

Imperial College London  
Department of Earth Science and Engineering  
MSc in Applied Computational Science and Engineering

Independent Research Project  
Project Plan

# Forecasting induced seismicity in Oklahoma

by

Zhiyong Liu

zl1220@ic.ac.uk

GitHub login: acse-liuzyon

Supervisors:

Dr. Stephen P. Hicks

June 2021

## **Abstract**

### **1 Introduction**

It is well known that humans can cause earthquakes through fluid injection and extraction. Fluid injection includes wastewater disposal and hydraulic fracturing, while fluid extraction includes oil, gas and groundwater. Such cases of induced seismicity have been well documented and confirmed in Oklahoma, where seismicity has been increased dramatically since 2010 (Ellsworth, 2013). In cases like Oklahoma, because the rate of earthquakes was very low before high-rate wastewater injection, it is relatively easy to determine the fundamental causes of induced seismicity. Therefore, due to the low background stress in intraplate regions, the human triggering signatures are clearly identifiable. However, in tectonically-active areas, it is more challenging to distinguish between natural and triggered seismicity. Large-scale big-data and statistical approaches are effective methods to predict the locations where human triggered seismicity may be expected. Big-data and statistical approaches will be crucial in separating natural causes from potential human triggering factors in these areas.

### **2 Literature Review**

Ellsworth (Ellsworth 2013) stated the current understanding of the causes and mechanics of human-induced earthquakes. It includes wastewater injection, emerging oil and gas recovery technologies, and other indirect induced activities, including deep fluid injection. Hincks et al. (Hincks et al. 2018) developed a Bayesian network to quantitatively evaluate correlations between well operational parameter, geological formation, and seismicity, finding that injection depth relative to crystalline basement was the most important parameter to seismic moment release. Wozniakowska and Eaton (Wozniakowska & Eaton 2020) used logistic regression to obtain a machine learning estimate of the seismogenic activation potential of each well and found that seismogenic activation potential is greatly influenced by injection depth.

### **3 Description of Problem and Objectives**

Although a causal relationship has been established between fluid injection and earthquakes, it is still not known why some areas are more susceptible to induced earthquakes. To solve this problem, big data and statistical methods are required. In this project, a unique and rich dataset of industrial activities from regions in California will be used to statistically evaluate and retrospectively forecast possible signatures of human-induced seismicity from existing high-quality earthquake in these regions. The database used contains well locations, geological metadata, monthly injection and extraction volumes. To test the existence of correlations between oil and gas activity and seismicity, some machine learning techniques will be developed to generate a model with predictive capabilities. Linear Regression and Logistic Regression will be used as one starting point to produce a simple model of earthquake occurrence. The main program will be coded in python, including using specific Python libraries such as Pandas, Scikit-Learn, statsmodels, matplotlib and some GIS modules.

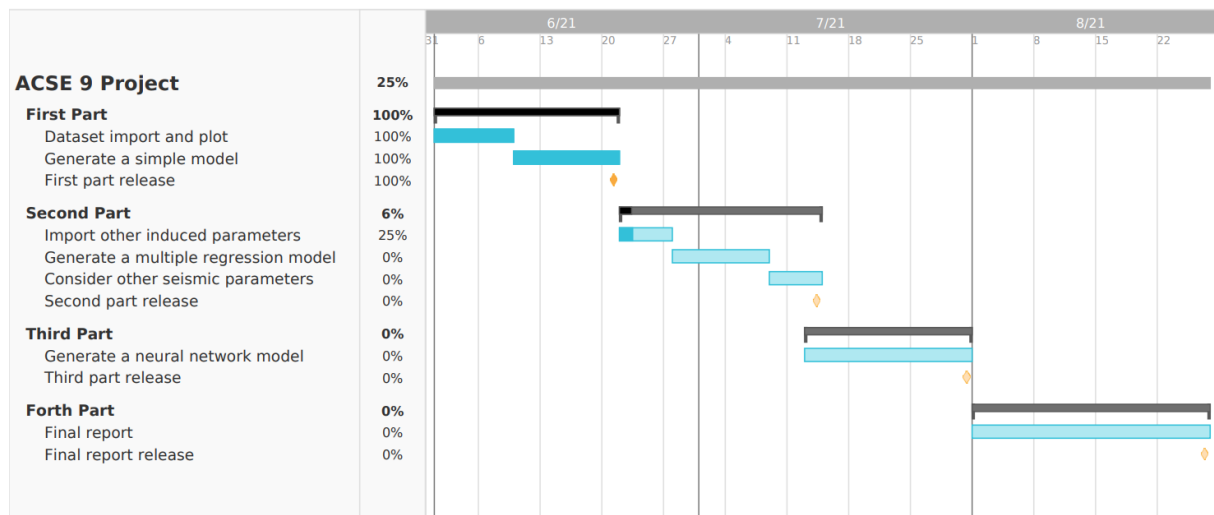


Figure 1: Gantt chart.

## 4 Progress to Date and Future Plan

In this project, three potential parameters to investigate correlations with earthquake activity, which are injection volume, well depth and geological formation being injected into. The parameters affected of earthquake activity considered are earthquake occurrence, maximum earthquake magnitude and earthquake moment rate. In the first part, dataset is successfully imported and plotted with geopandas. Then, as a starting point, injection volume will be concentrated on to find the correlation with the earthquake occurrence and produce a simple model. Here, some regression techniques such as linear or logistic regression will be used for the model generating. In the second part, other potential parameters (well depth and geological formation) will be introduced to consider together and a multiple regression model will be generated. Besides, other earthquake activity parameters (maximum earthquake magnitude and earthquake moment rate) will be investigated correlations with induced potential parameters. In the third part, a neural network model will be trained and generated with predictive capabilities. In the final part, the project report will be focused on.

## References

- Ellsworth, W. L. (2013), 'Injection-induced earthquakes', *Science* **341**(6142).
- Hincks, T., Aspinall, W., Cooke, R. & Gernon, T. (2018), 'Oklahoma's induced seismicity strongly linked to wastewater injection depth', *Science* **359**(6381), 1251–1255.
- Wozniakowska, P. & Eaton, D. W. (2020), 'Machine learning-based analysis of geological susceptibility to induced seismicity in the montney formation, canada', *Geophysical Research Letters* **47**(22), e2020GL089651.