



**HAL**  
open science

## Basin-Scale Machine Learning Prediction for Injection-Induced Seismicity in Oklahoma

Fnu Shamila Bolikoppa Palakshappa, Hideo Aochi, Zahra Derakhshandeh,  
Jonny Rutqvist, Mengsu Hu

► **To cite this version:**

Fnu Shamila Bolikoppa Palakshappa, Hideo Aochi, Zahra Derakhshandeh, Jonny Rutqvist, Mengsu Hu. Basin-Scale Machine Learning Prediction for Injection-Induced Seismicity in Oklahoma. 3rd International Conference on Coupled Processes in Fractured Geological Media: Observation, Modeling, and Application, Nov 2022, Berkeley (CA), United States. hal-04031710

**HAL Id: hal-04031710**

**<https://hal.science/hal-04031710v1>**

Submitted on 16 Mar 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## Basin-Scale Machine Learning Prediction for Injection-Induced Seismicity in Oklahoma

Fnu Sharmila Bolikoppa Palakshappa<sup>1)2)</sup>, Hideo Aochi<sup>3)4)</sup>, Zahra Derakhshandeh<sup>1)</sup>,  
Jonny Rutqvist<sup>2)</sup>, Mengsu Hu<sup>2)</sup>\*

1) California State University, East Bay, Hayward, USA

2) Energy Geosciences Division, Lawrence Berkeley National Laboratory, Berkeley, USA

3) BRGM, Orléans France

4) Ecole Normale Supérieure, CNRS 8538, PSL Research University, Paris, France

\*Corresponding author: mengsuhu@lbl.gov

### Abstract

Induced seismicity is a major concern for underground fluid injection, including injection of wastewater and CO<sub>2</sub> and H<sub>2</sub> for carbon neutrality. Machine learning (ML) appears to be promising for analyzing injection-induced seismic events (IISE) without explicitly solving complex multiphysics; however, applying ML for analyzing IISE in space and time has never been attempted. In this study, we apply ML to analyze IISE in Oklahoma at the basin scale. We applied a deep learning (DL) – Multilayer perceptron (MLP) model to correlate injection fluid volume and pressure with IISE at 53 areas of central and northwestern Oklahoma in different years. In the model, we implicitly considered the properties of geological formations (e.g., permeability, Young’s modulus, and strength) and fault locations by using a novel “neighboring” approach. In this approach, seismic events at each location correspond not only to the injection and seismic history of its own but also to those of the immediate neighbors. We use the MLP model to predict the injection-induced seismicity rate ( $\mu$ ) and the total seismicity rate ( $\lambda$ ) defined in an empirical model. We trained the MLP model by using 12-month injection volume and pressure, and historical values of  $\mu$  and  $\lambda$  of each location and its immediate neighbors, along with the coordinates. We then used the test data to assess the accuracy of the prediction. A Mean Square Error (MSE) of 0.003 was obtained for  $\mu$  and 0.033 for  $\lambda$ . We conclude that the new MLP model is promising for basin-scale IISE in space and time.

### 1. Introduction

Induced seismicity is a major concern associated with industrial underground fluid injection activities, including wastewater injection, injection for sequestering carbon dioxide and

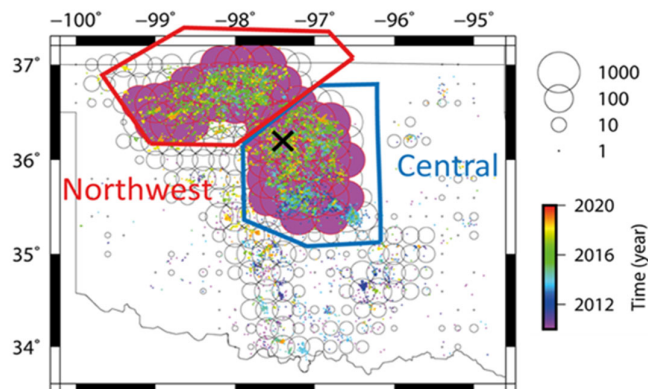


Figure 1. Fifty-three colored areas with at least 250 earthquakes of magnitude  $M_c \geq 2.3$  that were grouped into central and northwestern areas (Modified from (Aochi et al., 2021))

hydrogen storage. As a result of coupled hydro-mechanical (HM) processes, injecting fluid causes pore pressure to increase and shear strength to decrease in faults, thus may lead to reactivation of faults if certain criteria are satisfied (Ellsworth, 2013; Rutqvist et al., 2016). The induced seismic events due to wastewater fluid injection in Oklahoma is a typical example (Aochi et al., 2021). With sufficient data that is monitored in Oklahoma, we will use this massive amount

of data with a focus on the central and northwestern areas (as shown in Figure 1).

