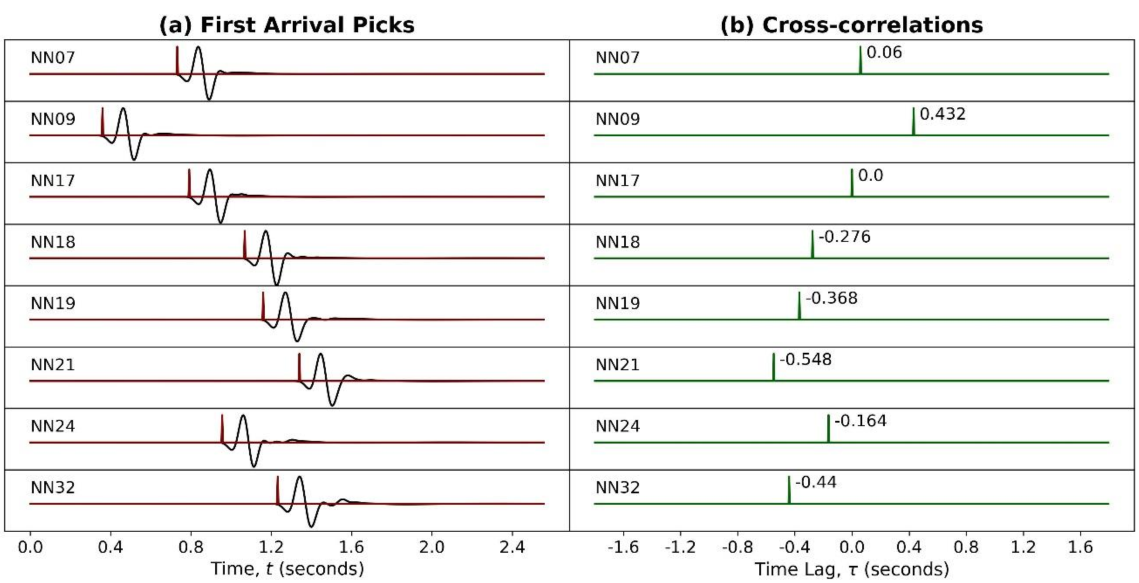

Figure 6: (a) shows an example of synthetic waveforms (black) and first arrival picks converted to delta functions (red). (b) is the corresponding cross-correlogram computed from using NN17 as master trace to cross-correlate with all traces in the seismic gather. Labeled numbers indicate time lags, which represent the moveout.

3.3 Probabilistic Multilayer Perceptron

Multilayer perceptrons (MLPs) are the fundamental building blocks of feedforward neural networks that consist of multiple layers of interconnected nodes and neurons. MLPs are also commonly referred to as artificial neural networks and deep neural networks. A simple MLP consists of an input layer, one or more hidden layers, and output layer (Figure 7). Each neuron in a layer is connected to all the neurons in the previous and next layers, with associated weights assigned to each connection. Additionally, each neuron has an activation function that determines its output based on the weighted sum of its inputs.

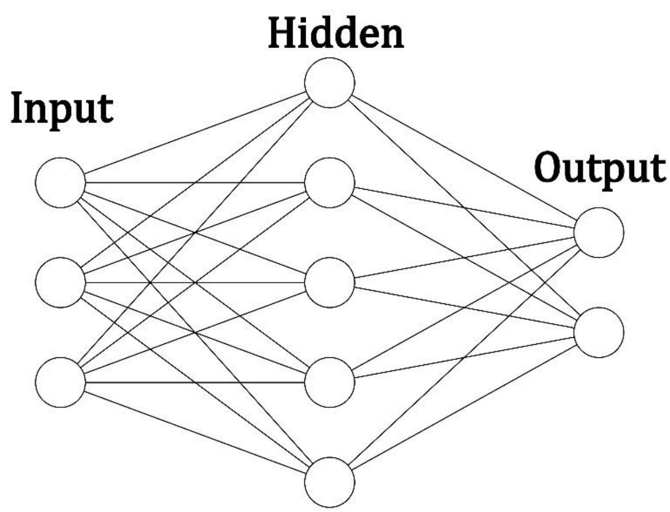

Figure 7: A simple MLP that consists of an input, hidden, and output layer. The circles in the hidden layer represent individual neurons.

To express MLPs mathematically, let X be an input vector as $X = (x_1, x_2, x_3 \dots, x_n)$. At the hidden layer, the neurons can be expressed as:

$$Z = a(W_h X + b_h),$$

where Z is the output of the hidden layer and a is the activation function, W_h and b_h are the weights and biases of hidden layer. The activation function introduces nonlinearity to the system so that the MLP can effectively learn the appropriate weights and biases to solve the nonlinear problem. Without the activation function, the system would be linear and the training of MLP would not converge. At the output layer:

$$\hat{y} = a(W_o Z + b_o),$$

where $\hat{y}$ is the MLP output (prediction). The training process of an MLP involves adjusting the weights and biases to minimize the error between the predicted output and the target output, typically using the backpropagation algorithm and gradient descent optimization (Lecun et al, 2015). Some examples of loss function include mean square error (L2 norm) and mean absolute error (L1 norm).

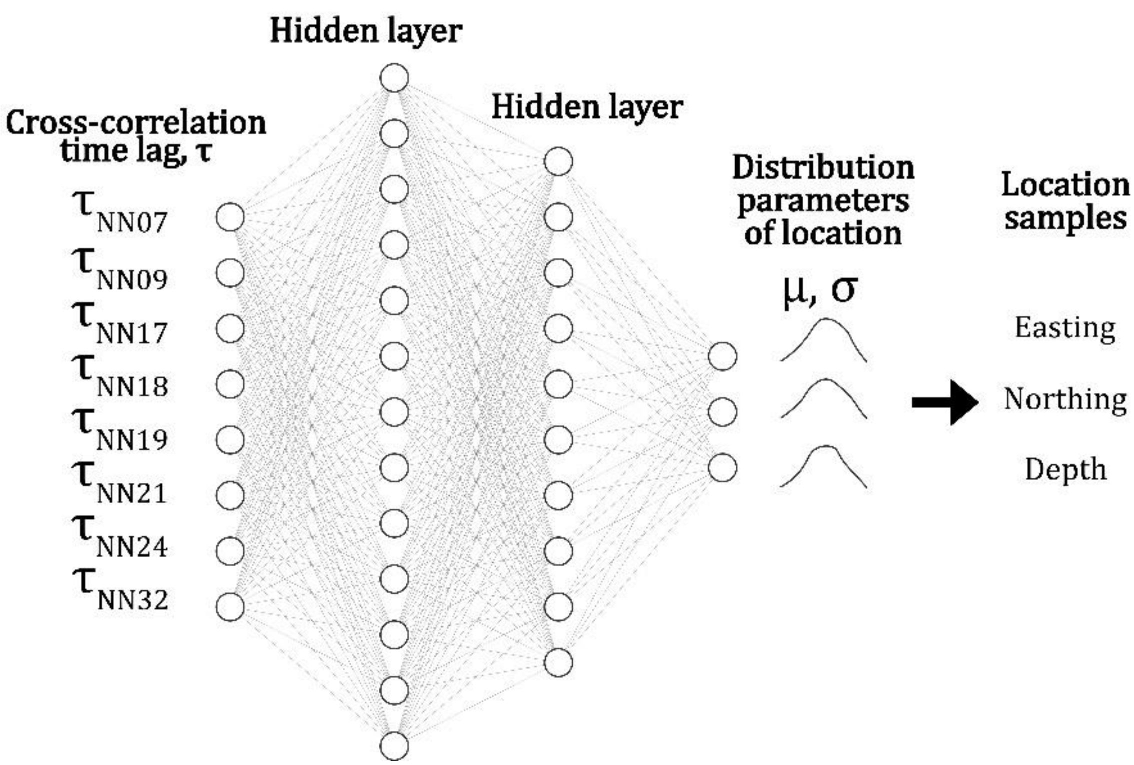

Figure 8: Probabilistic MLP (PMLP) architecture. It accepts cross-correlation time lags as input and outputs mean (μ) and standard deviation (σ) of MEQ location. μ and σ can then be used to sample from the Gaussian distribution to obtain MEQ location samples (easting, northing, depth). We compute the average location as the final location prediction.

In earthquake location prediction, uncertainties play a key role in the process of quantifying the reliability of NN predictions. Standard MLPs are deterministic, meaning they output deterministic point estimates. Here, we use the probabilistic MLP (PMLP) to predict MEQ locations from cross-correlation time lags. Figure 8 shows the architecture of PMLP. PMLP contains a preceding conventional MLP structure, however, instead of directly predicting the location of MEQs, it predicts the distribution parameters (mean and standard deviation) of MEQ locations which are assumed to follow a Gaussian distribution. Essentially, PMLP seeks to find the best distribution parameters that make the output training data (event locations) most probable. In mathematical terms, PMLP can be expressed as a general nonlinear regressor by:

$$PMLP(\tau_N) = [\mu_N, \sigma_N],$$

where τ is the cross-correlation time lags, μ and σ are mean and standard deviation across N number of input data (time lags).

To determine the set of μ and σ that can make MEQ locations most probable, we employ maximum likelihood estimation (MLE). MLE finds the parameters that maximize the likelihood of observing the MEQ locations given the PMLP regression model. In practice, it is easier to maximize the log of the likelihood, or equivalently, minimize the negative log-likelihood. The Gaussian likelihood function, L, for a single MEQ location (y), is given by:

$$L(y;\ \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(y-\mu)^2}{2\sigma^2}}$$

The negative log-likelihood simply means:

$$-\log\big(L(\tau;\ \mu,\sigma)\big) = -\log\left(\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(y-\mu)^2}{2\sigma^2}}\right)$$

By applying the logarithm properties:

$$-\log\left(\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(y-\mu)^2}{2\sigma^2}}\right) = -\log\left(\frac{1}{\sigma\sqrt{2\pi}}\right) - \log\left(e^{-\frac{(y-\mu)^2}{2\sigma^2}}\right)$$

$$= \log\big(\sigma\sqrt{2\pi}\big) + \frac{(y-\mu)^2}{2\sigma^2}$$

For optimization purposes, we can leave out the constant term $\log\big(\sqrt{2\pi}\big)$, and the resulting negative log-likelihood, NLL, (Nix and Weigend, 1994) can be defined as:

$$NLL = -\log\big(L(\tau;\ \mu,\sigma)\big) = \frac{1}{N}\sum_{i=1}^{N}\left[\log\sigma(\tau_i) + \frac{(y_i-\mu(\tau_i))^2}{2\sigma(\tau_i)^2}\right]$$

where τ_i is the cross-correlation time lags, σ is the standard deviation, y is the MEQ location values (easting, northing, depth), μ is the mean, and $i \in [1, N]$ where N is the number of training dataset. Simply put, the negative log-likelihood loss function finds the parameters (μ, σ) that best predict the MEQ locations in the training dataset.

We use ReLU as the activation function for all hidden layers. At the final layer, we only use fully-connected (dense) neurons without activation function for the easting and northing components as they contain negative and positive values. For the depth output component, we enforce a ReLU activation as the depth values are always positive.

In practice, we can apply the trained PMLP to unseen time lags, τ, to predict μ and σ. For example, for one set of time lags, the PMLP directly predicts one set of μ and σ of the MEQ location (easting, northing, depth). The predicted μ and σ are used to sample from the Gaussian distribution to obtain the realizations of predicted MEQ locations. Since this process is probabilistic, multiple sampling yields slightly different locations. This allows repeated sampling that produces a range of predictions, which we can then compute the mean as the final MEQ location prediction; and compute statistical uncertainties from the range of sampled predictions. Here, we note that the estimated uncertainties come from the trained PMLP regression model, instead of the error introduced from the input time lags. The uncertainties represent the range of values that the trained PMLP will produce.

4 Results

This section is divided into three main parts. First, we discuss the performance of PMLP on synthetic dataset, i.e., training and testing on time lags generated from synthetic events. Second, we discuss the results of applying the trained PMLP on the 2012 stimulation MEQ dataset. Third, we discuss the 2014 stimulation MEQ location estimates and interpret our results based on the location's geology.

4.1 Synthethic Tests

From the total 20,000 generated dataset, we remove a certain number of bad simulations due to edge effects, resulting in the new total to be 19,738 datasets. We randomly split 16,875 (85%) for training, 1,876 (10%) for validation and 987 (5%) for testing. To select the best trained weights, we evaluate the Euclidean distance between predictions and ground truth. The Euclidean distance, D, is calculated by:

$$D = \sqrt{\left(\hat{E}_i - E_i\right)^2 + \left(\hat{N}_i - N_i\right)^2 + \left(\hat{H}_i - H_i\right)^2}$$

where $\hat{E}, \hat{N}, \hat{H}$ are predicted easting, northing, and depth, and E, N, H are the respective ground truth. During training, we save the best weights that predict the lowest Euclidean distance on the validation dataset. Figure 9a shows the training progress in logarithms for better visualization, and the best weight is selected at epoch 236 along with the average validation loss of 42 m. During inference, the trained PMLP samples the Gaussian distribution associated with the MEQ location, thus generating slightly differing predictions for each sample. Figure 9b shows 300 samples of location predictions that are based on one input event. As expected, the samples are scattered around the mean location. For our study, we generate 3,000 samples and compute the corresponding mean as the final location prediction.

The average Euclidean distance loss on testing dataset is 41 m. We further examine the prediction errors on the testing dataset (Figure 10) and compute simple statistics tests to gauge the prediction performance. For example, the prediction errors have 90% likeliness to fall between [-53, 55] m in easting component; [-62, 59] m in northing component; and [-66, 68] m in depth component. In addition, we compute the 95% confidence interval, and the errors are approximately 10 m more on each side. In the broader context, the velocity model has dimensions of approximately 9 km x 9 km x 4 km, and PMLP's prediction errors are less than 100 m, corresponding to about a 1% error in each dimension.

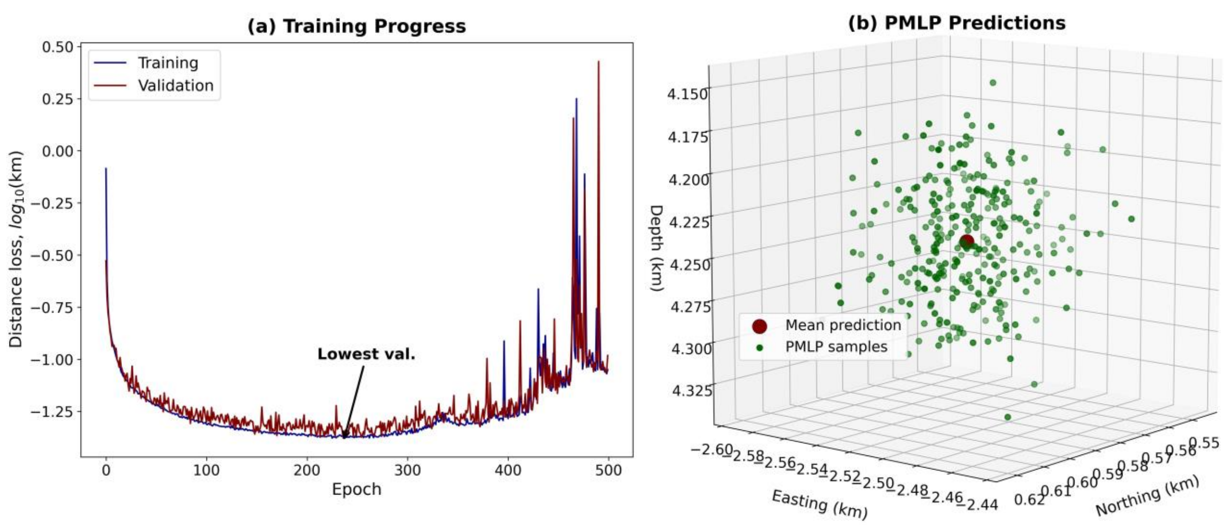

Figure 9: (a) shows the training progress of PMLP. Blue curve is the training loss and red curve is the validation loss. The loss refers to the Euclidean distance. We use the model weights at

epoch 236 as the final weights as that is when the validation loss is the lowest. (b) shows the PMLP predictions (300 samples) for one input event. The mean prediction is computed as the final prediction.

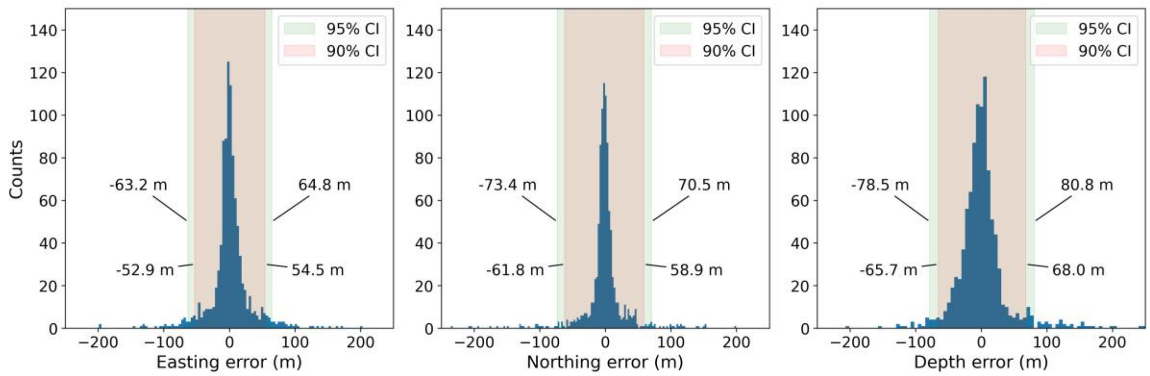

Figure 10: Prediction on testing dataset (n=987). The 95% confidence interval of the prediction error at easting is [-63.2, 64.8] m, northing is [-73.4, 70.5] m, and depth is [-78.5, 80.8] m. As for the 90% confidence interval, the prediction error at easting is [-52.9, 54.5] m, northing is [-61.8, 58.9] m, and depth is [-65.7, 68.0] m.

4.2 Field Application – 2012 Stimulation

Out of the 149 triggered waveforms, we consider 10 events to be outliers as they are out of bounds i.e., located above stations and far away from the injection zone. The PMLP model requires input from all eight borehole stations for accurate predictions. Therefore, we can only consider events that have recorded waveforms at each of these eight stations. This criterion further narrows our analysis to 113 events with waveforms in those borehole receivers. As the P-wave synthetic waveforms used in training, we only consider the first arrival picks of the vertical component in all field waveforms. We assume this is reasonable because all the induced MEQs are located below the receivers and the vertical component sensor can sufficiently pick up the first arrival waves.

Before picking the first arrivals, we apply these preprocessing steps to the field waveforms:

1) We use trim the waveforms using the same start and end time to ensure the event waveforms are aligned at the same time window.
2) We apply a bandpass of 5 – 15 Hz.
3) Lastly, we normalize the traces based on their maximum value.

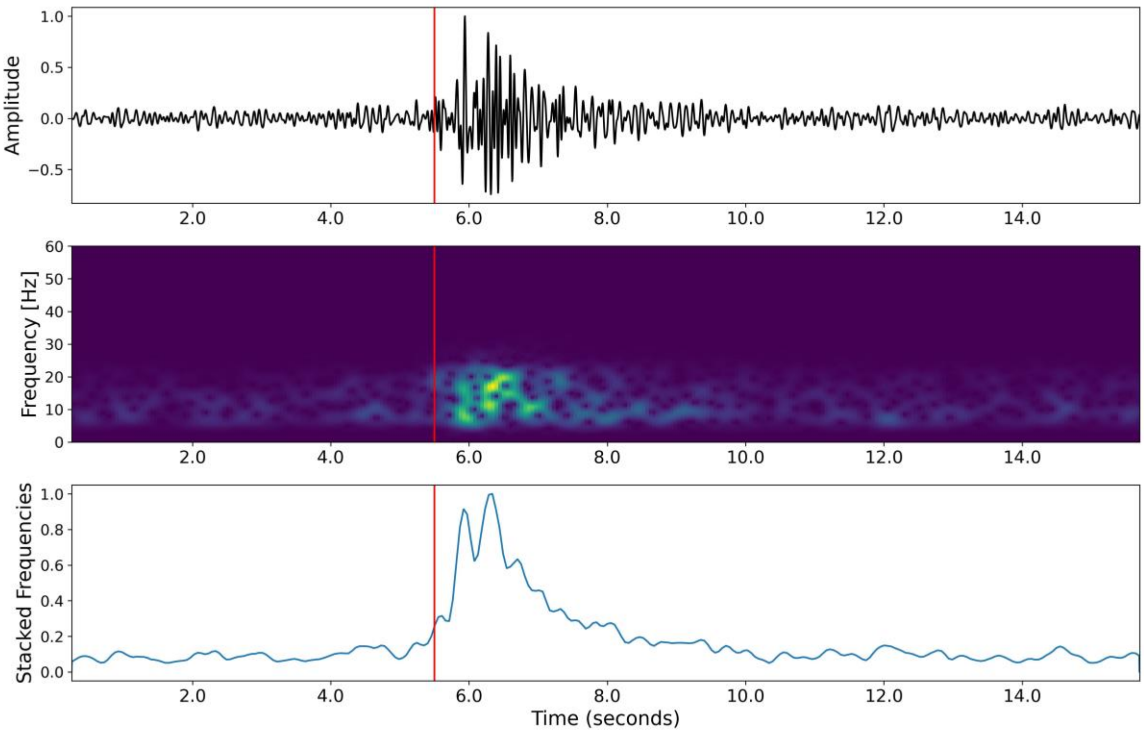

Figure 11: Top panel shows an example of seismic trace. Middle panel shows the corresponding bandpassed frequency spectra. In the bottom panel, the frequency spectra are summed up in the vertical component and normalized based on its absolute maximum value. The red line shows the handpicked first arrival for which the picking location is guided by the onset of energy as depicted in the bottom panel.

