

Bayesian Methods for Modeling Induced Seismicity

Bryan Azbill

March 2023

1 Introduction

Fluid induced seismicity is a well known consequence of subsurface fluid injections, operations typically for Wastewater disposal or fracking. Rock fracturing by fluid injection is well known as used for oil recovery but also has other purposes, such as enhancing geothermal systems. While most induced seismicity is small and of little consequence, seismic events large enough to cause potential damage and alert public concern can and have occurred. Therefore, it is prudent to take measure of induced seismic events to predict potential hazard during injection and judge when an injection procedure should be stalled or aborted.[1] Traffic light systems are used in practiced, where predictions of the largest potential magnitude of future seismic events, are compared to an acceptable limit which in turn tells the operator if they should slow, or stop injection. However, designing effective algorithms for this function is an ongoing process that needs development.[2] The purpose of this project is to develop bayesian algorithms that uses the locations and magnitudes of previously recorded seismic events to predicts the seismic events about to occur as a function of injection volume, with their distribution of magnitudes, the largest magnitude of the events, and updates to prediction after ever new recorded seismic event. This project will study induced seismicity records from the Utah "FORGE" project in particular, which is an enhanced geothermal systems research project that has been injecting in wells at Beaver County Utah and reading the seismic data with high precision. Additionally, another round of injections fore research is scheduled to be performed at the FORGE site in April 2023

2 Analysis

A current routine analysis of the problem starts with the following assumptions:

1. That by the time all induced seismicity is released for some injection Volume V , the cumulative induced seismic moment $M_0 = \sum_i(m_{0i})$ will be proportional to volume

$$dM_0/dV = c \tag{1}$$

2. That the relationship between the pdf of seismic events **moment magnitudes** m_i above some threshold M can be effectively approximated with an exponential distribution, known as the Gutenberg-richter law

$$P(m < M + x | m > M) = 1 - e^{-bx} \quad (2)$$

Actually, the Gutenberg richter law does not accurately describe induced seismicity, as the seismic moment probability distributions of induced seismicity often differs from the exponential at the tail of the distribution, because limiting sizes of the stimulated reservoir also puts a limit on the impacted fault area. [3] these assumptions overestimate the risk of induced seismicity. The goal of this project is to create a model that better accounts for the differences at the tail of the induced moment magnitude distribution. Nonetheless, It's useful start with derivation assuming GR-law and then proceed to try account for a more complex relationship. Suppose N_M is the number of earthquakes with magnitude greater than M occur under assumptions 1 and 2. Then for some threshold x the probability p that not any one of the them is greater than x above is

$$p = p(m_1 < M + x) * p(m_2 < M + x) * \dots * p(m_{N_M} < M + x) = (1 - e^{-bx})^{N_M} \quad (3)$$

$$\ln(p) = N_M \ln((1 - e^{-bx})) \quad (\text{apply log})$$

$$\ln(p) = -N_M e^{-bx} \quad (\ln(1+dx) \approx dx)$$

$$e^{-bx} = \frac{-\ln(p)}{N_M} \quad (\text{simplify expression})$$

$$x = \ln\left(\frac{N_M}{-\ln(p)}\right) \frac{1}{b} \quad (\text{apply ln, simplify})$$

again, N_M defines the number of events above magnitude N to include in the model, b the shape of the exponential, and p the probability that the largest magnitude occurred is bigger than x

3 Results

Three stages of fluid injection were carried out in the forge site during April of 2022. The cumulative induced seismicity vs volume is plotted in the chart below:

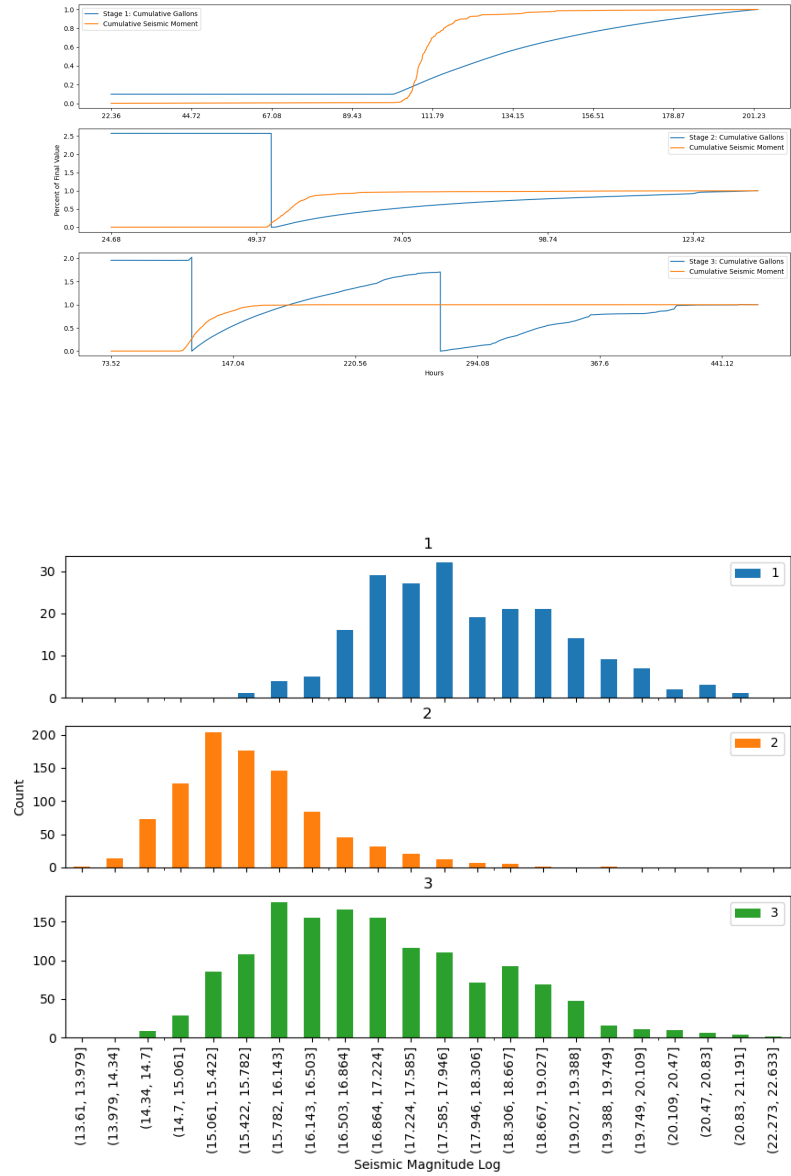


Figure 1: histograms of the seismic magnitude distribution

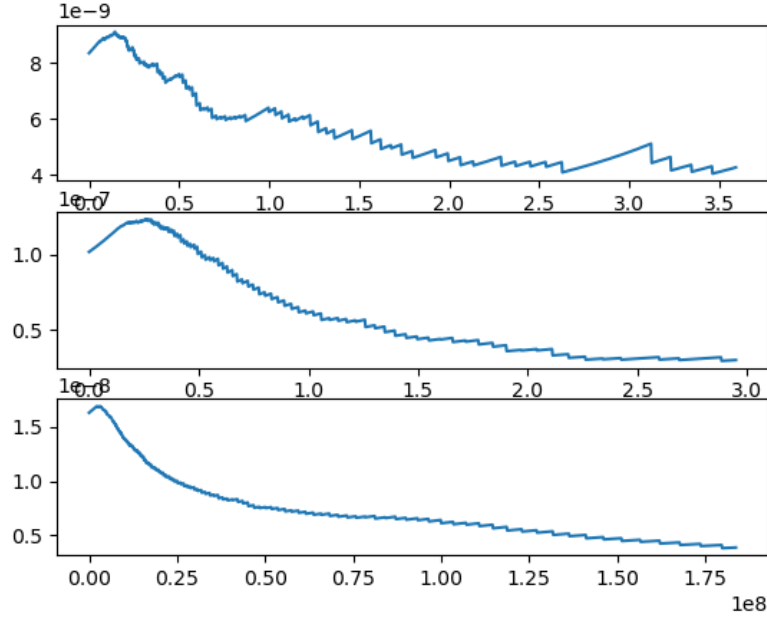


Figure 2: maximum likelihood b value vs M (minimum magnitude threshold)