In this study, we will explore applications of machine learning to investigate injection-induced seismicity in space and time at the basin scale. We will first introduce the methodology including the architecture of ML model that we used, and the “neighboring” approach for effectively considering temporal and spatial efforts in order to implicitly consider the properties of geological formations (e.g., permeability, Young’s modulus, and strength) and fault locations. Then we will present the preliminary results of the injection-induced seismicity rate ( $\mu$ ) and the total seismicity rate ( $\lambda$ ).

## 2. Methodology

We developed a Multilayer perceptron (MLP) model to correlate injection volume with seismic rate. The proposed methodology involves using a sliding window approach where we take the past 12 months’ data of input (injection volume and pore pressure) which gives a large amount of data to train the ML model.

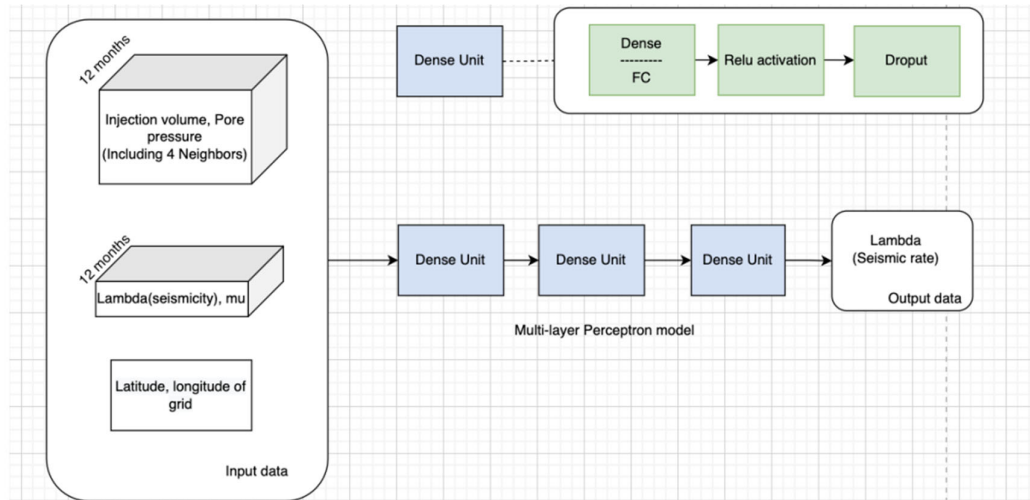


Figure 2. Proposed MLP model which is trained using the past 12 months of injection volume and pore pressure and seismicity ( $\lambda$ ,  $\mu$ ) for 5 coordinates (4 neighbors top, down, left, right). It also includes the latitude and longitude (i.e., the coordinates) as input for the model to learn coordinate-specific variations

The proposed MLP model consists of three layers—the input layer, the output layer, and the hidden layer (dense or Fully Connected (FC) layer), as shown in Figure 2. A dense or FC layer is a linear operation in which every input is connected to every output by weight (Figure 3). This layer gives sufficient flexibility to the model to solve possible non-linear relations between the input (injection volume, pore pressure) and output (seismic rate). Here we use the “sliding window” approach where we consider the past 12months’ data along with neighbor coordinates, and hence the input of 122 dimensions were fed to the MLP model.

As demonstrated in Figure 3, MLP architecture is a simple 4-layer feed-forward artificial neural network. The layers are an input layer having a fixed-length number of neurons (122), a dense layer followed by an activation layer Sigmoid, another dense layer followed by Sigmoid and Dropout, and the output layer with Sigmoid activation function. A Sigmoid activation inspects whether the result is negative or not. It transforms the negative values to zero and keeps the non-negative values as is. After the hidden layer, we used Sigmoid as activation before the information is sent to the output layer. This is to introduce non-linearity or even discontinuity to the linear output from the hidden layer. Sigmoid is the good function to use because it calculates the probability (ranging between 0 and 1) of the target output.