To obtain the most accurate first arrival picks, we compute the frequency spectra and stack the frequency's amplitudes to use as a guide for picking (Figure 11). The stacked frequencies illuminate the first arriving energy associated with the MEQ first arrivals. We carefully handpick the first arrivals, compute the cross-correlations and retrieve the corresponding time lags.

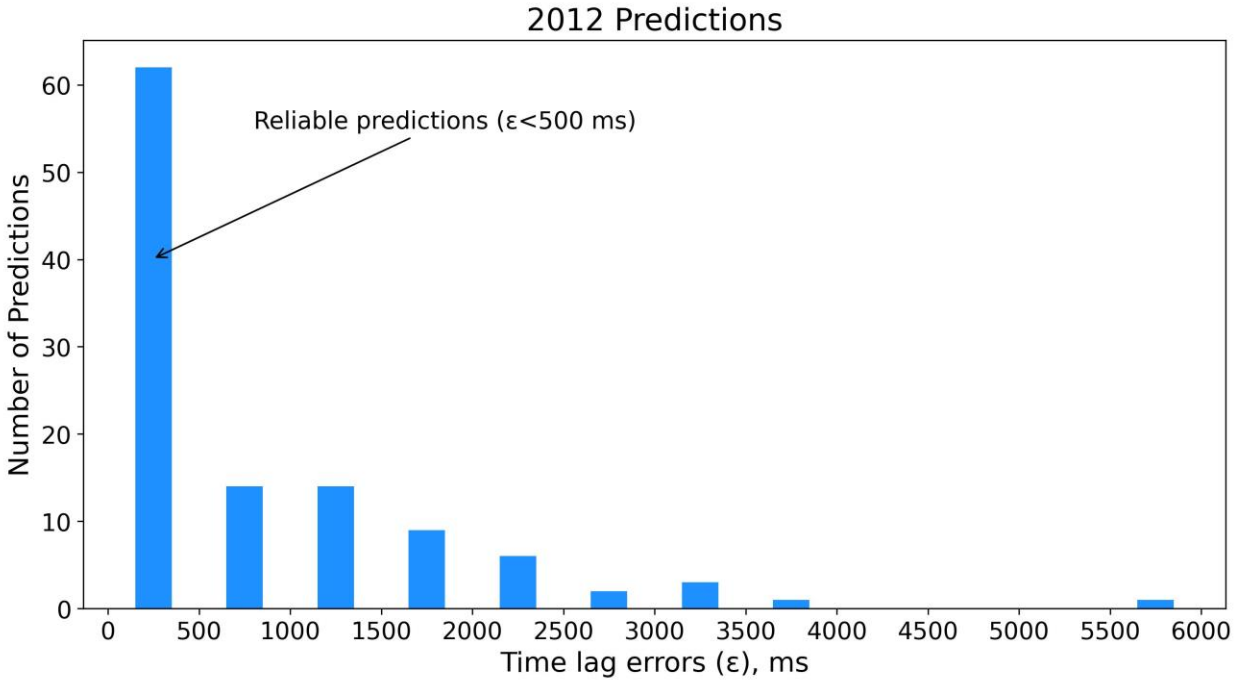

Figure 12: Histogram showing the mean-square-errors between predicted forward and field time lags for 2012 dataset. We consider the predictions falling within first bin as reliable.

Given the encouraging results observed from the application of PMLP in synthetic tests, we apply the trained PMLP model on the computed field time lags. However, preprocessing is needed due to the detection of unreliable location estimates within the raw predictions. This is evidenced by the significant discrepancies in time lag errors when comparing synthetic forward time lags with field-picked time lags. To address this, we calculate the mean-square-error for time lags (ε) between the field-picked and predicted forward time lags across all predictions, as shown in Figure 12. The histogram of ε guides our reliability criteria: predictions with ε below 500 ms are deemed reliable, which also corresponds to the most frequent histogram bin. Predicted locations in this bin have good match between the predicted forward time lags and that from field-picks. After the preprocessing step, we identify a total of 62 reliable predicted locations. Figure 13 shows the predicted locations and three examples of comparison of predicted forward first arrival picks vs. field first arrival picks.

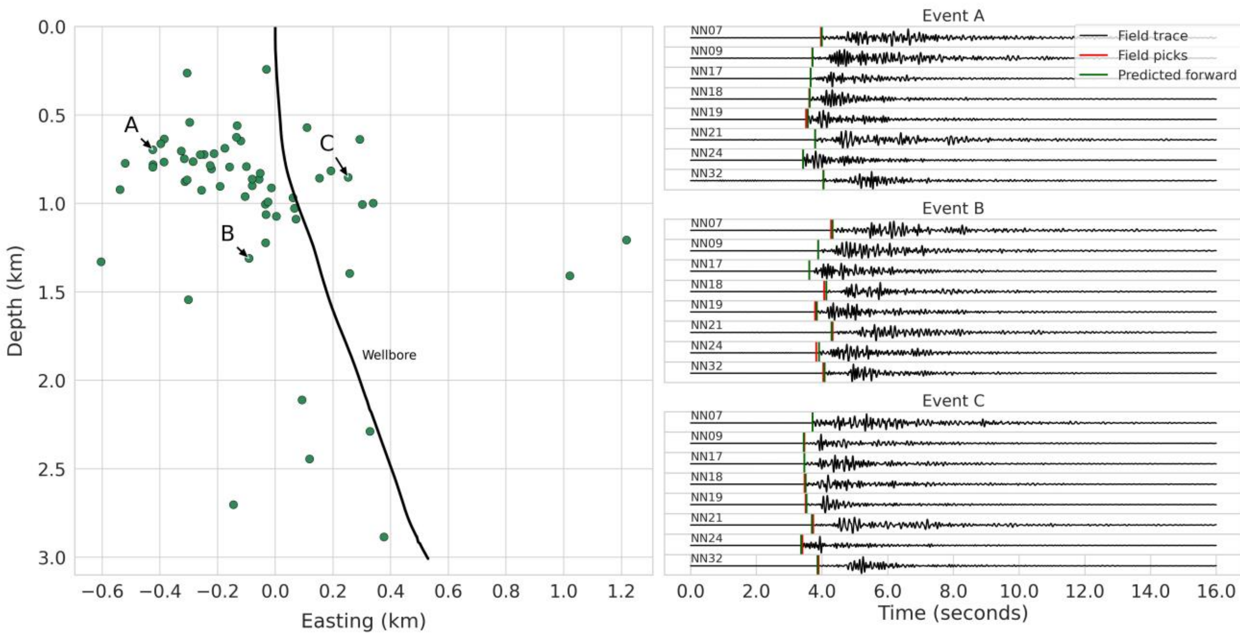

Figure 13: PMLP predicted MEQs for 2012 stimulation. The left panel shows the cross section of the predicted location of MEQs. Right panel highlights three examples (A, B, C) to show the comparison of synthetic (predicted forward) vs. field picked first arrivals.

From these predicted MEQ locations, we notice a majority concentrate at depths of 0.5 – 1.2 km. In comparison with the relocated events for 2012 stimulation done via physics-based inversion (Cladouhos et al., 2016), their predictions concentrate at depths of 0.5 – 1.3 km, which aligns with those from PMLP prediction. This directly corroborates with casing leak scenario which causes the induced microseismicity shallower than the intended depths (1.9 – 3.0 km). Events A, B, and C are three examples of predictions that display significantly good match between field (red vertical lines) and predicted forward picks (green vertical lines). The aligned depths of our predictions with those reported by Cladouhos et al. (2016), alongside the closely overlapping first arrival picks as evidenced in Events A, B, and C (in Figure 13; right panel), underscores the accuracy and reliability of PMLP in predicting microseismic locations.

4.3 Field Application – 2014 Stimulation

From the available 334 MEQ waveforms, we select 292 as the remaining events do not contain seismic traces in all eight stations. As the 2014 raw waveforms contain more noise and the original data format are less structured, it is essential to preprocess the field waveforms before picking the first arrivals. First, for each event, we find the most common start and end time within all traces because many waveforms have different start times. Second, we apply a bandpass filter of 6 – 20 Hz to remove noise of higher and lower frequencies. Third, we demean and normalize the traces so that the resulting seismograms can be picked easily.

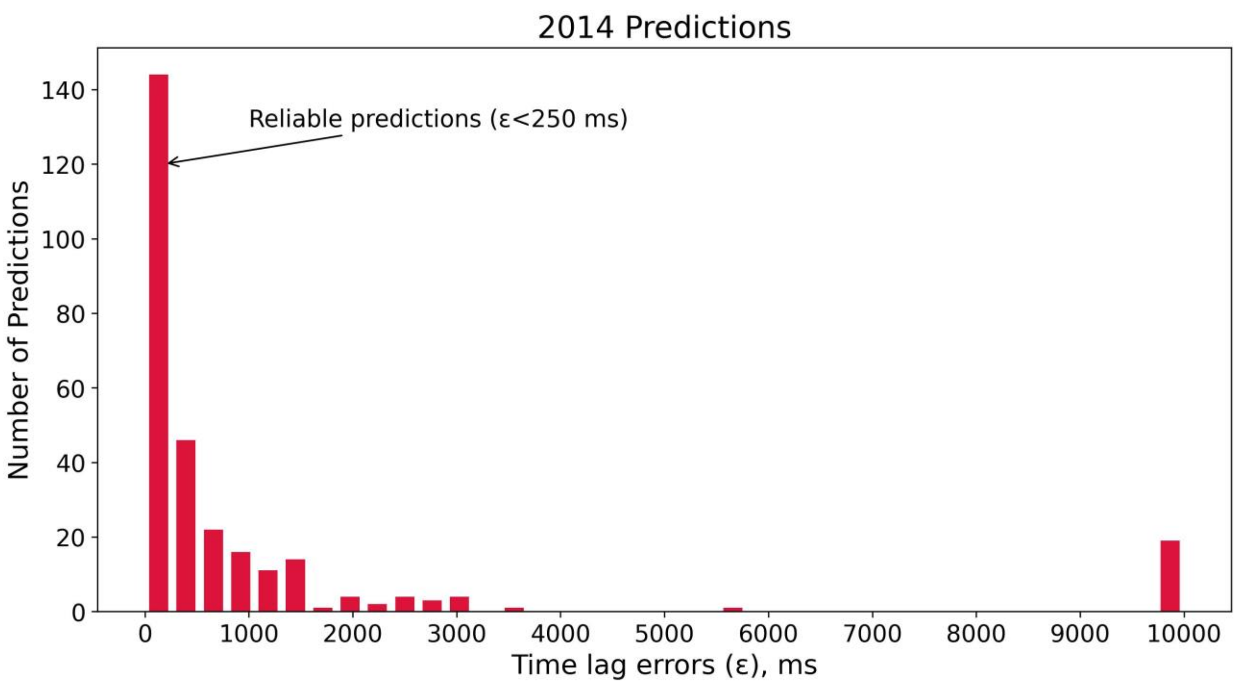

Figure 14: Similarly for 2014 dataset, we plot the histogram showing the mean-square-errors between predicted forward and field time lags. We consider the predictions falling within the first bin as reliable.

Similarly, we first apply the trained PMLP on the 2014 dataset, compute the cross-correlated time lags for each predicted location, and compare them with those from the field picks. The histogram of the errors is plotted in Figure 14. From the histogram, we see the first bin (250 ms) has the greatest number of predictions, which also means that these predictions are the most accurate due to their low error between the predicted forward and field picks. This entails a total of 142 reliable predictions (Figure S1) based on their first arrivals match. However, we notice that there are two clusters of predictions separated by a noticeable gap (lack of predictions) around 1.8 km depth. In Cladouhos et al. (2016), the physics-based inversion study did not show any location estimates above 1.8 km depth. Upon inspecting the first arrival match between the synthetic picks and field picks (Figure S2), we postulate that although the first arrival match is good, we think these predictions likely stem from incorrect first arrival picks. For instance, the waveforms in the six examples (Events A-F) contain relatively more noise and our picks may not best represent the real first arrivals. Following this, we consider it appropriate to only keep the events below 1.8 km depth (128 events) as the final predictions for interpretation (shown in Figure 15). In total, there are 128 events used as the final predictions.

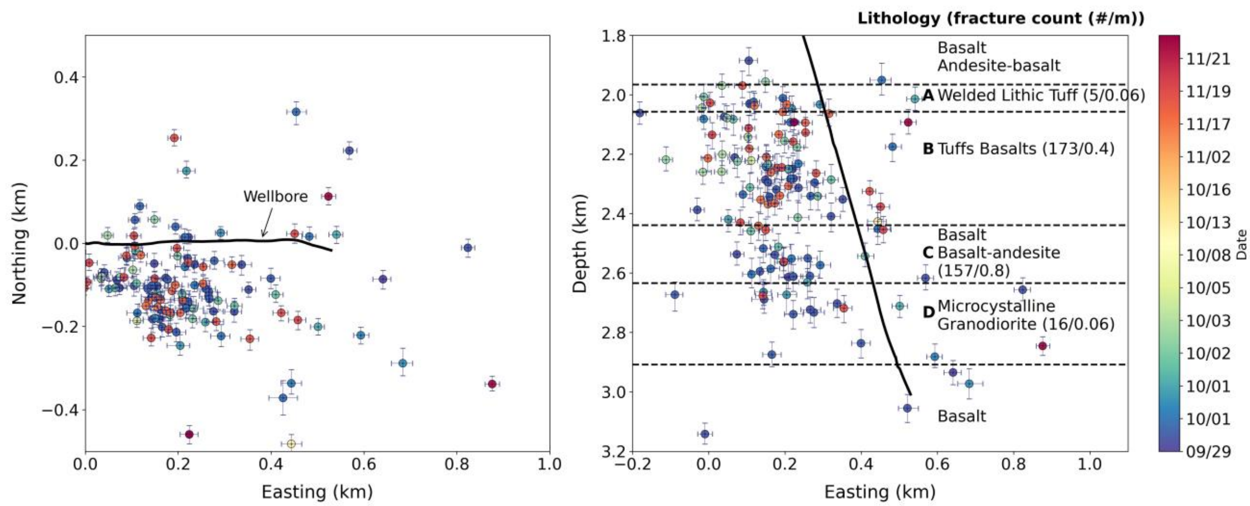

Figure 15: Overlay of 2014 stimulation location predictions with interpreted geologic zones from Cladouhos et al. (2016). The fracture count is determined by counting fractures within the NWG 55-29 borehole. The error bars are calculated by using the range of location samples predicted from repeated sampling of the trained PMLP.

We cross reference the final predictions with the subsurface geologic information determined from the NWG 55-29 borehole (Cladouhos et al., 2016) in Figure 15 (right panel). Additionally, we overlay the MEQ location predictions with the appropriate geologic zones. We re-reference the MEQ locations relative to the wellbore coordinates for better comparison. The fracture count in each zone is determined from images captured from a borehole televiewer (BHTV) survey. For example, in Zone A, there are 5 natural fractures per 0.06 m. We note that this fracture count is only representative of the region within the borehole. In general, we see that Zone B – Tuffs and Basalts (173/0.4 m) contains the most fractures, followed by Zone C (157/0.8 m), Zone D (16/0.6 m), and Zone A (5 / 0.06 m). Since natural fractures are more predisposed to rupture (Rutqvist et al., 2013; Lei et al., 2021) during fluid stimulation, it is reasonable to assume that the number of MEQs would be higher in regions that contain more natural fractures. Most of our MEQ predictions (n = 67) are in Zone B, which aligns with the zone having the highest fracture count. This correlation supports the validity of PMLP predictions in reflecting the geological conditions of the subsurface. Our confidence in these predictions is further bolstered by the excellent match of first arrival picks between synthetic and field traces. We select eight examples (see Figure S3) and compare their first arrivals with those picked from field, and the data match appears to match well (Figure S4). From the BHTV fracture counts, Zone C has the second highest number of fractures (157/0.8 m), also coincides with our predictions in terms of number of MEQ predictions. While Zones A and D show an equal number of MEQ predictions (n = 16), BHTV fracture counts reveal more fractures in Zone D (n = 16) than in Zone A (n = 5). It is important to note that these fracture counts are derived from BHTV images within the borehole and may not fully represent the entire geological strata. Additionally, fracture counts in Zones B and C are an order of magnitude higher than in Zones A and D, a pattern that is consistently echoed in our MEQ predictions.

The microseismic events predicted by the PMLP form a cloud analogous to that derived from the physics-based inversion (see Fig. 12 in Cladouhos et al. (2016)), albeit with most of our predicted events notably clustered to the west of the wellbore. The westward clustering observed in our predictions may be attributed to the DL model's reliance solely on P-wave picks, as opposed to the physics-based inversion which utilized both P- and S-wave arrivals. This methodological difference inherently results in slightly distinct MEQ location predictions. While we acknowledge that the refinement of first arrival picks could potentially enhance the model's accuracy, the strong data match between synthetic and field trace first arrivals suggests our current results are reasonable. The slight difference in the spatial distribution of predicted MEQs is counterbalanced by the significant advantage of the DL model in terms of inference speed. Once trained, our model can deliver predictions within seconds, a speed that significantly outperforms traditional source localization methods.

In general, our results indicate that most induced MEQs are located above 2.90 km depth (until Zone D), with a few outliers. In contrast, Cladouhos et al. (2016) demonstrate that the relocated MEQs extend down to 3.2 km, with a significant number of events occurring between depths of 3.05 and 3.2 km. A more in-depth study was done to interpret the natural fractures (Davatzes and Hickman, 2011), and it is found that fractures are only present at depths of 2.0 km to approximately 2.7 km. Since injection of fluid increases pore pressure within these fractures, it reduces the effective normal stress acting on the fracture walls. This reduction in normal stress decreases the frictional resistance to shear sliding along the fracture plane. As pore pressure continues to increase, it eventually overcomes the frictional resistance and causes the natural fractures to slip or rupture, a process known as shear reactivation (Das and Zoback, 2011; Rutqvist et al., 2013). Based on our findings, we can reasonably conclude that our predictions are accurate due to the presence of natural fractures matching the depths of predicted MEQs. Therefore, we postulate that our results could potentially be seen as an improvement to the method used in Cladouhos et al. (2016).

As for the uncertainties associated with the prediction, we compute the range (lower and upper bounds) of location samples generated from repeated predictions. We show the uncertainties in Figure 15 in the form of error bars. We avoid the use of standard deviation as uncertainties as we would want to know the full extent of location predictions produced from the repeated predictions. The uncertainties suggest a lesser variation and stronger confidence in the East-West component, while more variation in the depth component. This coincides with conventional earthquake location inversion methods in which the depth component typically shows larger uncertainties.

5 Picking Error Sensitivity Analysis

We perform a sensitivity analysis on the synthetic testing dataset to gauge how much the predicted locations would be affected by errors introduced in first arrival picks. We create four levels of errors in terms of number of time samples (nt): 5, 10, 15, and 20. We assume that 5-time samples would be an appropriate margin of error to characterize errors associated with handpicking. Next, we apply each error level to randomly selected receivers. Then, we compute the average Euclidean distance loss between the error-perturbed first arrivals and ground truth across the testing dataset. Table 1 shows the sensitivity of picking errors in accordance with the number of affected receivers. For example, when there are two random receivers affected with +/- 5-time samples (0.02 s) error, the average Euclidean distance loss is 0.1855 km. Our sensitivity analysis suggests that PMLP is highly sensitive to arrival picks. This is as expected because PMLP essentially only considers first-arrivals as input features. Therefore, in our study, we manually review all traces and re-pick first arrivals whenever it is necessary.

Picking errors (# nt (s))	**Number of affected receivers (selected by random)**								
	0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
5 (0.02)	0.0410	0.1320	0.1855	0.2170	0.2506	0.2873	0.3023	0.3357	0.3519
10 (0.04)	0.0410	0.2562	0.3518	0.4088	0.4893	0.5393	0.5624	0.6176	0.6770
15 (0.06)	0.0410	0.3740	0.5165	0.6118	0.7167	0.7729	0.8229	0.8903	0.9712
20 (0.08)	0.0411	0.4861	0.6623	0.8018	0.9194	0.9892	1.0799	1.1524	1.2401

Table 1: Corresponding Euclidean distance (in km) errors when first arrival picking errors are introduced. We test four picking errors: 5-, 10-, 15-, and 20-time samples. For each group, we test a variety of number of affected receivers, ranging from no (zero) receivers affected to all (eight) receivers affected.

