

# Reservoir Microearthquake Modelling Analysis: a Proof-of-concept Study and its Application to Injection Fluid-Induced Seismicity

Julius Marvin S. Rivera<sup>1</sup>, David E. Dempsey<sup>2</sup>

<sup>1</sup> Energy Development Corporation, One Corporate Center, Julia Vargas cor. Meralco Ave, Ortigas, Pasig City, Philippines

<sup>2</sup> Department of Engineering Science, The University of Auckland, 70 Symonds Street, Auckland Central, Auckland, New Zealand

[rivera.js@energy.com.ph](mailto:rivera.js@energy.com.ph), [d.dempsey@auckland.ac.nz](mailto:d.dempsey@auckland.ac.nz)

**Keywords:** microearthquake, microseismicity, reservoir modelling, induced seismicity, TOUGH2

## ABSTRACT

Microearthquakes (MEQs) occur when fluid is reinjected into the reservoir, raising the pressure in the vicinity of the injection well. The pressure build-up in the reservoir due to fluid injection decreases the rock yield strength, which causes shear failure, thus triggering a seismic event. This mechanism presents a further opportunity to use microseismicity as a means to calibrate reservoir parameters, particularly the active faults which tend to be the most conductive fluid flow pathways. The study aims to integrate MEQ modelling to the reservoir development workflow and to the calibration workflow to estimate the permeability of the formations and the faults.

The proof-of-concept study considers a synthetic induced seismicity model which represents an area where the fluid is being injected. Reservoir simulation is conducted to evaluate pressure migration through the reservoir for a given reservoir and fault parameters. The earthquake model uses the pressure change from the simulation to compute the average seismicity rate of the fault as well as the spatiotemporal evolution of the seismic events. Synthetic MEQ data is then generated from the earthquake simulation using the Poisson model, which serves as the data for calibration and inverse modelling. Synthetic inversion is then performed to estimate the permeability of both the reservoir and the fault using Markov Chain Monte Carlo (MCMC) sampling method. The study also includes the effects of variation in MEQ data and other uncertainties in the model in parameter estimation. The method developed in this study is then applied to an injection fluid-induced seismicity from a wastewater injection site.

## 1. INTRODUCTION

One of the most significant pieces of information obtained during reservoir characterization is the fluid-transport properties of rock, and understanding the flow of the fluid through the system is necessary for the development of a geothermal reservoir. Microearthquakes (MEQ), at these times, are used as a tool to assess the permeability of structures, to determine the geothermal reservoir boundaries, and to monitor migration patterns of the injection fluids (Pramono and Colombo, 2005). The data gathered through seismic monitoring support the management strategies in drilling and the field management of production and injection area (Sepulveda et al, 2013). The relationship between microseismicity and production and injection history has been an essential factor in enhancing conceptual models and reservoir simulations (Sewell et al., 2013).

Geothermal systems consist of complex physical processes of heat and mass transfer, and these systems involve deformation of solid rock matrix in a highly heterogeneous environment (O' Sullivan & Pruess, 2000). Thus, reservoir numerical modelling is usually undertaken to approximate the current physical state of the reservoir by mathematically characterizing the flow of fluid and heat in fractured porous media. Calibration is performed to improve the match between the reservoir model and the real system. Calibration is done by adjusting reservoir parameters so that the simulation will closely match the measured data from the field. Once a model is a suitable representation of the reservoir, it can be used in forecasting future responses. This study presents a further opportunity to use microseismicity as a tool for calibrating reservoir parameters; particularly, the permeability of the formation, and of the active faults that serve as major fluid pathways in the reservoir.

MEQ occurs when fluid is reinjected into the reservoir, raising the pressure in the vicinity of the injection well. The pressure build-up in the reservoir due to fluid injection decreases the rock yield strength and causes shear failure, thus triggering small earthquakes. The number of MEQs triggered, their location, and migration, provide information about the pressure changes and the nature of the fluid flow through the reservoir. The study aims to integrate MEQ data into the reservoir model development workflow and to the calibration workflow to estimate reservoir parameters, which in this study, are the permeability of the formation and the fault. This study builds on an earlier work by Rivera and Dempsey (2017). This study used synthetic examples to illustrate the reservoir model development workflow as a proof-of-concept. Some of these ideas were then tested on an induced seismicity case study in Arkansas.