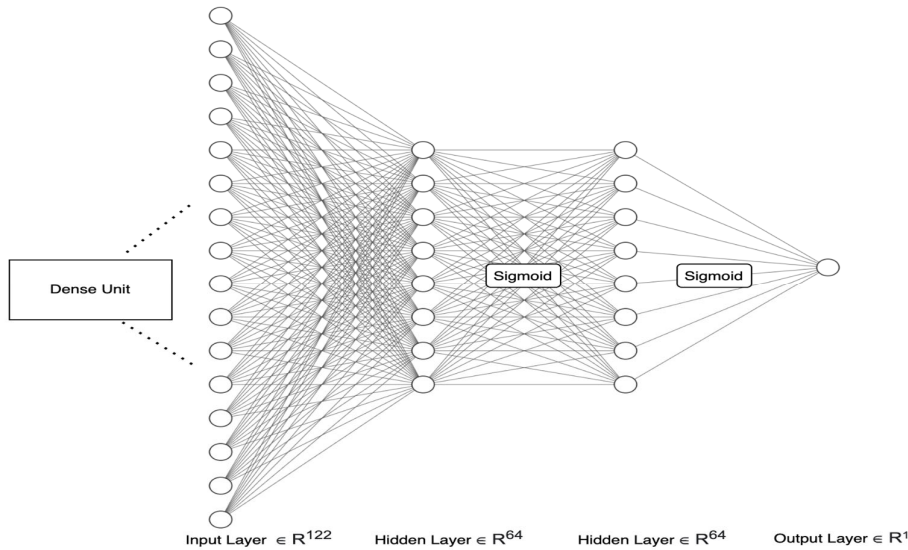


Figure 3. Multilayer Perceptron (MLP) model with one input, two hidden layers, and one output layer

### 3. Preliminary Results

As we validated the trained MLP model on the test set, the best inference obtained for the prediction of seismicity with provided Ground Truth is shown in Figures 4 and 5, where Ground Truth refers to the  $\mu$  and  $\lambda$  obtained from the Epidemic-Type Aftershock Sequence (ETAS) model (Aochi et al., 2021). The variable  $\mu$  is the injection-induced seismicity rate and  $\lambda$  is the total seismicity rate.

To evaluate our model, we used the Mean Square Error (MSE) as a performance metric:  $MSE = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2$ . The MSE is estimated between the model inference and the predicted seismicity. MSE is calculated by the sum of the square of prediction error, which is real output minus predicted output, and then divided by the number of data points.

The input features consist of 12-month injection volume and pressure, and historical values of  $\mu$  and  $\lambda$  of each location and its immediate neighbors along with the coordinates. Finally, we achieved an MSE of 0.003 for injection-induced seismicity rate ( $\mu$ ) and MSE of 0.033 for the total seismicity rate ( $\lambda$ ) on the test data.

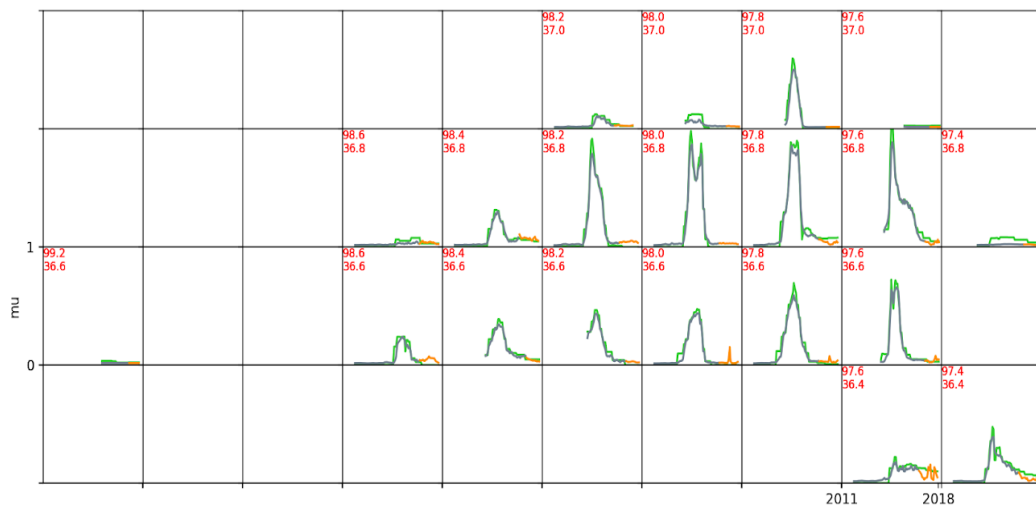


Figure 4. MLP model output for  $\mu$  with Green for Ground Truth, dark green for training predicted, and Orange for testing predicted