6 Discussion

In this study, we address and overcome the issue of implementing DL methods to locate field microseismic events for EGS. In practice, the foremost challenge is the scarcity of training data due to the nature of how induced MEQs are difficult to detect (high noise levels and low magnitude), and thus resulting in limited data samples for DL training. Existing studies that use DL to locate earthquakes depend on large training samples. We overcome this challenge by utilizing a field-derived 3D P-wave velocity model to simulate synthetic acoustic waveforms from numerous artificial MEQs that encompasses the whole velocity model's spatial extent. Consequently, we train a probabilistic neural network that trains on waveform features (cross-correlation time lags) and outputs MEQ locations (easting, northing, depth). Through the principle of transfer learning, this learned knowledge can be applied to previously unseen data, such as Newberry EGS field waveforms, to accurately obtain field MEQ locations. This technique overcomes the issue of lack of field training data by essentially generating a highly realistic training dataset that contains all relevant physics, which facilitates field prediction.

Our methodology combines speed (inference under few seconds) and precision (good data match) in predicting MEQ locations, making it suitable for automated picking routines. Using 3D acoustic wave propagation physics, it can significantly improve preliminary location estimates from automatically triggered waveforms. The adaptable workflow allows for various improvements, such as using updated velocity models or more complete waveforms (elastic). Owing to its core structure as a multilayer perceptron, retraining with new data is computationally efficient. Our proposed workflow can act as a blueprint for location predictions in other EGS sites. As long as the site's velocity model can be estimated, we can simulate a large number of artificial events and their corresponding realistic synthetic waveforms. Overall, our approach offers a streamlined, effective solution for MEQ location estimation in EGS sites.

A prevailing question arises: why opt for picking over directly using full waveforms in DL-based source localization methods? Many DL-based methods, especially those using CNNs, take in entire waveforms to predict locations. CNNs extract features through sliding convolution windows, a process that largely focuses on changes in waveform amplitudes. This feature extraction method is akin to seismic phase picking, which is also similar to computer vision for identifying contrasting edges or boundaries. Despite this, low signal-to-noise ratio waveforms can mislead the feature extraction, which would ultimately introduce errors in location predictions. It is evident that manual picks offer the most precise seismic phase arrivals. While established relocation methods such as HypoDD and GlowClust (Trugman and Shearer, 2017) may exist as a straightforward option to apply those curated picks to, these methods have their limitations. In sparse seismic networks, these algorithms struggle to optimally refine earthquake locations and can face convergence issues if initial locations are far from true solutions. As outlined in our approach, we supply the PMLP with an extensive synthetic dataset for training. The combination of NNs' ability to map complex data patterns and extensive training dataset offers a robust solution in situations with sparse receiver networks. Our proposed transfer learning method basically computes a global solution of the travel-time-location problem within the confines of the given velocity model. This avoids the need for initial location estimates. Essentially, the trained neural network acts as a comprehensive lookup table, correlating travel-times with locations. While further refinement of travel-time picks can enhance predicted MEQ location accuracy, our method could offer an improvement to estimate initial MEQ locations during EGS stimulation.

In our study, we operate under the assumption that the employed 3D velocity model, while broadly representative of the subsurface layers, may not fully capture all lateral heterogeneities. Despite this limitation, the model provides a more nuanced geological context compared to the simplified 1D layered models typically used in automated picking software. In practice, implementing this 3D model on-site could enhance initial-stage predictions, offering valuable, rapid feedback for the stimulation process and thereby improving operational efficiency.

In a relevant work by Chen et al. (2022), the study used RF that takes in P-wave travel-time and location coordinates, and predicts MEQ location (x,y,z coordinates). In the study's field application example at a hydraulic fracturing site, it used similar transfer learning techniques by simulating many artificial events in a 3D layered velocity model and apply their trained RF model on field data to obtain field location predictions. The study noted that when the RF model is trained on dataset generated from 1D velocity model and applied to unseen data generated from complex 3D velocity model, the prediction accuracy suffers from significant errors, especially for deep events and events near high contrasting velocity anomalies. The study also used the Monte Carlo method to estimate its prediction uncertainties by perturbing the original velocity model and examine its prediction variations. Our proposed workflow provides several improvements and introduces new advantages. In our proposed method, we use a field-informed high-resolution 3D P-wave velocity model to simulate synthetic MEQ-waveforms dataset. The velocity model is created by inverting the Green's function derived from ambient noise correlations from multiple Newberry seismic stations. As 3D velocity models contain more physics than that of 1D models, the forward simulated waveforms are more realistic. Furthermore, our proposed workflow simplifies the uncertainty estimation. The probabilistic design of the PMLP provides slightly varied outcomes for predicted samples, allowing for uncertainty quantification. Therefore, we eliminate the need for perturbing the velocity model and re-running forward modeling, a feature that could be beneficial for near real-time EGS monitoring.

There are a few improvements that can be made to our workflow. Acoustic wave modeling only generates the first arriving P-waves, and as such, our neural network is limited to P-wave features (cross-correlation time lags). In future, elastic modeling routines can be considered as they simulate S-wave components. Since neural networks work best with more features to train, having both P- and S-wave features is likely to improve the location prediction. This is also especially useful in noisy traces where the P-wave first arrival is not obvious, but the S-wave is due to its larger amplitude.

Moreover, as demonstrated in the sensitivity analysis, our workflow relies largely on the accuracy of first arrival picks. We attempt to overcome this issue by manually examining all the traces and repick the incorrect first arrival travel-times. However, we acknowledge that there may be picking inaccuracies especially when the traces are very noisy. We considered the use of state-of-the-art DL-based phase pickers such as PhaseNet and EQTransformer, but upon testing, we found that those pickers were not reliable due to inconsistent P-wave picks. In the future, we believe that developing DL-based phase pickers specifically tailored for induced MEQ waveforms could significantly improve P-wave picking. These specialized DL-based pickers could accurately identify first arrival picks. When integrated with our location estimation method, this could create a comprehensive DL suite for automated waveform picking and location estimation.

Understanding the limitations of how many synthetic MEQs can influence the inversion results can be beneficial for long-term EGS monitoring. Ideally, the more data the better it is for DL generalization at the EGS site, however, the abundance of synthetic MEQ dataset comes at cost of long simulation and DL training times. As such, it would be helpful to examine the boundaries of training data size that are needed to produce similar inversion results when compared to large number of datasets.

7 Conclusions

Locating microearthquakes (MEQs) in enhanced geothermal systems (EGS) is fundamentally challenging due to the nonlinear relationship between waveforms and location. While deep learning (DL) methodologies have shown proficiency in predicting natural earthquake locations using waveform data, the majority of these studies depend heavily on large amount of field data for training. As such, the scarcity of field training data in EGS presents considerable challenges for the implementation of DL-based approaches. To overcome the scarcity of training data, we leverage transfer learning principles by introducing a practical workflow that accurately predicts MEQ locations from cross-correlation time lags using probabilistic multilayer perceptrons (PMLP). Our results on the 2012 Newberry EGS stimulation show major microseismic activity at depths of 0.5 – 1.2 km, which agrees with the casing leakage scenario that took place in the well bore. We further apply our methodology to the 2014 stimulation data and found that most of the MEQs concentrate at 2.0 – 2.9 km depths. This finding aligns with the presence of natural fractures which extend from 2.0 – 2.7 km. Excellent data (time lags) match indicate the completeness of inversion and suggest predictions can be trusted. The combination of good data match and the predisposition of natural fractures having ruptures caused by fluid stimulation lead us to conclude that the majority of microseismic activity happens shallower than 3.0 km. Training with prior information specific to an EGS site holds promise for enabling real-time monitoring in such environments.

Acknowledgments

We extend our foremost acknowledgement to the Department of Energy (DOE) for their support through grant DE-EE0008763. We are especially thankful to Dr. Dennise Templeton and Dr. Eric M. Matzel for providing the Newberry velocity model. Our appreciation also goes to Prof. Derek Elsworth for providing the Newberry MEQ dataset. Additionally, we are grateful to Dr. Chao Guo for his insightful perspectives on our early MEQ data analysis.

Data and Software Availability

Newberry dataset was originally downloaded from https://fracture.lbl.gov/Newberry/Location.txt in October 2019. However, the link is now defunct, and the dataset is moved to https://gdr.openei.org/home. The Madagascar 3D acoustic seismic modeling codes are available at https://www.reproducibility.org/wiki/Main_Page. Machine learning codes associated with this research will be publicly available at https://github.com/zxleong in the near future.

References

AltaRock. (2014). Phase 2.1 Report, Newberry EGS Demonstration. https://gdr.openei.org/files/774/Phase%202.1%20Report_4.24.14.pdf

Bondár, I., Myers, S. C., & Engdahl, E. R. (2014). Earthquake location. In M. Beer, I. A. Kougioumtzoglou, E. Patelli, & I. S.-K. Au (Eds.), Encyclopedia of Earthquake Engineering (pp. 1–18). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-36197-5_184-1

Chang, W., & McMechan, G. A. (1994). 3-D elastic prestack, reverse-time depth migration. GEOPHYSICS, 59(4), 597–609. https://doi.org/10.1190/1.1443620

Chen, Y., Saad, O. M., Savvaidis, A., Chen, Y., & Fomel, S. (2022). 3d microseismic monitoring using machine learning. Journal of Geophysical Research: Solid Earth, 127(3), e2021JB023842. https://doi.org/10.1029/2021JB023842

Cladouhos, T. T., Petty, S., Swyer, M. W., Uddenberg, M. E., Grasso, K., & Nordin, Y. (2016). Results from newberry volcano egs demonstration, 2010–2014. Geothermics, 63, 44–61. https://doi.org/10.1016/j.geothermics.2015.08.009

Cladouhos, T.T., Petty, S., Callahan, O., Osborn, W., Hickman, S., Davatzes, N. (2011a). The role of stress modeling in stimulation planning at the Newberry VolcanoEGS Demonstration project. In: Proceedings: Thirty-Sixth Workshop onGeothermal Reservoir Engineering, Stanford University, Stanford, California,January 31-February 2, pp. 630–637, SGP-TR-1191.

Cladouhos, T.T., Clyne, M., Nichols, M., Petty, S., Osborn, W.L., Nofziger, L. (2011b). Newberry Volcano EGS Demonstration stimulation modeling. GRC Trans. 35,317–322.

Das, I., & Zoback, M. D. (2011). Long-period, long-duration seismic events during hydraulic fracture stimulation of a shale gas reservoir. The Leading Edge, 30(7), 778–786. https://doi.org/10.1190/1.3609093

Davatzes, N.C., Hickman, S.H., (2011). Preliminary analysis of stress in the NewberryEGS well NWG 55-29. GRC Trans. 35, 323–332.

Dokht, R. M. H., Kao, H., Visser, R., & Smith, B. (2019). Seismic event and phase detection using time–frequency representation and convolutional neural networks. Seismological Research Letters, 90(2A), 481–490. https://doi.org/10.1785/0220180308

EIA. (2023). Electric power monthly—U. S. Energy information administration. Retrieved from https://www.eia.gov/electricity/monthly/epm_table_grapher.php

Fang, Y., Den Hartog, S. A. M., Elsworth, D., Marone, C., & Cladouhos, T. (2016). Anomalous distribution of microearthquakes in the newberry geothermal reservoir: Mechanisms and implications. Geothermics, 63, 62–73. https://doi.org/10.1016/j.geothermics.2015.04.005

Gajewski, D., & Tessmer, E. (2005). Reverse modelling for seismic event characterization. Geophysical Journal International, 163(1), 276–284. https://doi.org/10.1111/j.1365-246X.2005.02732.x

Geiger, L., 1912, Probability method for the determination of earthquake epicenters from the arrival time only, St. Louis Univ. Bull. 8, 60–71.

Häring, M. O., Schanz, U., Ladner, F., & Dyer, B. C. (2008). Characterisation of the Basel 1 enhanced geothermal system. Geothermics, 37(5), 469–495. https://doi.org/10.1016/j.geothermics.2008.06.002

Izadi, G., & Elsworth, D. (2013). The effects of thermal stress and fluid pressure on induced seismicity during stimulation to production within fractured reservoirs. Terra Nova, 25(5), 374–380. https://doi.org/10.1111/ter.12046

Karasözen, E., & Karasözen, B. (2020). Earthquake location methods. GEM - International Journal on Geomathematics, 11(1), 13. https://doi.org/10.1007/s13137-020-00149-9

Kriegerowski, M., Petersen, G. M., Vasyura-Bathke, H., & Ohrnberger, M. (2019). A deep convolutional neural network for localization of clustered earthquakes based on multistation full waveforms. Seismological Research Letters, 90(2A), 510–516. https://doi.org/10.1785/0220180320

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521(7553), 436–444. https://doi.org/10.1038/nature14539

Lei, Q., Gholizadeh Doonechaly, N., & Tsang, C.-F. (2021). Modelling fluid injection-induced fracture activation, damage growth, seismicity occurrence and connectivity change in naturally fractured rocks. International Journal of Rock Mechanics and Mining Sciences, 138, 104598. https://doi.org/10.1016/j.ijrmms.2020.104598

Li, L., Tan, J., Schwarz, B., Staněk, F., Poiata, N., Shi, P., Diekmann, L., Eisner, L., & Gajewski, D. (2020). Recent advances and challenges of waveform-based seismic location methods at multiple scales. Reviews of Geophysics, 58(1), e2019RG000667. https://doi.org/10.1029/2019RG000667

Lu, S.-M. (2018). A global review of enhanced geothermal system (Egs). Renewable and Sustainable Energy Reviews, 81, 2902–2921. https://doi.org/10.1016/j.rser.2017.06.097

Majer, E. L., Baria, R., Stark, M., Oates, S., Bommer, J., Smith, B., & Asanuma, H. (2007). Induced seismicity associated with Enhanced Geothermal Systems. Geothermics, 36(3), 185–222. https://doi.org/10.1016/j.geothermics.2007.03.003

Matzel, E., Templeton, D., Petersson, A., Goebel, M. (2014). Imaging the Newberry EGS site using seismic interferometry. Thirty-Ninth Workshop on Geothermal Reservoir Engineering, SGP-TR-202.

McMechan, G. A., Clayton, R. W., & Mooney, W. D. (1982). Application of wave field continuation to the inversion of refraction data. Journal of Geophysical Research: Solid Earth, 87(B2), 927–935. https://doi.org/10.1029/JB087iB02p00927

Mousavi, S. M., & Beroza, G. C. (2020). Bayesian-deep-learning estimation of earthquake location from single-station observations. IEEE Transactions on Geoscience and Remote Sensing, 58(11), 8211–8224. https://doi.org/10.1109/TGRS.2020.2988770

Mousavi, S. M., Ellsworth, W. L., Zhu, W., Chuang, L. Y., & Beroza, G. C. (2020). Earthquake transformer—An attentive deep-learning model for simultaneous earthquake detection and phase picking. Nature Communications, 11(1), 3952. https://doi.org/10.1038/s41467-020-17591-w

Muffler, P., & Cataldi, R. (1978). Methods for regional assessment of geothermal resources. Geothermics, 7(2–4), 53–89. https://doi.org/10.1016/0375-6505(78)90002-0

Nix, D. A., & Weigend, A. S. (1994). Estimating the mean and variance of the target probability distribution. Proceedings of 1994 IEEE International Conference on Neural Networks (ICNN'94), 55–60 vol.1. https://doi.org/10.1109/ICNN.1994.374138

Perol, T., Gharbi, M., & Denolle, M. (2018). Convolutional neural network for earthquake detection and location. Science Advances, 4(2), e1700578. https://doi.org/10.1126/sciadv.1700578

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

Ross, Z. E., Yue, Y., Meier, M., Hauksson, E., & Heaton, T. H. (2019). Phaselink: A deep learning approach to seismic phase association. Journal of Geophysical Research: Solid Earth, 124(1), 856–869. https://doi.org/10.1029/2018JB016674

Rutqvist, J., Rinaldi, A. P., Cappa, F., & Moridis, G. J. (2013). Modeling of fault reactivation and induced seismicity during hydraulic fracturing of shale-gas reservoirs. Journal of Petroleum Science and Engineering, 107, 31–44. https://doi.org/10.1016/j.petrol.2013.04.023

Schill, E., Genter, A., Cuenot, N., & Kohl, T. (2017). Hydraulic performance history at the Soultz EGS reservoirs from stimulation and long-term circulation tests. Geothermics, 70, 110–124. https://doi.org/10.1016/j.geothermics.2017.06.003

Shen, H., & Shen, Y. (2021). Array-based convolutional neural networks for automatic detection and 4d localization of earthquakes in hawai‘i. Seismological Research Letters, 92(5), 2961–2971. https://doi.org/10.1785/0220200419

Tarantola, A., & Valette, B. (1981). Inverse problems = Quest for information. Journal of Geophysics, 50(1), 159-170. Retrieved from https://journal.geophysicsjournal.com/JofG/article/view/28.

Templeton, D. C., Wang, J., Goebel, M. K., Harris, D. B., & Cladouhos, T. T. (2020). Induced seismicity during the 2012 Newberry EGS stimulation: Assessment of two advanced earthquake detection techniques at an EGS site. Geothermics, 83, 101720. https://doi.org/10.1016/j.geothermics.2019.101720

Tester, J. W., Anderson, B. J., Batchelor, A. S., Blackwell, D. D., DiPippo, R., Drake, E. M., Garnish, J., Livesay, B., Moore, M. C., Nichols, K., Petty, S., Nafi Toksoz, M., Veatch, R. W., Baria, R., Augustine, C., Murphy, E., Negraru, P., & Richards, M. (2007). Impact of enhanced geothermal systems on US energy supply in the twenty-first century. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 365(1853), 1057–1094. https://doi.org/10.1098/rsta.2006.1964

Tomac, I., & Sauter, M. (2018). A review on challenges in the assessment of geomechanical rock performance for deep geothermal reservoir development. Renewable and Sustainable Energy Reviews, 82, 3972–3980. https://doi.org/10.1016/j.rser.2017.10.076

Trugman, D. T., & Shearer, P. M. (2017). Growclust: A hierarchical clustering algorithm for relative earthquake relocation, with application to the spanish springs and sheldon, nevada, earthquake sequences. Seismological Research Letters, 88(2A), 379–391. https://doi.org/10.1785/0220160188

Van den Ende, M. P. A., & Ampuero, J. -P. (2020). Automated seismic source characterization using deep graph neural networks. Geophysical Research Letters, 47(17). https://doi.org/10.1029/2020GL088690

Waldhauser, F., & Ellsworth, W. (2000). A double-difference earthquake location algorithm: Method and application to the northern hayward fault, california. Bulletin of the Seismological Society of America, 90(6), 1353–1368. https://doi.org/10.1785/0120000006

Zang, A., Oye, V., Jousset, P., Deichmann, N., Gritto, R., McGarr, A., Majer, E., & Bruhn, D. (2014). Analysis of induced seismicity in geothermal reservoirs – An overview. Geothermics, 52, 6–21. https://doi.org/10.1016/j.geothermics.2014.06.005

Zhang, X., Zhang, M., & Tian, X. (2021). Real-time earthquake early warning with deep learning: Application to the 2016 m 6. 0 Central Apennines, Italy earthquake. Geophysical Research Letters, 48(5). https://doi.org/10.1029/2020GL089394

Zhu, T. (2014). Time-reverse modelling of acoustic wave propagation in attenuating media. Geophysical Journal International, 197(1), 483–494. https://doi.org/10.1093/gji/ggt519

Zhu, W., & Beroza, G. C. (2018). Phasenet: A deep-neural-network-based seismic arrival time picking method. Geophysical Journal International. https://doi.org/10.1093/gji/ggy423

Zhu, W., McBrearty, I. W., Mousavi, S. M., Ellsworth, W. L., & Beroza, G. C. (2022). Earthquake phase association using a bayesian gaussian mixture model. Journal of Geophysical Research: Solid Earth, 127(5). https://doi.org/10.1029/2021JB023249

Microseismic Monitoring using Transfer Learning: Example from the Newberry EGS

Zi Xian Leong[1,†] and Tieyuan Zhu[1,2]

[1] Department of Geosciences, The Pennsylvania State University, University Park, PA, USA.

[2] EMS Energy Institute, The Pennsylvania State University, University Park, PA, USA.

†Currently at Chevron Technical Center, a division of Chevron U.S.A. Inc.

Corresponding author: Zi Xian Leong (zxnleong@gmail.com)

Key Points:

- We present a novel transfer learning workflow to predict microearthquake locations in EGS, addressing data scarcity for training
- Application to Newberry EGS reveals accurate microearthquake locations, validated against known geological features
- Employs probabilistic multilayer perceptrons that map cross-correlation time lags to microearthquake locations

Abstract

Enhanced geothermal systems (EGS) are promising for generating clean power by extracting heat energy from injection and extraction of water in geothermal reservoirs. The stimulation process involves hydroshearing which reactivates pre-existing cracks for creating permeability and meanwhile inducing microearthquakes. Locating these microearthquakes provide reliable feedback on the stimulation progress, but it poses a challenging nonlinear inverse problem. Current deep learning methods for locating earthquakes require extensive datasets for training, which is problematic as detected microearthquakes are often limited. To address the scarcity of training data, we propose a transfer learning workflow using probabilistic multilayer perceptron (PMLP) which predicts microearthquake locations from cross-correlation time lags in waveforms. Utilizing a 3D velocity model of Newberry site derived from ambient noise interferometry, we generate numerous synthetic microearthquakes and 3D acoustic waveforms for PMLP training. Accurate synthetic tests prompt us to apply the trained network to the 2012 and 2014 stimulation field waveforms. Predictions on the 2012 stimulation dataset show major microseismic activity at depths of 0.5–1.2 km, correlating with a known casing leakage scenario. In the 2014 dataset, the majority of predictions concentrate at 2.0–2.9 km depths, consistent with results obtained from conventional physics-based inversion, and align with the presence of natural fractures from 2.0–2.7 km. We validate our findings by comparing the synthetic and field picks, demonstrating a satisfactory match for the first arrivals. By combining the benefits of quick inference speeds and accurate location predictions, we demonstrate the feasibility of using transfer learning to locate microseismicity for EGS monitoring.