The long tail of the distribution can be further confirmed by checking the b fit value as a function of minimum magnitude.

Originally, geoscientists included M into the distribution to discount that fewer small events can be detected than fits an exponential distribution. However, raising it much further provides an even better fit.

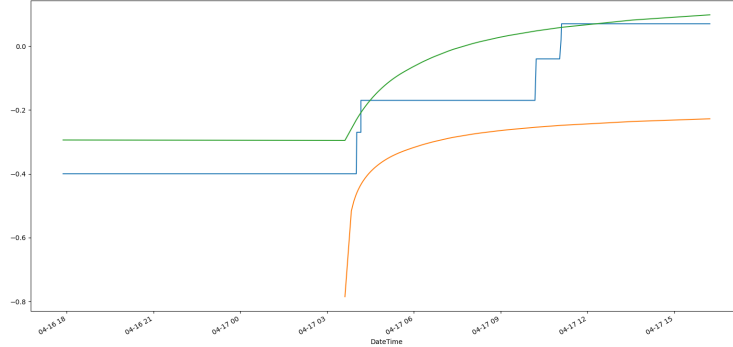


Figure 3: Fitting maximum magnitude prediction with a small vs large M value, and a ten percent probability of a larger maximum

Fitting to the maximum seismic event as a function of injection volume using a high M value and b value, matches the tail of the distribution much better, and therefore is accurate while the industry standard model, which doesn't account for the distribution tail, is not

4 Model Plan

This model uses the full data for the data it's predicting, but it's only appropriate for a model to only predict future maximum using past maximum events. Additionally, as I have learned in class, instead of raising the M value, the long tail of the distribution should be fit by updating the posterior distribution of b using a well designed prior. I want to write an algorithm that uses Bayesian methods on the prior of the b value distribution to can update the distribution after every past event. Additionally, it would be nice if studying the b value distribution can provide insight about why actual physics of induced seismicity produces. Ideally, the best statistical model should also relate to what's occurring in the subsurface, and could update prior involving probability distributions of things like rock stress and the size of faults that cross the reservoir. When relating to geomechanics, a different magnitude scale is used, called seismic moment, the random variable of which is the exponential of moment magnitude, making it a Pareto distribution. In seismic moment scale, magnitude is proportional to rupture area of the fault, and I would like to study seismic moment in order to gather insight about how the long tail of magnitude distribution might relate to the area of faults crossing the reservoir at the Forge site. I want to use the new model design to predict the seismicity that will be induced at the Utah FORGE site this April, but the raw seismic from that may or may not be processed into event magnitudes to validate by the end of the semester.

References

- [1] *Global review of human-induced earthquakes*. [Online; accessed 8. Mar. 2023]. Mar. 2018. DOI: 10.1016/j.earscirev.2017.07.008.
- [2] A. Mignan et al. “Induced seismicity closed-form traffic light system for actuarial decision-making during deep fluid injections”. In: *Sci. Rep.* 7.13607 (Oct. 2017), pp. 1–10. ISSN: 2045-2322. DOI: 10.1038/s41598-017-13585-9.
- [3] Pawel Urban et al. “Violations of Gutenberg–Richter Relation in Anthropogenic Seismicity”. In: *Pure Appl. Geophys.* 173.5 (May 2016), pp. 1517–1537. ISSN: 1420-9136. DOI: 10.1007/s00024-015-1188-5.