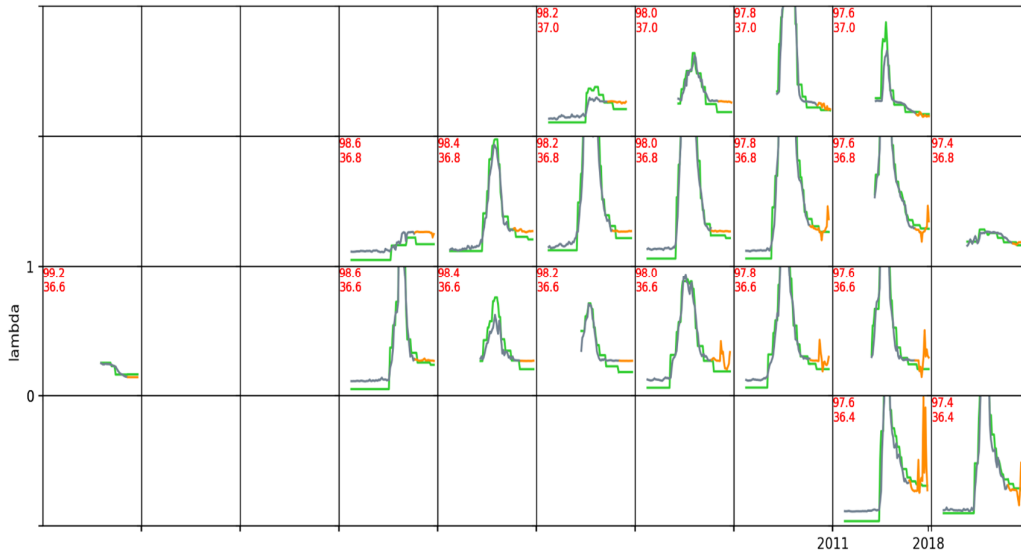


Figure 5. MLP model output for  $\mu$  where Green - Ground Truth, dark green - training predicted, Orange - testing predicted

## 4. Conclusions

Induced seismicity is a significant concern for underground fluid injection. Machine learning appears to be promising for analyzing injection-induced seismic events (IISE) without explicitly solving complex multiphysics. In this study, we made the first attempt to make use of machine learning to analyze IISE in Oklahoma at the basin scale. First, we took the data involving geometric and physical data with time evolution. The data was pre-processed for the new “neighboring approach for effectively considering temporal and spatial efforts in order to implicitly consider the properties of geological formations (e.g., permeability, Young’s modulus, and strength) and fault locations. Meanwhile, rectangular sub-grids of these 53 areas were generated based on the distribution of the geological features. Then we applied a deep learning (DL) – Multilayer perceptron (MLP) model to correlate injection fluid volume and pressure with IISE at 53 areas of central and northwestern Oklahoma in different years. We trained the MLP model by using 12-month injection volume and pressure, and historical values of  $\mu$  and  $\lambda$  of each location and its immediate neighbors along with the coordinates. We then used the test data to assess the accuracy of the prediction. Finally, we achieved a Mean Square Error (MSE) of 0.003 for  $\mu$  and 0.033 for  $\lambda$ . We conclude that the new MLP model is promising for basin-scale IISE in space and time.

## Acknowledgments

This research was conducted under the CSUEB Berkeley Lab Intern Program funded by Royalty funds provided by Lawrence Berkeley National Laboratory, operated for the U.S. Department of Energy under Contract No. DE-AC02-05CH11231, and was supported by the US Department of Energy (DOE), the Office of Nuclear Energy, Spent Fuel and Waste Disposition Campaign under Contract No. DE-AC02-05CH11231 with Berkeley Lab.

## References

- [1] Aochi, H., Maury, J., Le Guenan, T. (2021). How Do Statistical Parameters of Induced Seismicity Correlate with Fluid Injection? Case of Oklahoma. *Seismological Research Letter*, 92 (4): 2573–2590.
- [2] Ellsworth, W.L. (2013). Injection-induced earthquakes. *Science*, 341: 6142.
- [3] Rutqvist, J., Rinaldi, A.P., Cappa, F, et al. (2016). Fault activation and induced seismicity in geological carbon storage – Lessons learned from recent modeling studies. *Journal of Rock Mechanics and Geotechnical Engineering*, 8: 789-904.