Plain Language Summary

Enhanced geothermal systems (EGS) are an emerging technology that generates clean electricity by injecting water into underground hot rocks and pumping it back to the surface for power generation. However, this stimulation process causes tiny earthquakes, known as microearthquakes. Tracking the location of these microearthquakes is crucial for monitoring the EGS creation process. Unfortunately, finding where these microearthquakes occur is a complex task. Using deep learning methods is challenging because of the general lack of microearthquakes for training. To overcome this, we employ transfer learning, which allows computer models to train on realistic data, and eventually deploy to real-world EGS microearthquake data. We create a realistic geological model of the Newberry EGS site and generate many artificial microearthquake data for deep learning training. During the application on field data from 2012 and 2014 stimulation, the computer model successfully identifies the depth and location of MEQs. Our results match well with what we already know about the underground structure, such as the presence of natural fractures in the rock. This study shows that our approach can effectively predict microearthquake locations even when presented with limited earthquake data for training, which is promising for monitoring and improving EGS operations in the future.

1 Introduction

Geothermal energy has emerged as a valuable and sustainable resource in the global energy landscape, which harnesses the Earth's natural heat to generate electricity, providing a reliable and consistent supply, unlike intermittent resources such as solar or wind power (Muffler & Cataldi, 1978). As a low-emission energy form, geothermal power mitigates greenhouse gas emissions and reduces the nation's reliance on fossil fuels (Tester et al., 2006). According to the U.S. Energy Information Administration (EIA, 2023), geothermal energy generation in 2022 reached approximately 17 billion kWh, positioning the United States as the leading global producer of geothermal electricity. Moreover, electricity generated from geothermal plants is projected to increase to 37.2 billion kWh in 2050. Enhanced geothermal systems (EGS) technology harnesses heat energy produced from areas of young tectonism and volcanism, but contains relatively low permeability (e.g., Häring et al., 2008; Cladouhos et al., 2016; Schill et al., 2017; Lu, 2018; Tomac and Sauter, 2018). In an EGS, fluid is injected into the subsurface under carefully controlled conditions, which caused pre-existing fractures to reopen, enhancing permeability. Increased permeability allows fluids to circulate in the now-fractured rock and to transport heat to the surface where electricity can be generated.

The creation of EGS has been widely known to induce microearthquakes (MEQs) (Zang et al., 2014; Majer et al., 2007). These MEQs, serving as reservoir stimulation diagnostic indicators, can locate fluid-induced fractures and monitor EGS stimulation progress such as crack propagation, permeability evolution, and temperature changes (Izadi and Elsworth, 2013; Fang et al., 2016). However, elevated occurrence of MEQs may lead to negative public perception regarding EGS deployment, particularly felt seismicity may be perceived as an isolated annoyance. Furthermore, there is concern about the cumulative effects of recurrent events and the potential of larger earthquakes in the future (Majer et al., 2007). Consequently, there is both a scientific and societal need to locate and monitor MEQs associated with EGS reservoir stimulation.

Conventional earthquake location methods involve iteratively minimizing the difference between picked P- and/or S-wave first arrival times and predicted data at multiple seismic stations (Geiger, 1912; Tarantola & Valette, 1982; Bondar et al., 2014; Karasözen & Karasözen, 2020). While these methods have been widely employed in seismology, they exhibit certain limitations. The accuracy of earthquake location estimates can be affected by convergence issues, particularly when the initial location guess is not sufficiently close to the true hypocenter, the solution may converge at a local minimum, leading to inaccurate location estimates. Additionally, conventional methods can be computationally intensive, particularly when applied to large datasets or in regions with complex geology (local heterogeneities). As such, most location algorithms rely on one-dimensional (1D) velocity models, where the velocity changes only with depth. Furthermore, waveform-based methods that are based on time-reversal imaging principles utilize finite difference to compute time-reversed seismograms and the actual source location is determined by identifying the point of highest energy concentration (e.g., McMechan, 1982; Chang and McMechan, 1994; Gajewski and Tessmer, 2005; Zhu, 2014; Li et al., 2020). Wavefield simulation method is unsurprisingly computationally expensive, and the energy focusing can be ambiguous for noisy data and very heterogeneous models. Waldhauser and Ellsworth (2000) proposed hypoDD, a widely used location inversion method that iteratively minimizes the misfit between theoretical and observed differential travel-times for pairs of earthquakes (double-difference) at each station. Nonetheless, the system can get very large if all event pairs are used in double-difference methods and reducing the efficiency of location estimation.

Deep learning (DL) techniques have been increasingly applied in earthquake seismology. For example, DL has seen significant developments in earthquake event phase detection (Ross et al., 2018; Dokht et al., 2019), phase picking (Zhu and Beroza; 2018; Mousavi et al., 2020), and phase association (Ross et al., 2019; Zhu et al., 2022). For DL-based earthquake location inversion, a large majority of studies rely heavily on training with labeled field data. Perol et al. (2018) used convolutional neural network (CNN) that trained on ~2,900 single station events near Guthrie, Oklahoma, in which the CNN accepts three-component waveforms and predicts earthquake location groups of six clusters. Later studies improved the earthquake location inversion method by employing more advanced DL algorithms and utilize multi-station three-component waveforms as input to predict three-dimensional (3D) locations. For example, Kriegerowski et al. (2019) employed deep CNN to predict easting, northing, and depth of earthquakes based on ~3,000 events from Western Bohemia, Czech Republic. Van den Ende and Ampuero (2020) used graph neural networks to predict the source latitude, longitude, depth, and magnitude based on ~1,300 events from Southern California. Shen and Shen (2021) used deep CNNs that trained on USGS Combined Cataglog earthquakes (~1,800) to predict latitude, longitude, depth, and origin time of events. Zhang et al. (2021) adapted deep CNNs to predict 3D event location probabilities based on ~1,000 events from Central Apennies, Italy. Using single-station waveforms, Mousavi and Beroza (2020) employed Bayesian neural networks to predict epicenter, depth, and origin time based on the Stanford Earthquake Data Set (~450k events).

Comparing natural earthquakes to geothermal induced MEQs reveals several distinct differences, particularly in terms of their detectability (Fang et al., 2016; Templeton et al., 2020). MEQs are generally characterized by lower magnitudes and higher scarcity compared to natural earthquakes. The lower magnitudes make MEQs more challenging to detect, as they are often masked by background noise. This results in fewer MEQ events detected in conventional catalogs. This scarcity of MEQs poses a significant challenge for DL training, as the limited amount of available data restricts the ability to build robust and accurate DL models for solving the nonlinear MEQ location inversion problem. Consequently, even though DL algorithms are strong solvers for nonlinear problems and have quick inference speeds, the data scarcity for training presents as the major challenge for using DL guided solutions to accurately locate MEQs. Moreover, the accuracy of predicted locations using conventional earthquake location methods (e.g., minimizing travel-time misfit) highly depends on the velocity model used. Simplified velocity models can result in less precise location predictions due to the lack of local heterogeneities present in the model. Using higher resolution velocity models that include more local geological features will incur higher computation costs. As such, it is pivotal to develop a practical method that combines the benefits of DL (quick inference times and strong nonlinear solving abilities), address the paucity of field training data, and integrates high-resolution realistic velocity models, to estimate induced MEQ locations for EGS monitoring.

In this study, we present a transfer learning workflow using probabilistic multilayer perceptron (PMLP) to accurately predict MEQ locations from waveform data. Transfer learning involves applying a machine learning model, initially trained on one dataset, to a different but related dataset. The knowledge transfer technique is especially beneficial in applications scenarios where collecting extensive training data is impractical or unfeasible. This approach serves as the basis of our study to locate field MEQs at the Newberry EGS site. The workflow encompasses three parts. Firstly, we use a high-resolution 3D velocity model created by Matzel et al. (2014) to simulate numerous synthetic MEQ events using 3D acoustic finite-difference modeling. From the synthetic waveforms, we extract its first arrivals. In practice, since we do not have the MEQs event origin time, we compute the cross-correlation of the first arrivals such that the first arrival of the master trace is at zero time lag. The time lags at other receivers contain the same moveout pattern as the first arrivals. Secondly, we train a PMLP that inputs cross-correlation time lags and outputs the locations (x, y, z) of MEQs. Lastly, we apply the trained PMLP onto the field dataset to obtain field MEQ location predictions. We are essentially leveraging transfer learning principles by allowing the neural network to train on realistic or *physics-informed* synthetic dataset, and then apply its knowledge learned onto field waveforms to predict the induced MEQ locations.

This manuscript is organized as follows. Firstly, we provide some background on the Newberry EGS and its field collected dataset. Secondly, we introduce our methodology, including the Newberry 3D velocity model, synthetic training dataset generation, and PMLP. Lastly, we discuss and interpret our results, and showcase our potential improvements to the previous understanding of Newberry EGS microseismicity.

2 Newberry EGS

Newberry Volcano is a shield volcano located in central Oregon, about 20 mi (35 km) south of the city of Bend and approximately 40 mi (65 km) east of the crest of the Cascade Range. The Newberry EGS was operated by AltaRock Energy and Davenport Newberry to test and demonstrate the EGS technology. After an extensive study of the state of the stress for the area (Cladouhos et al., 2011a; Davatzes and Hickman, 2011), this location was selected due to a very low permeability rate as well as a large conductive thermal anomaly that yields high-temperatures (Cladouhos et al., 2011b), making it ideal to test the creation of an EGS. Borehole logs reveal natural fractures extending from approximately depths of 2,000 m to 2,700 m. At these depths, the interpreted lithology consists of tuffs, basalt-andesite, and granodiorite. The EGS demonstration was stimulated two times, first in 2012 and later in 2014, to induce hydroshearing in the reservoir and enhance the movement of fluids through the system (Cladouhos et al., 2016).

In the 2012 fluid stimulation, there was a suspected casing leakage which caused induced MEQs at shallower than the intended depths. In the fall of 2014, casing repairs and re-stimulations were made. In the drilling well, the perforated liner is used to create multiple pathways for fluid injection into the rock formation, enabling efficient fracturing and increased heat exchange between the injected fluid and the surrounding hot rocks. The perforated liner starts at 1,912 m (6,272 ft) to 3,045 m (9,990 ft), along with a blank liner extending from 2,289 m (7,509 ft) to 2,493 m (8,177 ft). The depths at which the perforated liner is installed (1,912 m - 3,045 m) is considered the targeted depth for EGS stimulations. The experiment had a monitoring array of seven surface seismic stations and eight borehole stations. Figure 1 shows the general vicinity of Newberry EGS site. For the purposes of our study, we only show the borehole stations because the recorded waveforms are frequently missing at the surface stations. As such, we only work on data traces from the eight borehole stations throughout our study.

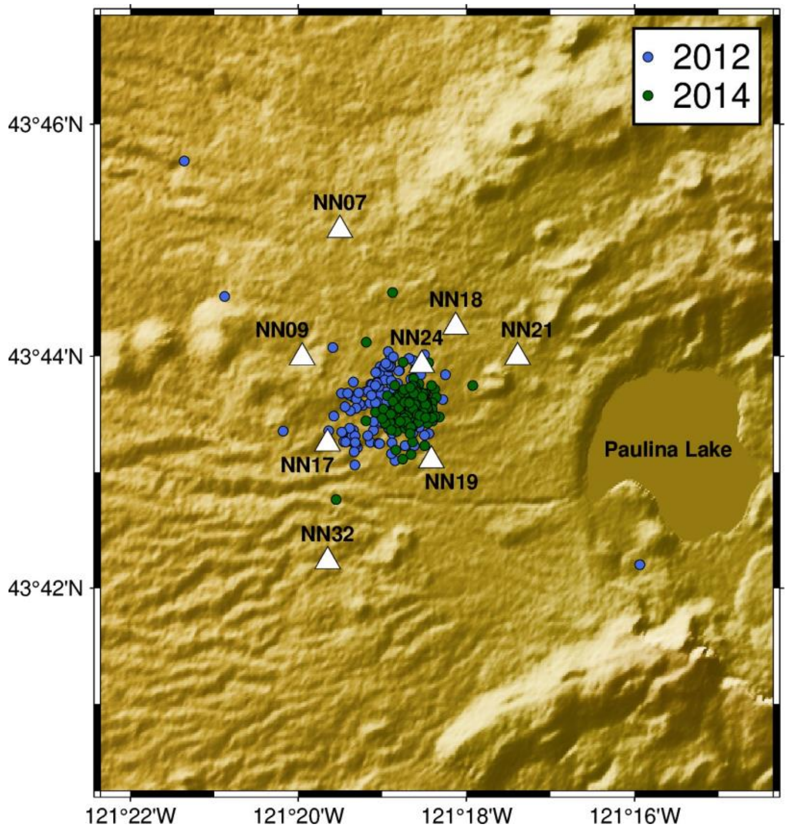

Figure 1: Aerial view of Newberry EGS site. The eight NN stations are borehole seismic stations. Events in blue are from the initial location catalog from the 2012 stimulation. Events in green are the corresponding locations of 2014 stimulation.

2.1 Microseismicity of the 2012 and 2014 EGS Stimulation

The 2012 stimulation lasted from Sept. 1, 2012, to Dec. 31, 2012. About 40,000 m^3 of water were injected with about 90% of the events were above the casing shoe (depths less than 1,830m (6,000 ft)), suggesting that injected fluid had leaked out of the casing to stimulate relatively shallow and cool rock. In the summer of 2013, caliper and video logs confirmed that there was both a horizontal crack in the casing at 683 m (2,240 ft) depth and a leak in the parasitic aeration string (AltaRock, 2014). In 2014, casing repairs were made, and second stimulation was conducted on Sept. 22, 2014, until Nov. 30, 2014. As for microseismicity, the seismic acquisition software automatically identified events, generated preliminary P- and S-wave picks and locations.

During the 2012 stimulation, about 175 events were located with magnitudes between M 0.0 and M 2.3. As for the 2014 stimulation, about 398 events were located with magnitudes between M 0.0 and M 2.2 (Cladouhos et al., 2016).

As for the data availability (http://fracture.lbl.gov/Newberry/Location.txt – assessed October 2019), there were only 149 datasets comprising waveforms and locations for the 2012 stimulation. For the 2014 stimulation, only 334 datasets are available.

3 Methodology

The main objective of this study is to develop DL algorithms to predict the locations of MEQs induced in the Newberry EGS, using waveform features, specifically cross-correlation time lags. The workflow is summarized in Figure 2. The workflow methodology can be divided into four parts. Firstly, we obtain a realistic seismic velocity model that is derived from field observations. Secondly, we simulate numerous synthetic MEQs, and their corresponding waveforms based on the field-informed velocity model. Thirdly, we use a neural network (PMLP in this study) to map the relationship from cross-correlation time lags (derived from waveforms) to MEQ locations (x,y,z). Lastly, we apply the trained PMLP onto the field waveforms to obtain Newberry MEQ location predictions.

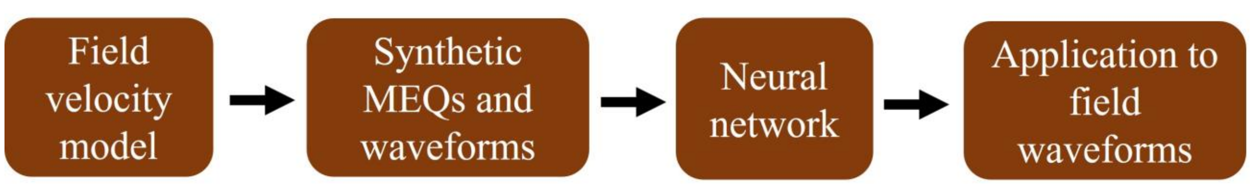

Figure 2: The workflow of this study begins with using a realistic velocity model derived from field measurements to generate numerous random MEQs. Next, we simulate the corresponding MEQ waveforms using 3D acoustic forward modeling. Following this, we extract cross-correlation time lags from these waveforms. These time lags are then utilized as inputs for our neural network, with the MEQ locations serving as outputs. After the neural network is adequately trained, we implement *transfer learning*, by applying this trained neural network to the actual field waveforms to obtain accurate location predictions.

3.1 Newberry Seismic Velocity Model

Matzel et al. (2014) computed ambient noise correlations from 22 seismic stations in the Newberry network, together with 12 additional stations from the nearby CC (Cascade Chain), UO (University of Oregon), and UW (University of Washington) seismic networks. The Green's functions that emerged from the cross-correlation waveforms were treated as seismic record and inverted for the best fitting 1D model along each path, resulting in Vp, Vs, and Qs models. For this study, we use the Vp model as a basis for our study.

The original format of velocity model is in latitude, longitude, and altitude (elevation above sea level). As such, we apply these preprocessing steps to convert the location to appropriate scales:

1) We first convert the latitude and longitude to easting and northing coordinates using the open-source software UTM (https://github.com/Turbo87/utm).
2) Next, we convert the altitude to depth below ground by subtracting altitude from the local topography.
3) Due to the significantly larger easting and northing values compared to depth, we normalize these values by subtracting them from the easting and northing coordinates

of centroid of the 15 stations. This ensures the new coordinates system is centered around the seismic stations.

4) Finally, we upsample the original velocity model from spatial sampling (dx, dy, dz) of 500 m to 25 m to satisfy seismic modeling numerical stability requirements (Igel, 2017).

Similarly, we also preprocess the locations of the field MEQ events. We overlay the velocity model with 2012 and 2014 stimulation initially located MEQ events in Figure 3. The 2012 stimulation MEQs are scattered as far as ~2 km away from the well bore, with the majority of events lying at depths of 2.0 – 3 km. These initial location estimates are incorrect (see Figures 3a and 3b) as there was a casing leak and most of the MEQs were later relocated to much shallower depths (0.6 – 1.3 km). As for the 2014 stimulation MEQs, the initial locations are noticeable at the wrong depths (Figures 3a and 3b) as the fluid injection was correctly stimulated at intended depths of ~ 1.9 – 3.0 km (Cladouhos et al., 2016). Moreover, we note that the velocity model completely covers the spatial extent of all the MEQs. This allows us to generate synthetic MEQs anywhere within the velocity model and simulate their corresponding waveforms.

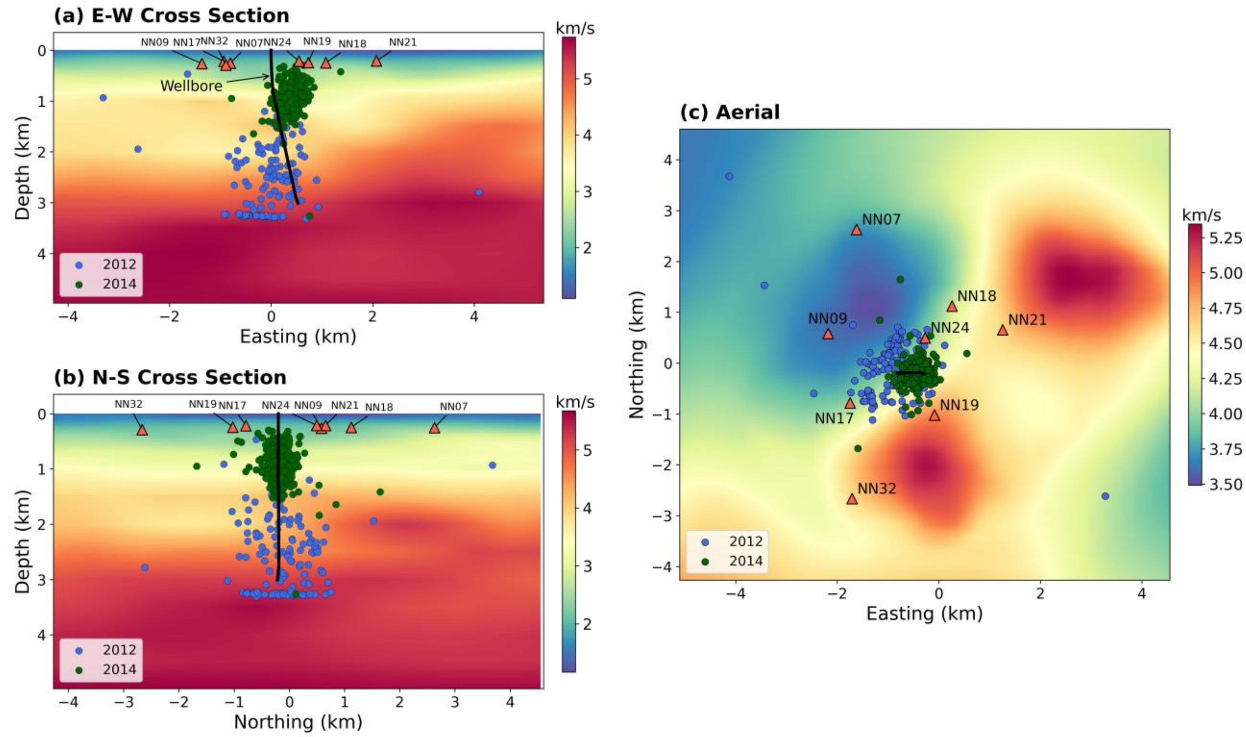

Figure 3: 3D P-wave velocity model of Newberry EGS site generated by Matzel et al. (2014). (a) represents the East-West cross section, (b) is the North-South cross section, and (c) is the aerial view of the location. Blue dots are initially located events from 2012 stimulation and green dots are from 2014 stimulation. In the downloaded raw dataset, there are 149 events for 2012 stimulation and 344 events for 2014 stimulation.

3.2 Synthetic MEQs, 3D Acoustic Waveforms, and Cross-correlation Time Lags

From the velocity model, we generate 10,000 artificial events across the entire extent of velocity model, and another 10,000 events to focus on the regions below the seismic stations which is also the injection zone (Figure 4). We note that the artificial MEQ events concentrate at the regions with field events. Next, we perform acoustic wave seismic modeling using the open-source Madagascar software (https://www.reproducibility.org/wiki/Main_Page) to generate the synthetic waveforms.

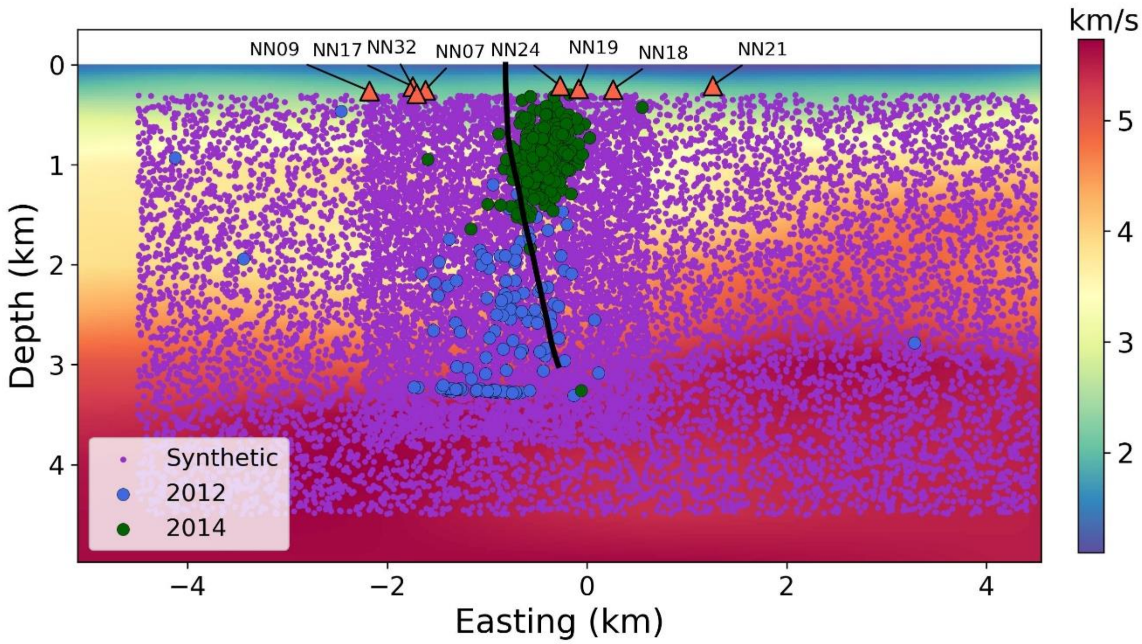

Figure 4: 10,000 synthetic events (purple) covering almost the entire spatial extent of velocity model. There are an additional 10,000 events covering the regions below the seismic stations. Blue dots are initially located events from 2012 stimulation and green dots are from 2014 stimulation.

Figure 5 shows an example snapshot of P-wave arriving at the receivers. The P-wave is analogous to the first arrivals emanated from induced MEQs during fluid stimulation. Figure 6a shows an example of waveforms generated (in black) from seismic modeling. It is important to highlight that the moveout pattern is caused by the relative MEQ location to receivers. For different MEQs at other locations, the time taken for first arrivals to arrive at the receivers cause different moveout patterns. We pick the first arrivals from the waveforms and create corresponding delta functions (red spikes in Figure 6a). Next, we use the trace at NN17 as the master trace to cross-correlate with all traces within a seismic gather. The cross-correlations aims to preserve the moveout information such that the time lag at the master trace (NN17) is zero, while the time lags at other traces correspond to the moveout pattern. Figure 6b shows the resulting cross-correlations with labeled time lags. The time lags are directly indicative of the moveout caused by the relative location of MEQ and receiver locations. The time lags are treated as the input of the neural network whereas, the location information (easting, northing, depth) is treated as the output.

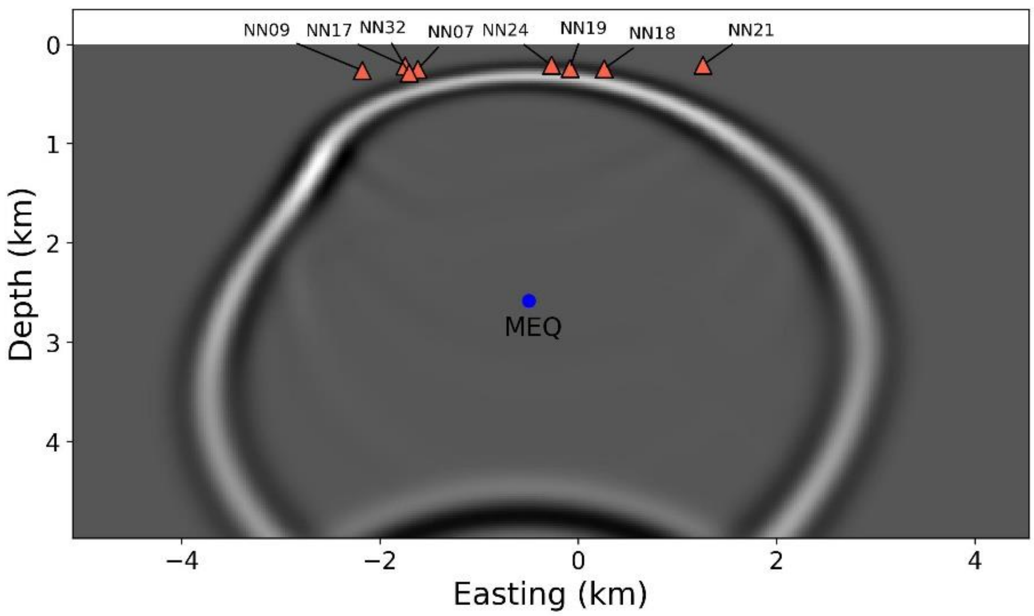

Figure 5: Example snapshot of pressure wave arriving at receivers.

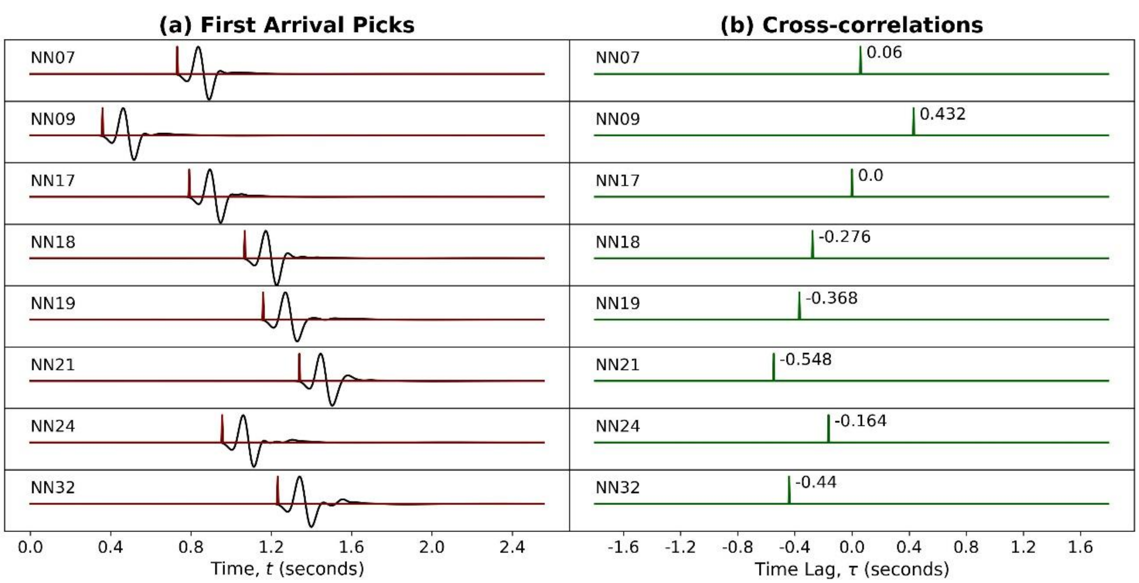

Figure 6: (a) shows an example of synthetic waveforms (black) and first arrival picks converted to delta functions (red). (b) is the corresponding cross-correlogram computed from using NN17 as master trace to cross-correlate with all traces in the seismic gather. Labeled numbers indicate time lags, which represent the moveout.

3.3 Probabilistic Multilayer Perceptron

Multilayer perceptrons (MLPs) are the fundamental building blocks of feedforward neural networks that consist of multiple layers of interconnected nodes and neurons. MLPs are also commonly referred to as artificial neural networks and deep neural networks. A simple MLP consists of an input layer, one or more hidden layers, and output layer (Figure 7). Each neuron in a layer is connected to all the neurons in the previous and next layers, with associated weights assigned to each connection. Additionally, each neuron has an activation function that determines its output based on the weighted sum of its inputs.

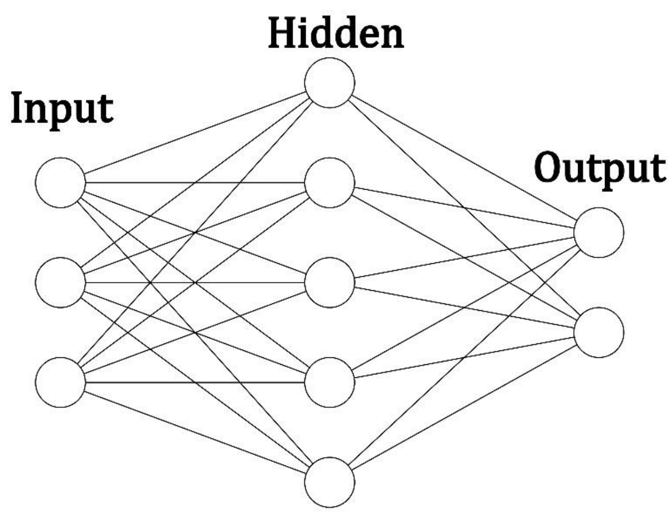

Figure 7: A simple MLP that consists of an input, hidden, and output layer. The circles in the hidden layer represent individual neurons.

To express MLPs mathematically, let X be an input vector as $X = (x_1, x_2, x_3 \dots, x_n)$. At the hidden layer, the neurons can be expressed as:

$$Z = a(W_h X + b_h),$$

where Z is the output of the hidden layer and a is the activation function, W_h and b_h are the weights and biases of hidden layer. The activation function introduces nonlinearity to the system so that the MLP can effectively learn the appropriate weights and biases to solve the nonlinear problem. Without the activation function, the system would be linear and the training of MLP would not converge. At the output layer:

$$\hat{y} = a(W_o Z + b_o),$$

where $\hat{y}$ is the MLP output (prediction). The training process of an MLP involves adjusting the weights and biases to minimize the error between the predicted output and the target output, typically using the backpropagation algorithm and gradient descent optimization (Lecun et al, 2015). Some examples of loss function include mean square error (L2 norm) and mean absolute error (L1 norm).

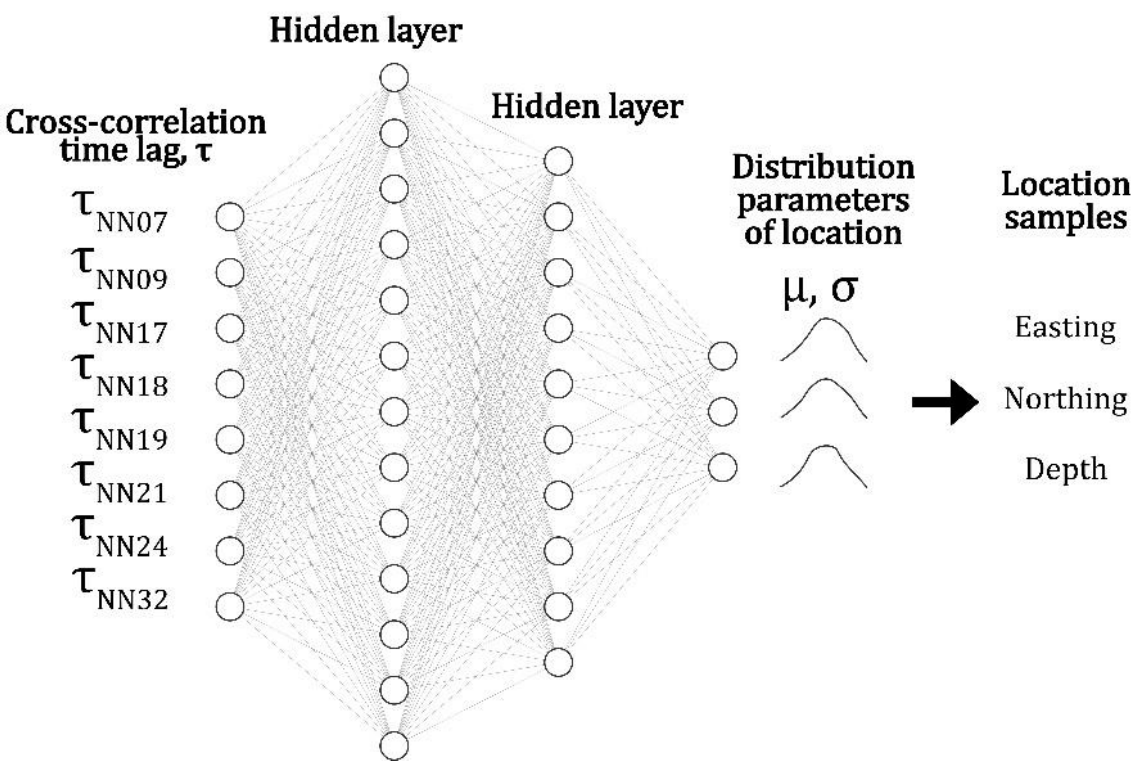

Figure 8: Probabilistic MLP (PMLP) architecture. It accepts cross-correlation time lags as input and outputs mean (μ) and standard deviation (σ) of MEQ location. μ and σ can then be used to sample from the Gaussian distribution to obtain MEQ location samples (easting, northing, depth). We compute the average location as the final location prediction.

In earthquake location prediction, uncertainties play a key role in the process of quantifying the reliability of NN predictions. Standard MLPs are deterministic, meaning they output deterministic point estimates. Here, we use the probabilistic MLP (PMLP) to predict MEQ locations from cross-correlation time lags. Figure 8 shows the architecture of PMLP. PMLP contains a preceding conventional MLP structure, however, instead of directly predicting the location of MEQs, it predicts the distribution parameters (mean and standard deviation) of MEQ locations which are assumed to follow a Gaussian distribution. Essentially, PMLP seeks to find the best distribution parameters that make the output training data (event locations) most probable. In mathematical terms, PMLP can be expressed as a general nonlinear regressor by:

$$PMLP(\tau_N) = [\mu_N, \sigma_N],$$

where τ is the cross-correlation time lags, μ and σ are mean and standard deviation across N number of input data (time lags).

To determine the set of μ and σ that can make MEQ locations most probable, we employ maximum likelihood estimation (MLE). MLE finds the parameters that maximize the likelihood of observing the MEQ locations given the PMLP regression model. In practice, it is easier to maximize the log of the likelihood, or equivalently, minimize the negative log-likelihood. The Gaussian likelihood function, L, for a single MEQ location (y), is given by:

$$L(y;\ \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(y-\mu)^2}{2\sigma^2}}$$

The negative log-likelihood simply means:

$$-\log\big(L(\tau;\ \mu,\sigma)\big) = -\log\left(\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(y-\mu)^2}{2\sigma^2}}\right)$$

By applying the logarithm properties:

$$-\log\left(\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(y-\mu)^2}{2\sigma^2}}\right) = -\log\left(\frac{1}{\sigma\sqrt{2\pi}}\right) - \log\left(e^{-\frac{(y-\mu)^2}{2\sigma^2}}\right)$$

$$= \log\big(\sigma\sqrt{2\pi}\big) + \frac{(y-\mu)^2}{2\sigma^2}$$

For optimization purposes, we can leave out the constant term $\log\big(\sqrt{2\pi}\big)$, and the resulting negative log-likelihood, NLL, (Nix and Weigend, 1994) can be defined as:

$$NLL = -\log\big(L(\tau;\ \mu,\sigma)\big) = \frac{1}{N}\sum_{i=1}^{N}\left[\log\sigma(\tau_i) + \frac{(y_i-\mu(\tau_i))^2}{2\sigma(\tau_i)^2}\right]$$

where τ_i is the cross-correlation time lags, σ is the standard deviation, y is the MEQ location values (easting, northing, depth), μ is the mean, and $i \in [1, N]$ where N is the number of training dataset. Simply put, the negative log-likelihood loss function finds the parameters (μ, σ) that best predict the MEQ locations in the training dataset.

We use ReLU as the activation function for all hidden layers. At the final layer, we only use fully-connected (dense) neurons without activation function for the easting and northing components as they contain negative and positive values. For the depth output component, we enforce a ReLU activation as the depth values are always positive.

In practice, we can apply the trained PMLP to unseen time lags, τ, to predict μ and σ. For example, for one set of time lags, the PMLP directly predicts one set of μ and σ of the MEQ location (easting, northing, depth). The predicted μ and σ are used to sample from the Gaussian distribution to obtain the realizations of predicted MEQ locations. Since this process is probabilistic, multiple sampling yields slightly different locations. This allows repeated sampling that produces a range of predictions, which we can then compute the mean as the final MEQ location prediction; and compute statistical uncertainties from the range of sampled predictions. Here, we note that the estimated uncertainties come from the trained PMLP regression model, instead of the error introduced from the input time lags. The uncertainties represent the range of values that the trained PMLP will produce.

4 Results

This section is divided into three main parts. First, we discuss the performance of PMLP on synthetic dataset, i.e., training and testing on time lags generated from synthetic events. Second, we discuss the results of applying the trained PMLP on the 2012 stimulation MEQ dataset. Third, we discuss the 2014 stimulation MEQ location estimates and interpret our results based on the location's geology.

4.1 Synthethic Tests

From the total 20,000 generated dataset, we remove a certain number of bad simulations due to edge effects, resulting in the new total to be 19,738 datasets. We randomly split 16,875 (85%) for training, 1,876 (10%) for validation and 987 (5%) for testing. To select the best trained weights, we evaluate the Euclidean distance between predictions and ground truth. The Euclidean distance, D, is calculated by:

$$D = \sqrt{\left(\hat{E}_i - E_i\right)^2 + \left(\hat{N}_i - N_i\right)^2 + \left(\hat{H}_i - H_i\right)^2}$$

where $\hat{E}, \hat{N}, \hat{H}$ are predicted easting, northing, and depth, and E, N, H are the respective ground truth. During training, we save the best weights that predict the lowest Euclidean distance on the validation dataset. Figure 9a shows the training progress in logarithms for better visualization, and the best weight is selected at epoch 236 along with the average validation loss of 42 m. During inference, the trained PMLP samples the Gaussian distribution associated with the MEQ location, thus generating slightly differing predictions for each sample. Figure 9b shows 300 samples of location predictions that are based on one input event. As expected, the samples are scattered around the mean location. For our study, we generate 3,000 samples and compute the corresponding mean as the final location prediction.

The average Euclidean distance loss on testing dataset is 41 m. We further examine the prediction errors on the testing dataset (Figure 10) and compute simple statistics tests to gauge the prediction performance. For example, the prediction errors have 90% likeliness to fall between [-53, 55] m in easting component; [-62, 59] m in northing component; and [-66, 68] m in depth component. In addition, we compute the 95% confidence interval, and the errors are approximately 10 m more on each side. In the broader context, the velocity model has dimensions of approximately 9 km x 9 km x 4 km, and PMLP's prediction errors are less than 100 m, corresponding to about a 1% error in each dimension.

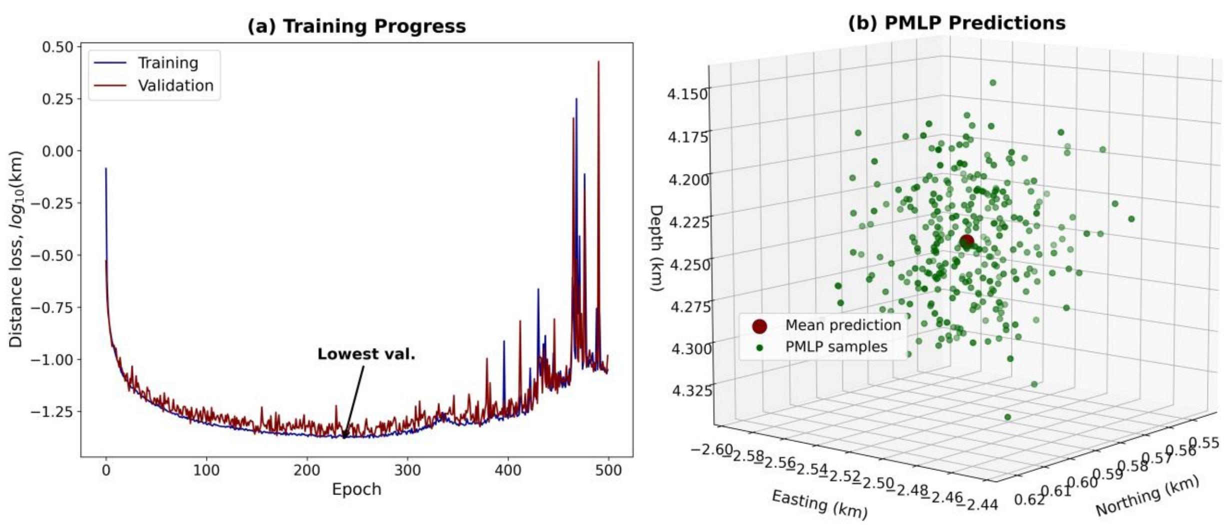

Figure 9: (a) shows the training progress of PMLP. Blue curve is the training loss and red curve is the validation loss. The loss refers to the Euclidean distance. We use the model weights at

epoch 236 as the final weights as that is when the validation loss is the lowest. (b) shows the PMLP predictions (300 samples) for one input event. The mean prediction is computed as the final prediction.

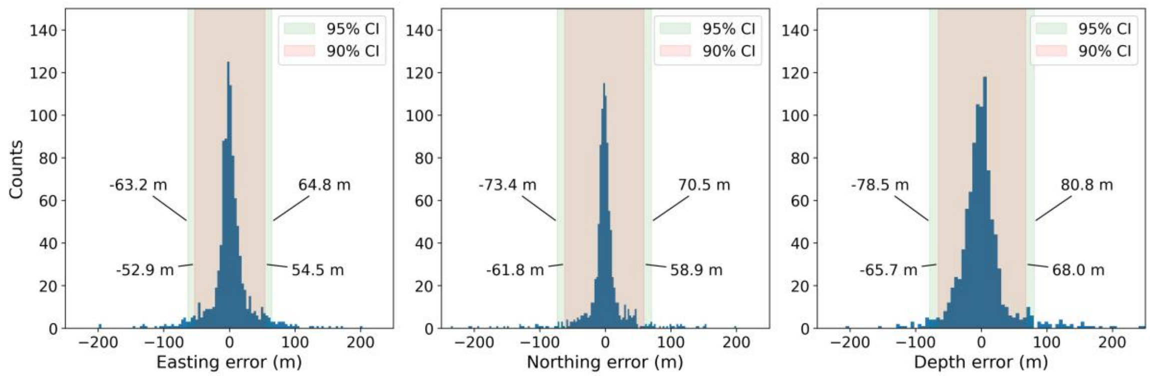

Figure 10: Prediction on testing dataset (n=987). The 95% confidence interval of the prediction error at easting is [-63.2, 64.8] m, northing is [-73.4, 70.5] m, and depth is [-78.5, 80.8] m. As for the 90% confidence interval, the prediction error at easting is [-52.9, 54.5] m, northing is [-61.8, 58.9] m, and depth is [-65.7, 68.0] m.

4.2 Field Application – 2012 Stimulation

Out of the 149 triggered waveforms, we consider 10 events to be outliers as they are out of bounds i.e., located above stations and far away from the injection zone. The PMLP model requires input from all eight borehole stations for accurate predictions. Therefore, we can only consider events that have recorded waveforms at each of these eight stations. This criterion further narrows our analysis to 113 events with waveforms in those borehole receivers. As the P-wave synthetic waveforms used in training, we only consider the first arrival picks of the vertical component in all field waveforms. We assume this is reasonable because all the induced MEQs are located below the receivers and the vertical component sensor can sufficiently pick up the first arrival waves.

Before picking the first arrivals, we apply these preprocessing steps to the field waveforms:

1) We use trim the waveforms using the same start and end time to ensure the event waveforms are aligned at the same time window.
2) We apply a bandpass of 5 – 15 Hz.
3) Lastly, we normalize the traces based on their maximum value.

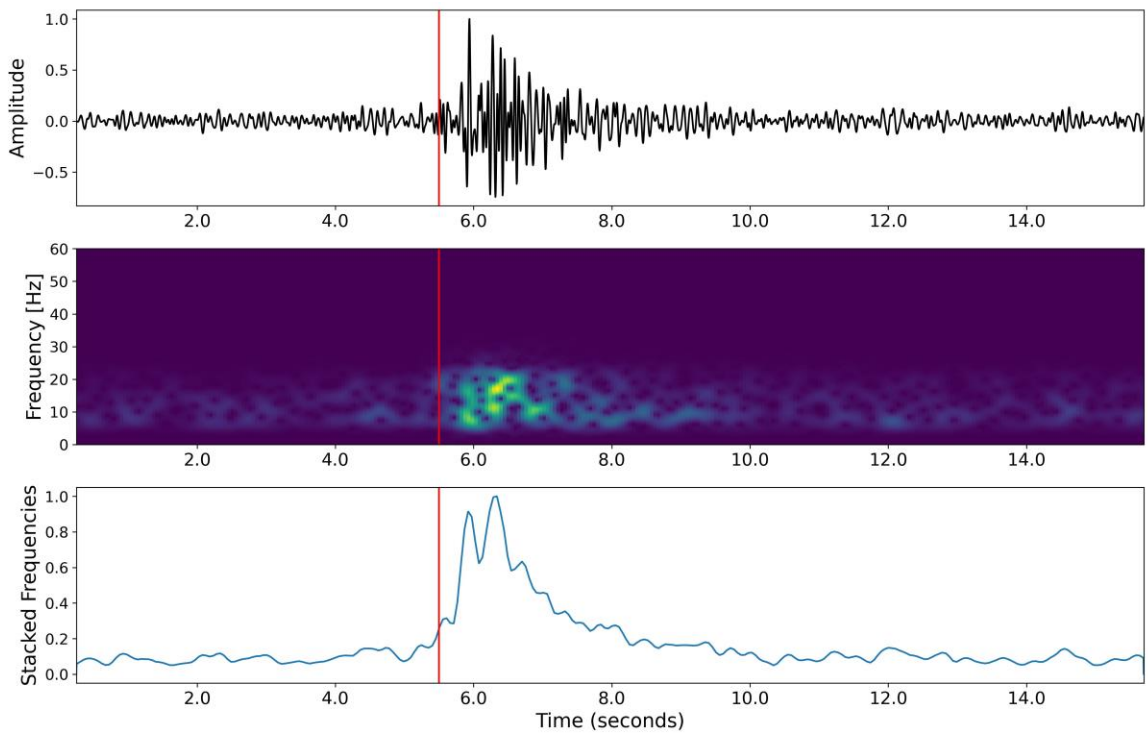

Figure 11: Top panel shows an example of seismic trace. Middle panel shows the corresponding bandpassed frequency spectra. In the bottom panel, the frequency spectra are summed up in the vertical component and normalized based on its absolute maximum value. The red line shows the handpicked first arrival for which the picking location is guided by the onset of energy as depicted in the bottom panel.

To obtain the most accurate first arrival picks, we compute the frequency spectra and stack the frequency's amplitudes to use as a guide for picking (Figure 11). The stacked frequencies illuminate the first arriving energy associated with the MEQ first arrivals. We carefully handpick the first arrivals, compute the cross-correlations and retrieve the corresponding time lags.

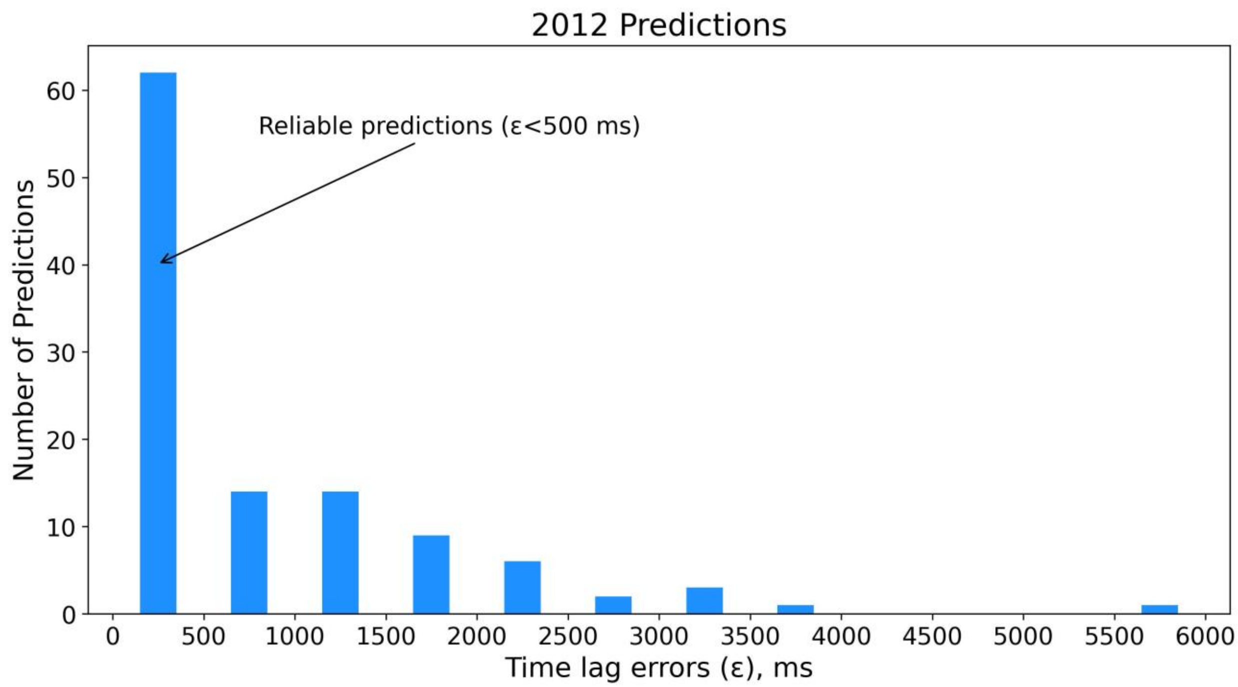

Figure 12: Histogram showing the mean-square-errors between predicted forward and field time lags for 2012 dataset. We consider the predictions falling within first bin as reliable.

Given the encouraging results observed from the application of PMLP in synthetic tests, we apply the trained PMLP model on the computed field time lags. However, preprocessing is needed due to the detection of unreliable location estimates within the raw predictions. This is evidenced by the significant discrepancies in time lag errors when comparing synthetic forward time lags with field-picked time lags. To address this, we calculate the mean-square-error for time lags (ε) between the field-picked and predicted forward time lags across all predictions, as shown in Figure 12. The histogram of ε guides our reliability criteria: predictions with ε below 500 ms are deemed reliable, which also corresponds to the most frequent histogram bin. Predicted locations in this bin have good match between the predicted forward time lags and that from field-picks. After the preprocessing step, we identify a total of 62 reliable predicted locations. Figure 13 shows the predicted locations and three examples of comparison of predicted forward first arrival picks vs. field first arrival picks.

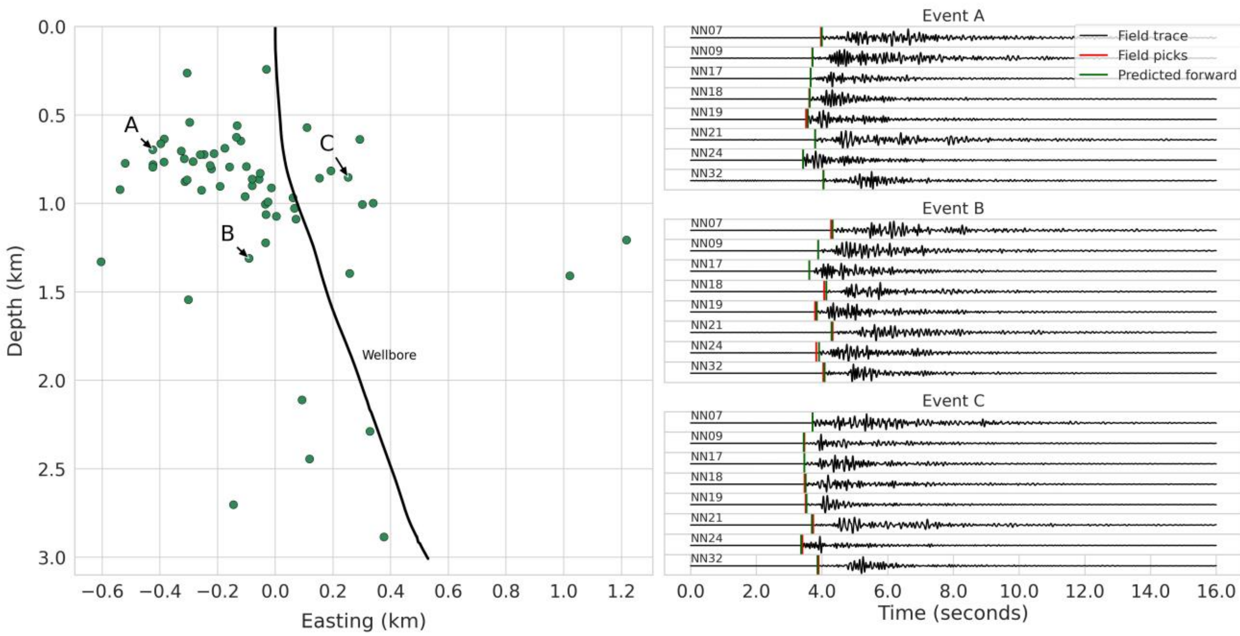

Figure 13: PMLP predicted MEQs for 2012 stimulation. The left panel shows the cross section of the predicted location of MEQs. Right panel highlights three examples (A, B, C) to show the comparison of synthetic (predicted forward) vs. field picked first arrivals.

From these predicted MEQ locations, we notice a majority concentrate at depths of 0.5 – 1.2 km. In comparison with the relocated events for 2012 stimulation done via physics-based inversion (Cladouhos et al., 2016), their predictions concentrate at depths of 0.5 – 1.3 km, which aligns with those from PMLP prediction. This directly corroborates with casing leak scenario which causes the induced microseismicity shallower than the intended depths (1.9 – 3.0 km). Events A, B, and C are three examples of predictions that display significantly good match between field (red vertical lines) and predicted forward picks (green vertical lines). The aligned depths of our predictions with those reported by Cladouhos et al. (2016), alongside the closely overlapping first arrival picks as evidenced in Events A, B, and C (in Figure 13; right panel), underscores the accuracy and reliability of PMLP in predicting microseismic locations.

4.3 Field Application – 2014 Stimulation

From the available 334 MEQ waveforms, we select 292 as the remaining events do not contain seismic traces in all eight stations. As the 2014 raw waveforms contain more noise and the original data format are less structured, it is essential to preprocess the field waveforms before picking the first arrivals. First, for each event, we find the most common start and end time within all traces because many waveforms have different start times. Second, we apply a bandpass filter of 6 – 20 Hz to remove noise of higher and lower frequencies. Third, we demean and normalize the traces so that the resulting seismograms can be picked easily.

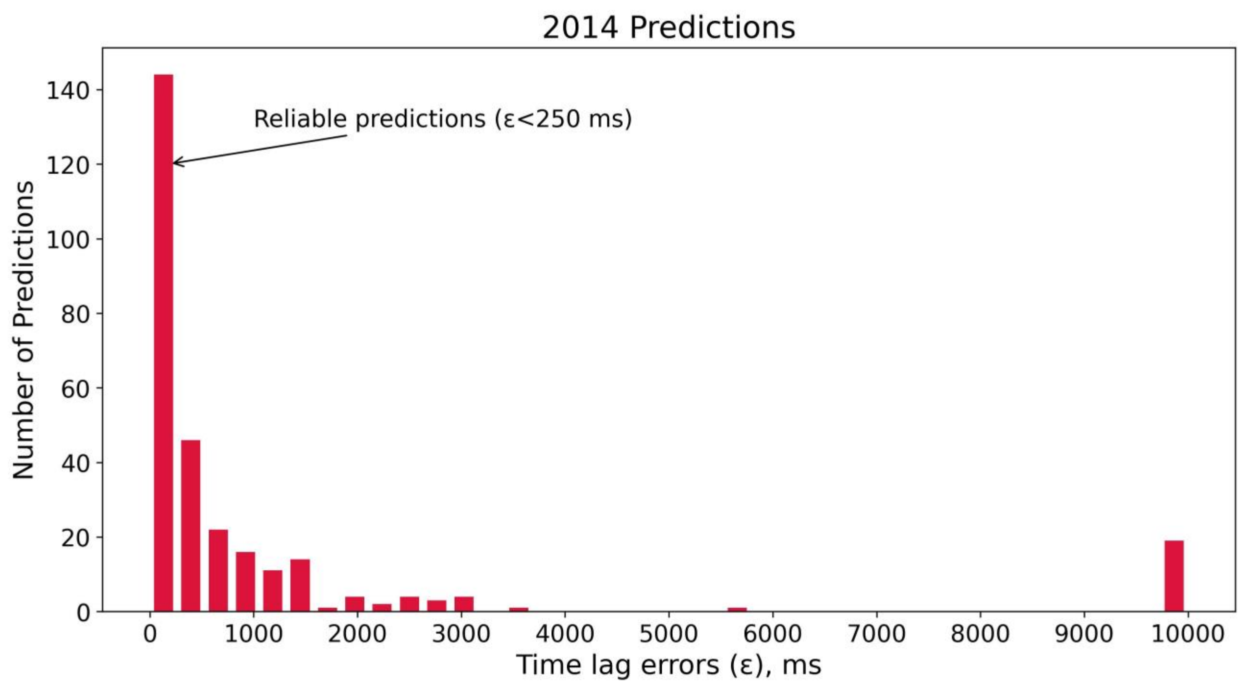

Figure 14: Similarly for 2014 dataset, we plot the histogram showing the mean-square-errors between predicted forward and field time lags. We consider the predictions falling within the first bin as reliable.

Similarly, we first apply the trained PMLP on the 2014 dataset, compute the cross-correlated time lags for each predicted location, and compare them with those from the field picks. The histogram of the errors is plotted in Figure 14. From the histogram, we see the first bin (250 ms) has the greatest number of predictions, which also means that these predictions are the most accurate due to their low error between the predicted forward and field picks. This entails a total of 142 reliable predictions (Figure S1) based on their first arrivals match. However, we notice that there are two clusters of predictions separated by a noticeable gap (lack of predictions) around 1.8 km depth. In Cladouhos et al. (2016), the physics-based inversion study did not show any location estimates above 1.8 km depth. Upon inspecting the first arrival match between the synthetic picks and field picks (Figure S2), we postulate that although the first arrival match is good, we think these predictions likely stem from incorrect first arrival picks. For instance, the waveforms in the six examples (Events A-F) contain relatively more noise and our picks may not best represent the real first arrivals. Following this, we consider it appropriate to only keep the events below 1.8 km depth (128 events) as the final predictions for interpretation (shown in Figure 15). In total, there are 128 events used as the final predictions.

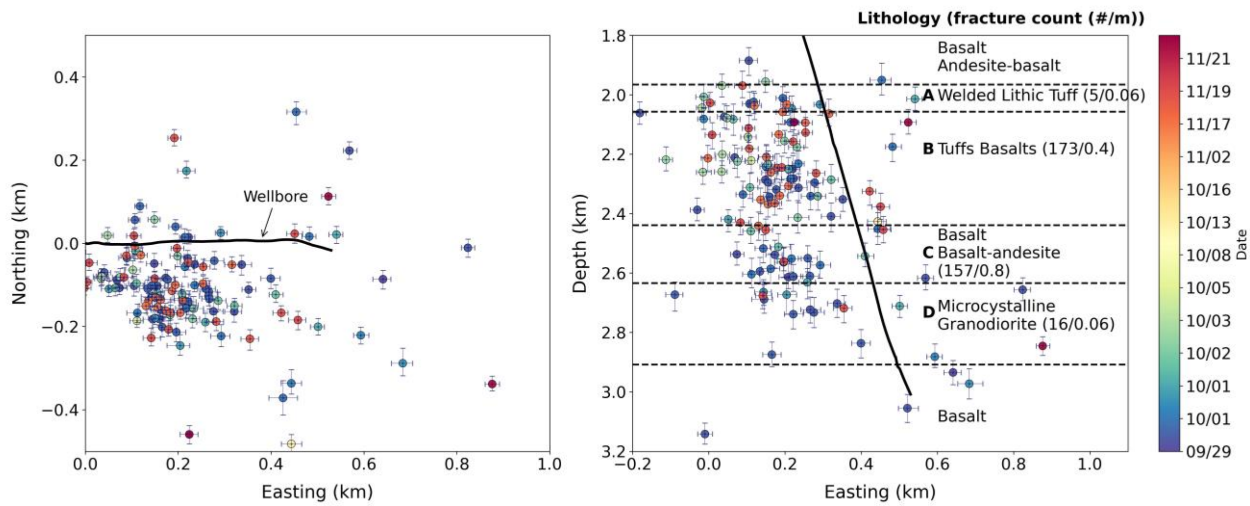

Figure 15: Overlay of 2014 stimulation location predictions with interpreted geologic zones from Cladouhos et al. (2016). The fracture count is determined by counting fractures within the NWG 55-29 borehole. The error bars are calculated by using the range of location samples predicted from repeated sampling of the trained PMLP.

We cross reference the final predictions with the subsurface geologic information determined from the NWG 55-29 borehole (Cladouhos et al., 2016) in Figure 15 (right panel). Additionally, we overlay the MEQ location predictions with the appropriate geologic zones. We re-reference the MEQ locations relative to the wellbore coordinates for better comparison. The fracture count in each zone is determined from images captured from a borehole televiewer (BHTV) survey. For example, in Zone A, there are 5 natural fractures per 0.06 m. We note that this fracture count is only representative of the region within the borehole. In general, we see that Zone B – Tuffs and Basalts (173/0.4 m) contains the most fractures, followed by Zone C (157/0.8 m), Zone D (16/0.6 m), and Zone A (5 / 0.06 m). Since natural fractures are more predisposed to rupture (Rutqvist et al., 2013; Lei et al., 2021) during fluid stimulation, it is reasonable to assume that the number of MEQs would be higher in regions that contain more natural fractures. Most of our MEQ predictions (n = 67) are in Zone B, which aligns with the zone having the highest fracture count. This correlation supports the validity of PMLP predictions in reflecting the geological conditions of the subsurface. Our confidence in these predictions is further bolstered by the excellent match of first arrival picks between synthetic and field traces. We select eight examples (see Figure S3) and compare their first arrivals with those picked from field, and the data match appears to match well (Figure S4). From the BHTV fracture counts, Zone C has the second highest number of fractures (157/0.8 m), also coincides with our predictions in terms of number of MEQ predictions. While Zones A and D show an equal number of MEQ predictions (n = 16), BHTV fracture counts reveal more fractures in Zone D (n = 16) than in Zone A (n = 5). It is important to note that these fracture counts are derived from BHTV images within the borehole and may not fully represent the entire geological strata. Additionally, fracture counts in Zones B and C are an order of magnitude higher than in Zones A and D, a pattern that is consistently echoed in our MEQ predictions.

The microseismic events predicted by the PMLP form a cloud analogous to that derived from the physics-based inversion (see Fig. 12 in Cladouhos et al. (2016)), albeit with most of our predicted events notably clustered to the west of the wellbore. The westward clustering observed in our predictions may be attributed to the DL model's reliance solely on P-wave picks, as opposed to the physics-based inversion which utilized both P- and S-wave arrivals. This methodological difference inherently results in slightly distinct MEQ location predictions. While we acknowledge that the refinement of first arrival picks could potentially enhance the model's accuracy, the strong data match between synthetic and field trace first arrivals suggests our current results are reasonable. The slight difference in the spatial distribution of predicted MEQs is counterbalanced by the significant advantage of the DL model in terms of inference speed. Once trained, our model can deliver predictions within seconds, a speed that significantly outperforms traditional source localization methods.

In general, our results indicate that most induced MEQs are located above 2.90 km depth (until Zone D), with a few outliers. In contrast, Cladouhos et al. (2016) demonstrate that the relocated MEQs extend down to 3.2 km, with a significant number of events occurring between depths of 3.05 and 3.2 km. A more in-depth study was done to interpret the natural fractures (Davatzes and Hickman, 2011), and it is found that fractures are only present at depths of 2.0 km to approximately 2.7 km. Since injection of fluid increases pore pressure within these fractures, it reduces the effective normal stress acting on the fracture walls. This reduction in normal stress decreases the frictional resistance to shear sliding along the fracture plane. As pore pressure continues to increase, it eventually overcomes the frictional resistance and causes the natural fractures to slip or rupture, a process known as shear reactivation (Das and Zoback, 2011; Rutqvist et al., 2013). Based on our findings, we can reasonably conclude that our predictions are accurate due to the presence of natural fractures matching the depths of predicted MEQs. Therefore, we postulate that our results could potentially be seen as an improvement to the method used in Cladouhos et al. (2016).

As for the uncertainties associated with the prediction, we compute the range (lower and upper bounds) of location samples generated from repeated predictions. We show the uncertainties in Figure 15 in the form of error bars. We avoid the use of standard deviation as uncertainties as we would want to know the full extent of location predictions produced from the repeated predictions. The uncertainties suggest a lesser variation and stronger confidence in the East-West component, while more variation in the depth component. This coincides with conventional earthquake location inversion methods in which the depth component typically shows larger uncertainties.

5 Picking Error Sensitivity Analysis

We perform a sensitivity analysis on the synthetic testing dataset to gauge how much the predicted locations would be affected by errors introduced in first arrival picks. We create four levels of errors in terms of number of time samples (nt): 5, 10, 15, and 20. We assume that 5-time samples would be an appropriate margin of error to characterize errors associated with handpicking. Next, we apply each error level to randomly selected receivers. Then, we compute the average Euclidean distance loss between the error-perturbed first arrivals and ground truth across the testing dataset. Table 1 shows the sensitivity of picking errors in accordance with the number of affected receivers. For example, when there are two random receivers affected with +/- 5-time samples (0.02 s) error, the average Euclidean distance loss is 0.1855 km. Our sensitivity analysis suggests that PMLP is highly sensitive to arrival picks. This is as expected because PMLP essentially only considers first-arrivals as input features. Therefore, in our study, we manually review all traces and re-pick first arrivals whenever it is necessary.

Picking errors (# nt (s))	**Number of affected receivers (selected by random)**								
	0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
5 (0.02)	0.0410	0.1320	0.1855	0.2170	0.2506	0.2873	0.3023	0.3357	0.3519
10 (0.04)	0.0410	0.2562	0.3518	0.4088	0.4893	0.5393	0.5624	0.6176	0.6770
15 (0.06)	0.0410	0.3740	0.5165	0.6118	0.7167	0.7729	0.8229	0.8903	0.9712
20 (0.08)	0.0411	0.4861	0.6623	0.8018	0.9194	0.9892	1.0799	1.1524	1.2401

Table 1: Corresponding Euclidean distance (in km) errors when first arrival picking errors are introduced. We test four picking errors: 5-, 10-, 15-, and 20-time samples. For each group, we test a variety of number of affected receivers, ranging from no (zero) receivers affected to all (eight) receivers affected.

6 Discussion

In this study, we address and overcome the issue of implementing DL methods to locate field microseismic events for EGS. In practice, the foremost challenge is the scarcity of training data due to the nature of how induced MEQs are difficult to detect (high noise levels and low magnitude), and thus resulting in limited data samples for DL training. Existing studies that use DL to locate earthquakes depend on large training samples. We overcome this challenge by utilizing a field-derived 3D P-wave velocity model to simulate synthetic acoustic waveforms from numerous artificial MEQs that encompasses the whole velocity model's spatial extent. Consequently, we train a probabilistic neural network that trains on waveform features (cross-correlation time lags) and outputs MEQ locations (easting, northing, depth). Through the principle of transfer learning, this learned knowledge can be applied to previously unseen data, such as Newberry EGS field waveforms, to accurately obtain field MEQ locations. This technique overcomes the issue of lack of field training data by essentially generating a highly realistic training dataset that contains all relevant physics, which facilitates field prediction.

Our methodology combines speed (inference under few seconds) and precision (good data match) in predicting MEQ locations, making it suitable for automated picking routines. Using 3D acoustic wave propagation physics, it can significantly improve preliminary location estimates from automatically triggered waveforms. The adaptable workflow allows for various improvements, such as using updated velocity models or more complete waveforms (elastic). Owing to its core structure as a multilayer perceptron, retraining with new data is computationally efficient. Our proposed workflow can act as a blueprint for location predictions in other EGS sites. As long as the site's velocity model can be estimated, we can simulate a large number of artificial events and their corresponding realistic synthetic waveforms. Overall, our approach offers a streamlined, effective solution for MEQ location estimation in EGS sites.

A prevailing question arises: why opt for picking over directly using full waveforms in DL-based source localization methods? Many DL-based methods, especially those using CNNs, take in entire waveforms to predict locations. CNNs extract features through sliding convolution windows, a process that largely focuses on changes in waveform amplitudes. This feature extraction method is akin to seismic phase picking, which is also similar to computer vision for identifying contrasting edges or boundaries. Despite this, low signal-to-noise ratio waveforms can mislead the feature extraction, which would ultimately introduce errors in location predictions. It is evident that manual picks offer the most precise seismic phase arrivals. While established relocation methods such as HypoDD and GlowClust (Trugman and Shearer, 2017) may exist as a straightforward option to apply those curated picks to, these methods have their limitations. In sparse seismic networks, these algorithms struggle to optimally refine earthquake locations and can face convergence issues if initial locations are far from true solutions. As outlined in our approach, we supply the PMLP with an extensive synthetic dataset for training. The combination of NNs' ability to map complex data patterns and extensive training dataset offers a robust solution in situations with sparse receiver networks. Our proposed transfer learning method basically computes a global solution of the travel-time-location problem within the confines of the given velocity model. This avoids the need for initial location estimates. Essentially, the trained neural network acts as a comprehensive lookup table, correlating travel-times with locations. While further refinement of travel-time picks can enhance predicted MEQ location accuracy, our method could offer an improvement to estimate initial MEQ locations during EGS stimulation.

In our study, we operate under the assumption that the employed 3D velocity model, while broadly representative of the subsurface layers, may not fully capture all lateral heterogeneities. Despite this limitation, the model provides a more nuanced geological context compared to the simplified 1D layered models typically used in automated picking software. In practice, implementing this 3D model on-site could enhance initial-stage predictions, offering valuable, rapid feedback for the stimulation process and thereby improving operational efficiency.

In a relevant work by Chen et al. (2022), the study used RF that takes in P-wave travel-time and location coordinates, and predicts MEQ location (x,y,z coordinates). In the study's field application example at a hydraulic fracturing site, it used similar transfer learning techniques by simulating many artificial events in a 3D layered velocity model and apply their trained RF model on field data to obtain field location predictions. The study noted that when the RF model is trained on dataset generated from 1D velocity model and applied to unseen data generated from complex 3D velocity model, the prediction accuracy suffers from significant errors, especially for deep events and events near high contrasting velocity anomalies. The study also used the Monte Carlo method to estimate its prediction uncertainties by perturbing the original velocity model and examine its prediction variations. Our proposed workflow provides several improvements and introduces new advantages. In our proposed method, we use a field-informed high-resolution 3D P-wave velocity model to simulate synthetic MEQ-waveforms dataset. The velocity model is created by inverting the Green's function derived from ambient noise correlations from multiple Newberry seismic stations. As 3D velocity models contain more physics than that of 1D models, the forward simulated waveforms are more realistic. Furthermore, our proposed workflow simplifies the uncertainty estimation. The probabilistic design of the PMLP provides slightly varied outcomes for predicted samples, allowing for uncertainty quantification. Therefore, we eliminate the need for perturbing the velocity model and re-running forward modeling, a feature that could be beneficial for near real-time EGS monitoring.

There are a few improvements that can be made to our workflow. Acoustic wave modeling only generates the first arriving P-waves, and as such, our neural network is limited to P-wave features (cross-correlation time lags). In future, elastic modeling routines can be considered as they simulate S-wave components. Since neural networks work best with more features to train, having both P- and S-wave features is likely to improve the location prediction. This is also especially useful in noisy traces where the P-wave first arrival is not obvious, but the S-wave is due to its larger amplitude.

Moreover, as demonstrated in the sensitivity analysis, our workflow relies largely on the accuracy of first arrival picks. We attempt to overcome this issue by manually examining all the traces and repick the incorrect first arrival travel-times. However, we acknowledge that there may be picking inaccuracies especially when the traces are very noisy. We considered the use of state-of-the-art DL-based phase pickers such as PhaseNet and EQTransformer, but upon testing, we found that those pickers were not reliable due to inconsistent P-wave picks. In the future, we believe that developing DL-based phase pickers specifically tailored for induced MEQ waveforms could significantly improve P-wave picking. These specialized DL-based pickers could accurately identify first arrival picks. When integrated with our location estimation method, this could create a comprehensive DL suite for automated waveform picking and location estimation.

Understanding the limitations of how many synthetic MEQs can influence the inversion results can be beneficial for long-term EGS monitoring. Ideally, the more data the better it is for DL generalization at the EGS site, however, the abundance of synthetic MEQ dataset comes at cost of long simulation and DL training times. As such, it would be helpful to examine the boundaries of training data size that are needed to produce similar inversion results when compared to large number of datasets.

7 Conclusions

Locating microearthquakes (MEQs) in enhanced geothermal systems (EGS) is fundamentally challenging due to the nonlinear relationship between waveforms and location. While deep learning (DL) methodologies have shown proficiency in predicting natural earthquake locations using waveform data, the majority of these studies depend heavily on large amount of field data for training. As such, the scarcity of field training data in EGS presents considerable challenges for the implementation of DL-based approaches. To overcome the scarcity of training data, we leverage transfer learning principles by introducing a practical workflow that accurately predicts MEQ locations from cross-correlation time lags using probabilistic multilayer perceptrons (PMLP). Our results on the 2012 Newberry EGS stimulation show major microseismic activity at depths of 0.5 – 1.2 km, which agrees with the casing leakage scenario that took place in the well bore. We further apply our methodology to the 2014 stimulation data and found that most of the MEQs concentrate at 2.0 – 2.9 km depths. This finding aligns with the presence of natural fractures which extend from 2.0 – 2.7 km. Excellent data (time lags) match indicate the completeness of inversion and suggest predictions can be trusted. The combination of good data match and the predisposition of natural fractures having ruptures caused by fluid stimulation lead us to conclude that the majority of microseismic activity happens shallower than 3.0 km. Training with prior information specific to an EGS site holds promise for enabling real-time monitoring in such environments.

Acknowledgments

We extend our foremost acknowledgement to the Department of Energy (DOE) for their support through grant DE-EE0008763. We are especially thankful to Dr. Dennise Templeton and Dr. Eric M. Matzel for providing the Newberry velocity model. Our appreciation also goes to Prof. Derek Elsworth for providing the Newberry MEQ dataset. Additionally, we are grateful to Dr. Chao Guo for his insightful perspectives on our early MEQ data analysis.

Data and Software Availability

Newberry dataset was originally downloaded from https://fracture.lbl.gov/Newberry/Location.txt in October 2019. However, the link is now defunct, and the dataset is moved to https://gdr.openei.org/home. The Madagascar 3D acoustic seismic modeling codes are available at https://www.reproducibility.org/wiki/Main_Page. Machine learning codes associated with this research will be publicly available at https://github.com/zxleong in the near future.

References

AltaRock. (2014). Phase 2.1 Report, Newberry EGS Demonstration. https://gdr.openei.org/files/774/Phase%202.1%20Report_4.24.14.pdf

Bondár, I., Myers, S. C., & Engdahl, E. R. (2014). Earthquake location. In M. Beer, I. A. Kougioumtzoglou, E. Patelli, & I. S.-K. Au (Eds.), Encyclopedia of Earthquake Engineering (pp. 1–18). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-36197-5_184-1

Chang, W., & McMechan, G. A. (1994). 3-D elastic prestack, reverse-time depth migration. GEOPHYSICS, 59(4), 597–609. https://doi.org/10.1190/1.1443620

Chen, Y., Saad, O. M., Savvaidis, A., Chen, Y., & Fomel, S. (2022). 3d microseismic monitoring using machine learning. Journal of Geophysical Research: Solid Earth, 127(3), e2021JB023842. https://doi.org/10.1029/2021JB023842

Cladouhos, T. T., Petty, S., Swyer, M. W., Uddenberg, M. E., Grasso, K., & Nordin, Y. (2016). Results from newberry volcano egs demonstration, 2010–2014. Geothermics, 63, 44–61. https://doi.org/10.1016/j.geothermics.2015.08.009

Cladouhos, T.T., Petty, S., Callahan, O., Osborn, W., Hickman, S., Davatzes, N. (2011a). The role of stress modeling in stimulation planning at the Newberry VolcanoEGS Demonstration project. In: Proceedings: Thirty-Sixth Workshop onGeothermal Reservoir Engineering, Stanford University, Stanford, California,January 31-February 2, pp. 630–637, SGP-TR-1191.

Cladouhos, T.T., Clyne, M., Nichols, M., Petty, S., Osborn, W.L., Nofziger, L. (2011b). Newberry Volcano EGS Demonstration stimulation modeling. GRC Trans. 35,317–322.

Das, I., & Zoback, M. D. (2011). Long-period, long-duration seismic events during hydraulic fracture stimulation of a shale gas reservoir. The Leading Edge, 30(7), 778–786. https://doi.org/10.1190/1.3609093

Davatzes, N.C., Hickman, S.H., (2011). Preliminary analysis of stress in the NewberryEGS well NWG 55-29. GRC Trans. 35, 323–332.

Dokht, R. M. H., Kao, H., Visser, R., & Smith, B. (2019). Seismic event and phase detection using time–frequency representation and convolutional neural networks. Seismological Research Letters, 90(2A), 481–490. https://doi.org/10.1785/0220180308

EIA. (2023). Electric power monthly—U. S. Energy information administration. Retrieved from https://www.eia.gov/electricity/monthly/epm_table_grapher.php

Fang, Y., Den Hartog, S. A. M., Elsworth, D., Marone, C., & Cladouhos, T. (2016). Anomalous distribution of microearthquakes in the newberry geothermal reservoir: Mechanisms and implications. Geothermics, 63, 62–73. https://doi.org/10.1016/j.geothermics.2015.04.005

Gajewski, D., & Tessmer, E. (2005). Reverse modelling for seismic event characterization. Geophysical Journal International, 163(1), 276–284. https://doi.org/10.1111/j.1365-246X.2005.02732.x

Geiger, L., 1912, Probability method for the determination of earthquake epicenters from the arrival time only, St. Louis Univ. Bull. 8, 60–71.

Häring, M. O., Schanz, U., Ladner, F., & Dyer, B. C. (2008). Characterisation of the Basel 1 enhanced geothermal system. Geothermics, 37(5), 469–495. https://doi.org/10.1016/j.geothermics.2008.06.002

Izadi, G., & Elsworth, D. (2013). The effects of thermal stress and fluid pressure on induced seismicity during stimulation to production within fractured reservoirs. Terra Nova, 25(5), 374–380. https://doi.org/10.1111/ter.12046

Karasözen, E., & Karasözen, B. (2020). Earthquake location methods. GEM - International Journal on Geomathematics, 11(1), 13. https://doi.org/10.1007/s13137-020-00149-9

Kriegerowski, M., Petersen, G. M., Vasyura-Bathke, H., & Ohrnberger, M. (2019). A deep convolutional neural network for localization of clustered earthquakes based on multistation full waveforms. Seismological Research Letters, 90(2A), 510–516. https://doi.org/10.1785/0220180320

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521(7553), 436–444. https://doi.org/10.1038/nature14539

Lei, Q., Gholizadeh Doonechaly, N., & Tsang, C.-F. (2021). Modelling fluid injection-induced fracture activation, damage growth, seismicity occurrence and connectivity change in naturally fractured rocks. International Journal of Rock Mechanics and Mining Sciences, 138, 104598. https://doi.org/10.1016/j.ijrmms.2020.104598

Li, L., Tan, J., Schwarz, B., Staněk, F., Poiata, N., Shi, P., Diekmann, L., Eisner, L., & Gajewski, D. (2020). Recent advances and challenges of waveform-based seismic location methods at multiple scales. Reviews of Geophysics, 58(1), e2019RG000667. https://doi.org/10.1029/2019RG000667

Lu, S.-M. (2018). A global review of enhanced geothermal system (Egs). Renewable and Sustainable Energy Reviews, 81, 2902–2921. https://doi.org/10.1016/j.rser.2017.06.097

Majer, E. L., Baria, R., Stark, M., Oates, S., Bommer, J., Smith, B., & Asanuma, H. (2007). Induced seismicity associated with Enhanced Geothermal Systems. Geothermics, 36(3), 185–222. https://doi.org/10.1016/j.geothermics.2007.03.003

Matzel, E., Templeton, D., Petersson, A., Goebel, M. (2014). Imaging the Newberry EGS site using seismic interferometry. Thirty-Ninth Workshop on Geothermal Reservoir Engineering, SGP-TR-202.

McMechan, G. A., Clayton, R. W., & Mooney, W. D. (1982). Application of wave field continuation to the inversion of refraction data. Journal of Geophysical Research: Solid Earth, 87(B2), 927–935. https://doi.org/10.1029/JB087iB02p00927

Mousavi, S. M., & Beroza, G. C. (2020). Bayesian-deep-learning estimation of earthquake location from single-station observations. IEEE Transactions on Geoscience and Remote Sensing, 58(11), 8211–8224. https://doi.org/10.1109/TGRS.2020.2988770

Mousavi, S. M., Ellsworth, W. L., Zhu, W., Chuang, L. Y., & Beroza, G. C. (2020). Earthquake transformer—An attentive deep-learning model for simultaneous earthquake detection and phase picking. Nature Communications, 11(1), 3952. https://doi.org/10.1038/s41467-020-17591-w

Muffler, P., & Cataldi, R. (1978). Methods for regional assessment of geothermal resources. Geothermics, 7(2–4), 53–89. https://doi.org/10.1016/0375-6505(78)90002-0

Nix, D. A., & Weigend, A. S. (1994). Estimating the mean and variance of the target probability distribution. Proceedings of 1994 IEEE International Conference on Neural Networks (ICNN'94), 55–60 vol.1. https://doi.org/10.1109/ICNN.1994.374138

Perol, T., Gharbi, M., & Denolle, M. (2018). Convolutional neural network for earthquake detection and location. Science Advances, 4(2), e1700578. https://doi.org/10.1126/sciadv.1700578

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

Ross, Z. E., Yue, Y., Meier, M., Hauksson, E., & Heaton, T. H. (2019). Phaselink: A deep learning approach to seismic phase association. Journal of Geophysical Research: Solid Earth, 124(1), 856–869. https://doi.org/10.1029/2018JB016674

Rutqvist, J., Rinaldi, A. P., Cappa, F., & Moridis, G. J. (2013). Modeling of fault reactivation and induced seismicity during hydraulic fracturing of shale-gas reservoirs. Journal of Petroleum Science and Engineering, 107, 31–44. https://doi.org/10.1016/j.petrol.2013.04.023

Schill, E., Genter, A., Cuenot, N., & Kohl, T. (2017). Hydraulic performance history at the Soultz EGS reservoirs from stimulation and long-term circulation tests. Geothermics, 70, 110–124. https://doi.org/10.1016/j.geothermics.2017.06.003

Shen, H., & Shen, Y. (2021). Array-based convolutional neural networks for automatic detection and 4d localization of earthquakes in hawai‘i. Seismological Research Letters, 92(5), 2961–2971. https://doi.org/10.1785/0220200419

Tarantola, A., & Valette, B. (1981). Inverse problems = Quest for information. Journal of Geophysics, 50(1), 159-170. Retrieved from https://journal.geophysicsjournal.com/JofG/article/view/28.

Templeton, D. C., Wang, J., Goebel, M. K., Harris, D. B., & Cladouhos, T. T. (2020). Induced seismicity during the 2012 Newberry EGS stimulation: Assessment of two advanced earthquake detection techniques at an EGS site. Geothermics, 83, 101720. https://doi.org/10.1016/j.geothermics.2019.101720

Tester, J. W., Anderson, B. J., Batchelor, A. S., Blackwell, D. D., DiPippo, R., Drake, E. M., Garnish, J., Livesay, B., Moore, M. C., Nichols, K., Petty, S., Nafi Toksoz, M., Veatch, R. W., Baria, R., Augustine, C., Murphy, E., Negraru, P., & Richards, M. (2007). Impact of enhanced geothermal systems on US energy supply in the twenty-first century. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 365(1853), 1057–1094. https://doi.org/10.1098/rsta.2006.1964

Tomac, I., & Sauter, M. (2018). A review on challenges in the assessment of geomechanical rock performance for deep geothermal reservoir development. Renewable and Sustainable Energy Reviews, 82, 3972–3980. https://doi.org/10.1016/j.rser.2017.10.076

Trugman, D. T., & Shearer, P. M. (2017). Growclust: A hierarchical clustering algorithm for relative earthquake relocation, with application to the spanish springs and sheldon, nevada, earthquake sequences. Seismological Research Letters, 88(2A), 379–391. https://doi.org/10.1785/0220160188

Van den Ende, M. P. A., & Ampuero, J. -P. (2020). Automated seismic source characterization using deep graph neural networks. Geophysical Research Letters, 47(17). https://doi.org/10.1029/2020GL088690

Waldhauser, F., & Ellsworth, W. (2000). A double-difference earthquake location algorithm: Method and application to the northern hayward fault, california. Bulletin of the Seismological Society of America, 90(6), 1353–1368. https://doi.org/10.1785/0120000006

Zang, A., Oye, V., Jousset, P., Deichmann, N., Gritto, R., McGarr, A., Majer, E., & Bruhn, D. (2014). Analysis of induced seismicity in geothermal reservoirs – An overview. Geothermics, 52, 6–21. https://doi.org/10.1016/j.geothermics.2014.06.005

Zhang, X., Zhang, M., & Tian, X. (2021). Real-time earthquake early warning with deep learning: Application to the 2016 m 6. 0 Central Apennines, Italy earthquake. Geophysical Research Letters, 48(5). https://doi.org/10.1029/2020GL089394

Zhu, T. (2014). Time-reverse modelling of acoustic wave propagation in attenuating media. Geophysical Journal International, 197(1), 483–494. https://doi.org/10.1093/gji/ggt519

Zhu, W., & Beroza, G. C. (2018). Phasenet: A deep-neural-network-based seismic arrival time picking method. Geophysical Journal International. https://doi.org/10.1093/gji/ggy423

Zhu, W., McBrearty, I. W., Mousavi, S. M., Ellsworth, W. L., & Beroza, G. C. (2022). Earthquake phase association using a bayesian gaussian mixture model. Journal of Geophysical Research: Solid Earth, 127(5). https://doi.org/10.1029/2021JB023249

Journal of Geophysical Research: Machine Learning and Computation

Supporting Information for

Microseismic Monitoring using Transfer Learning: Example from the Newberry EGS

Zi Xian Leong[1,†] and Tieyuan Zhu[1,2]

[1] Department of Geosciences, The Pennsylvania State University, University Park, PA, USA.

[2] EMS Energy Institute, The Pennsylvania State University, University Park, PA, USA.

†Currently at Chevron Technical Center, a division of Chevron U.S.A. Inc.

Contents of this file

Figures S1 to S4

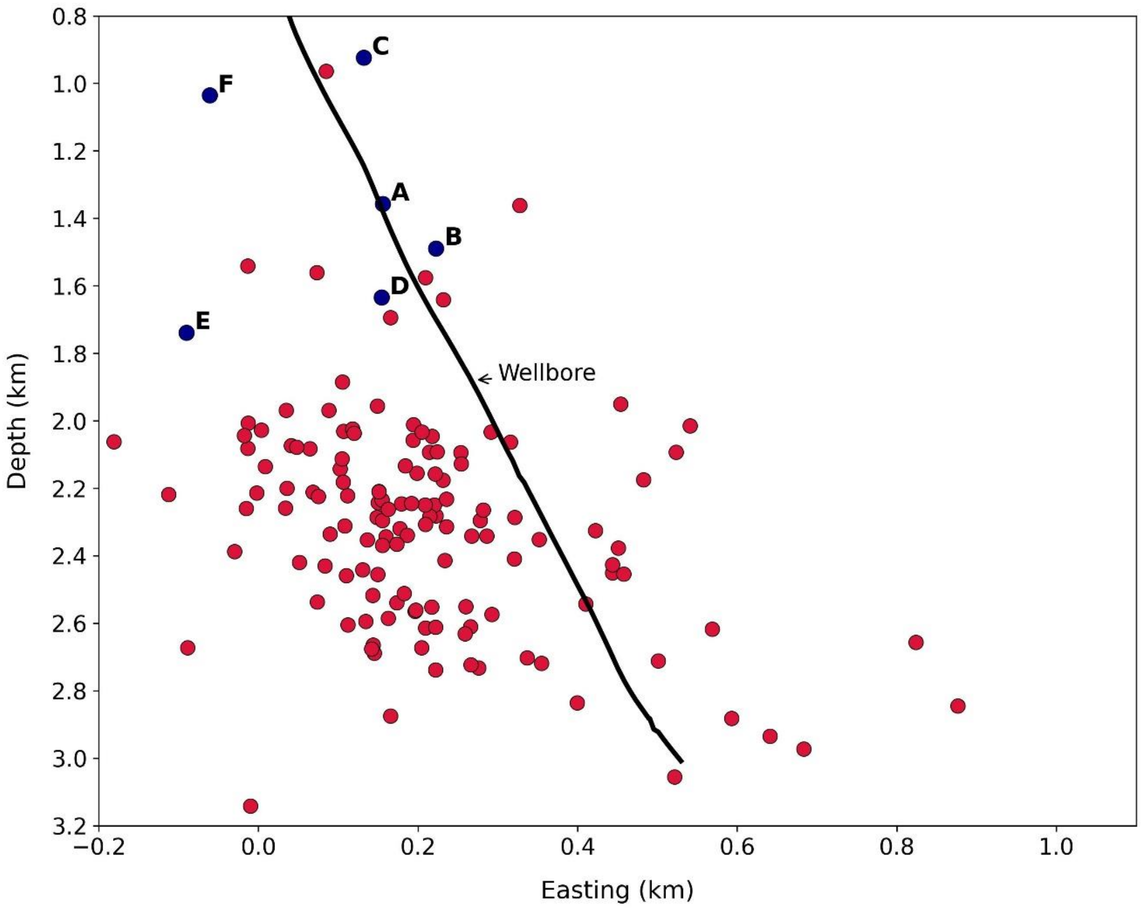

Figure S1. Predictions (n=142) from the first bin in the histogram of Figure 14. Here, we select six events (Events A, B, C, D, E, and F) to compare their first arrivals match between the synthetic and that of field. There is a noticeable lack of predictions around 1.8 km depth.

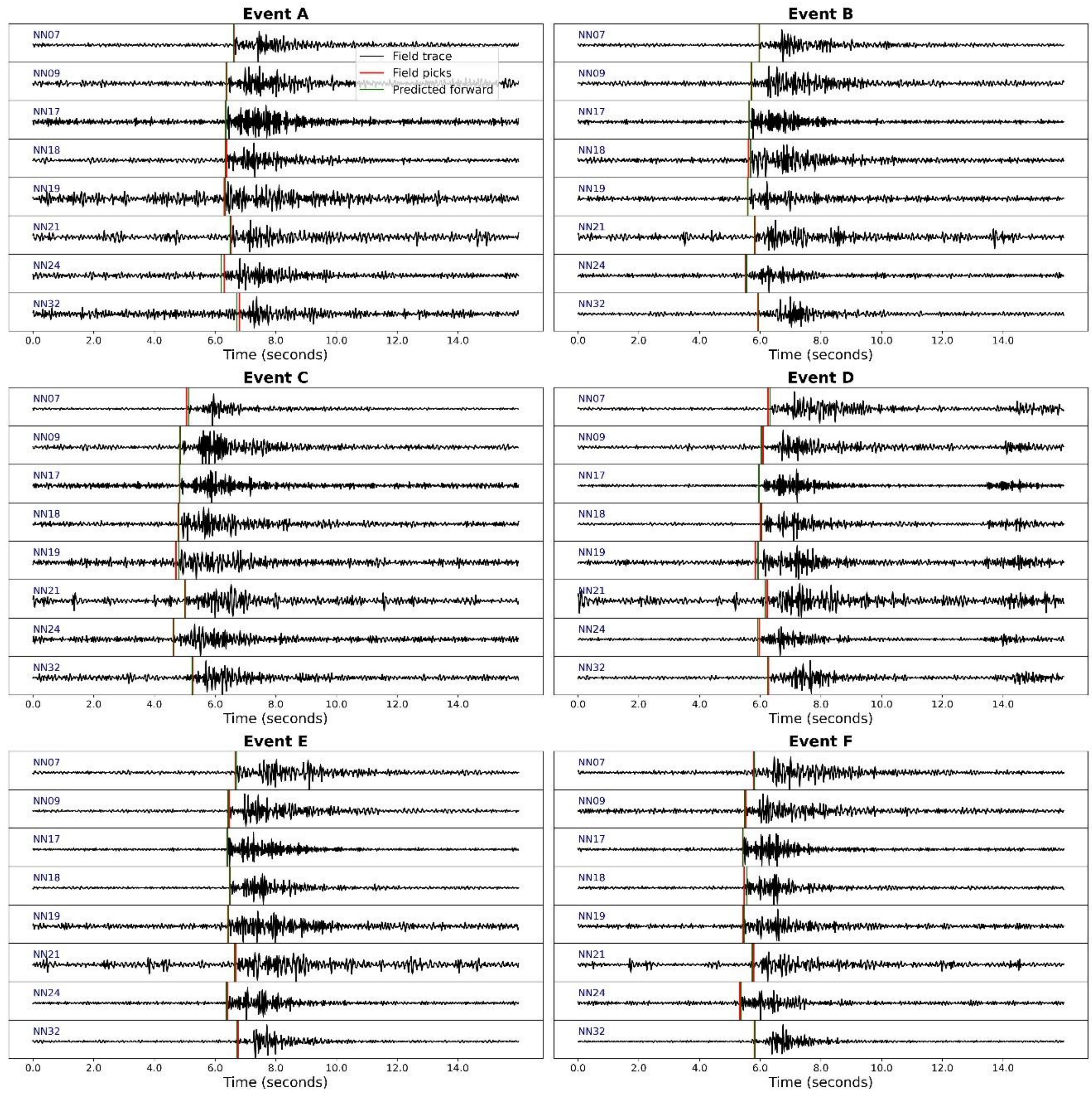

Figure S2. Comparison of synthetic (green vertical line) vs. field (red vertical line) first arrivals for selected events in Figure S1.

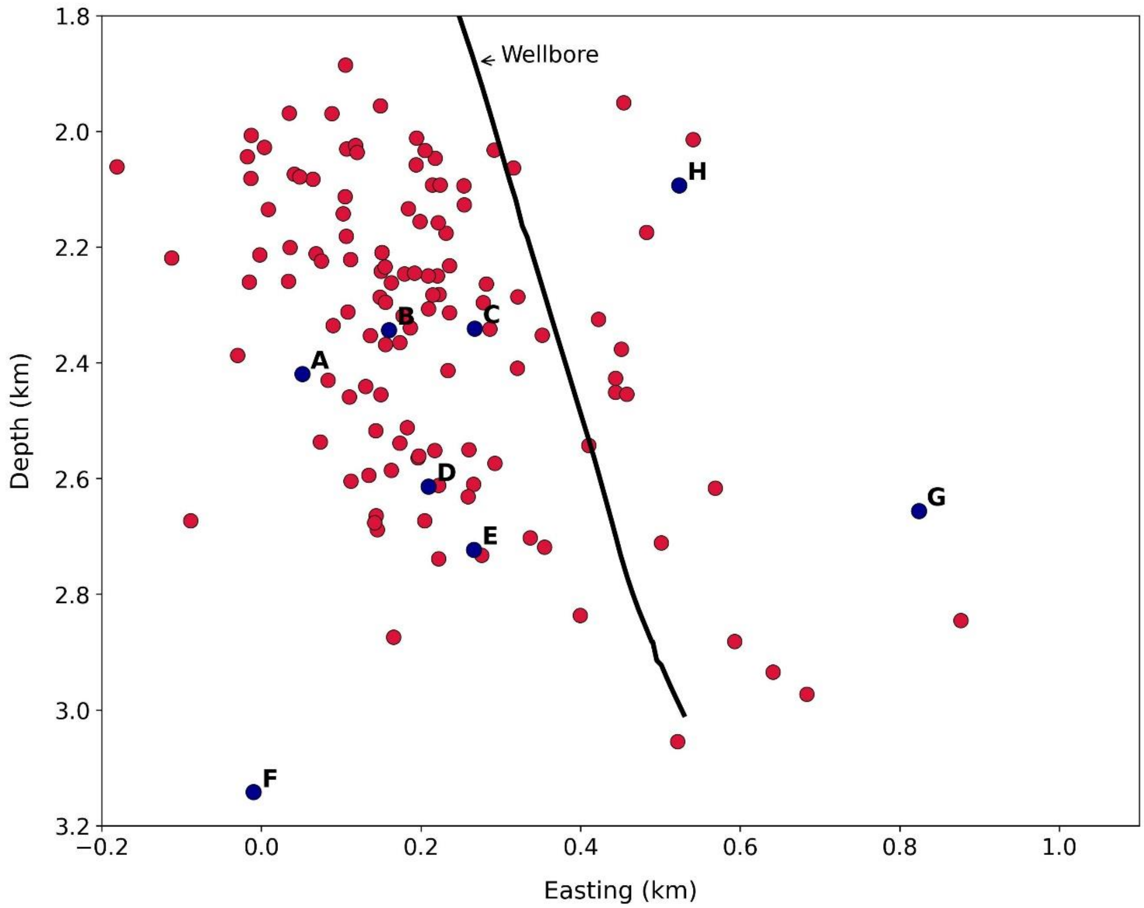

Figure S3. Reproduced from Figure 15 in the main text, but without labels. We randomly select eight example events (Event A, B, C, D, E, F, G, and H) for the comparison of synthetic vs. field first arrivals.

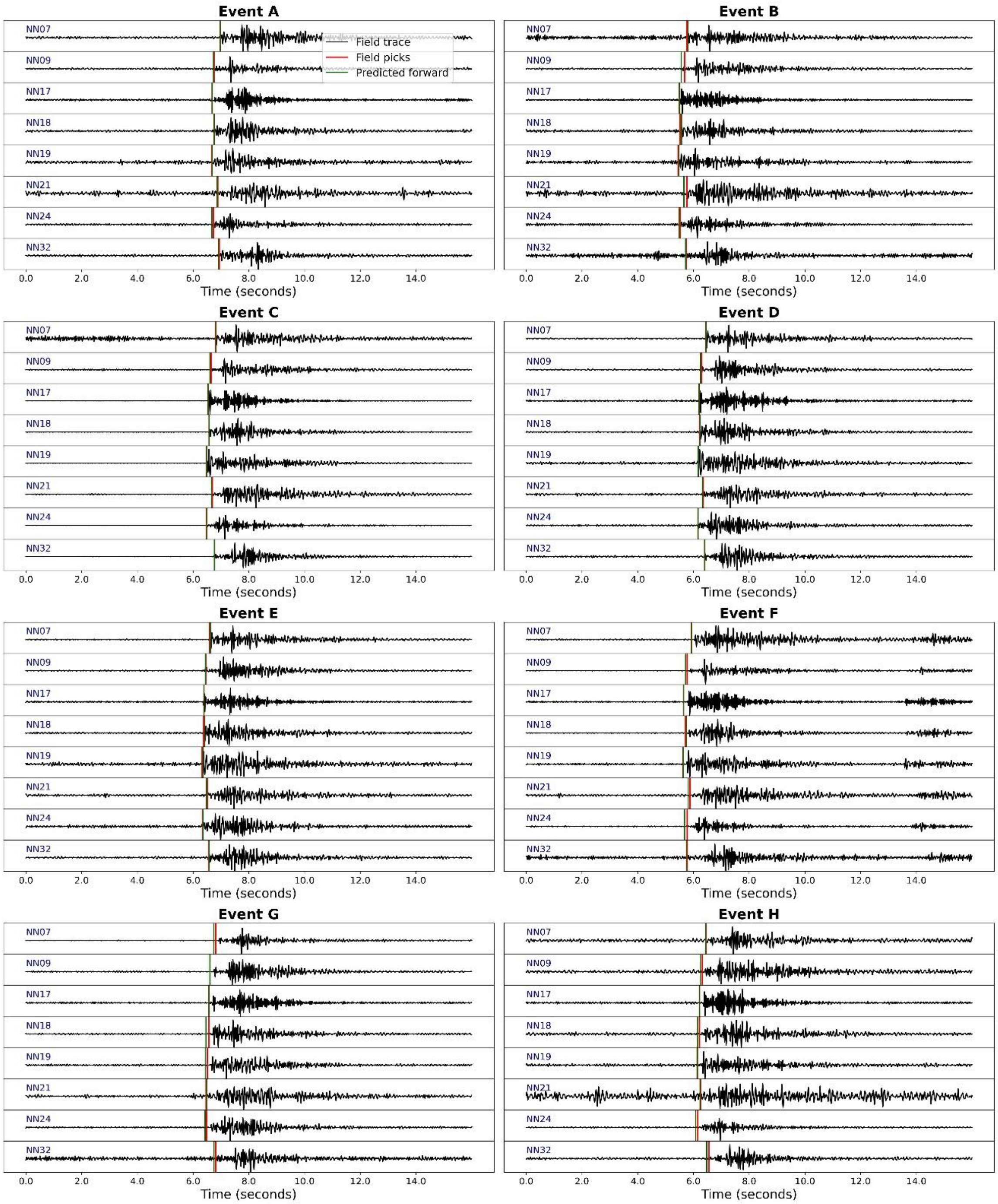

Figure S4. Comparison of synthetic (green line) vs. field (red line) first arrivals from the selected events in Figure S3.