## 2. METHODOLOGY

The study presents a synthetic example as the initial steps in this paper are a proof-of-concept. The study assumes a simple geothermal reservoir which focuses mainly on the injection area. It is assumed that there is a fault that lies somewhere within the reservoir, which serves as the major pathway of the fluid. As the fluid flows through the fault, the pressure changes along the fault that causes rock failure and triggers seismicity.

This paper follows a general workflow for the development of the synthetic induced seismicity model. Figure 1 shows the workflow, which includes the synthetic model development and its sub-processes together with the inverse modelling process. The process starts with the creation of a synthetic numerical model which replicates the physical characteristics of the system. The numerical model includes the model grid and all the reservoir parameters needed for the simulation. TOUGH2 simulation is performed to generate the simulated pressure evolution in the reservoir due to fluid injection. The pressure output from the

synthetic model is the input in the earthquake simulation that estimates the number of events triggered along the fault at a given period. The spatiotemporal evolution of the seismic events is used to compute for the average seismicity rate of the fault. The seismicity rate is then used to generate MEQ data using a Poisson point process. The synthetic MEQ data generated from the synthetic model is used in inverse modelling to estimate reservoir parameters, specifically the permeability of the reservoir and the fault.

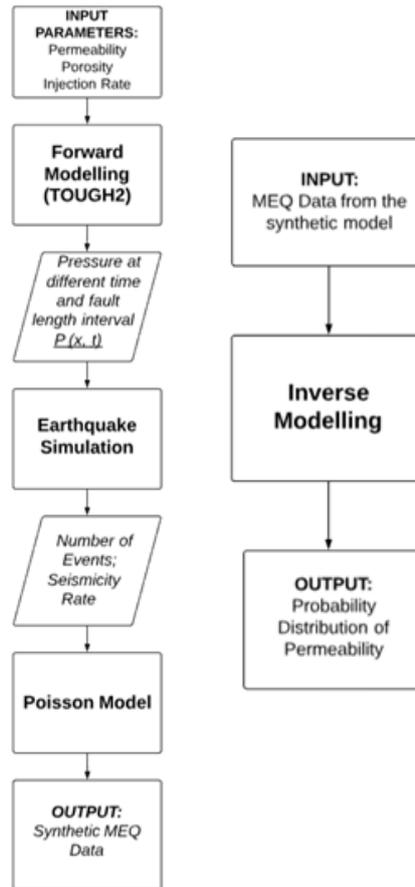


Figure 1: General workflow of synthetic model development and inverse modelling using MEQ data.

## 2.1 Two-Parameter Synthetic Model Development

### 2.1.1 Model Grid Generation

The first step in the development of the synthetic model is the generation of the grid for flow simulation. A simple rectangular grid with a dimension of 10.4 km x 10.4 km x 1 km is created and assumes an injection well at the center with fluid injecting at a constant rate. The model assumes that there is a fault in the reservoir located directly at the right of the injection well. Initially, the dimension of the grid blocks used in the model is 400 m x 400 m x 1000 m. It is known that the dimensions of the blocks play a crucial role in the efficiency and accuracy of the output on running both the forward and inverse models using TOUGH2, thus making the dimension of the blocks smaller will improve the simulation. However, it will be time-consuming if all blocks are of small size, mainly if the study is only concerned with the output of the blocks near the injection well. PyTOUGH gives a mechanism for altering the dimension and resolution of the grid, together with optimizing grid structure (O’Sullivan et al., 2013). Using PyTOUGH, several grid refinements are performed with dimensions of 200 m and 100 m towards the center of the grid near the injection well. Figure 2a shows the model grid with refinement together with the location of the fault and the injection well.

In this study, the fluid is assumed to flow radially in x and y directions from the center of the grid. There is an inherent symmetry in the x-axis given that the permeability of the reservoir and the fault are different. Thus, the model will consider the upper half of the grid to reduce the number of grid blocks in the simulation (Figure 2b). A model with less number of grid blocks aids in reducing the time for inverse modelling since it involves running several forward models.

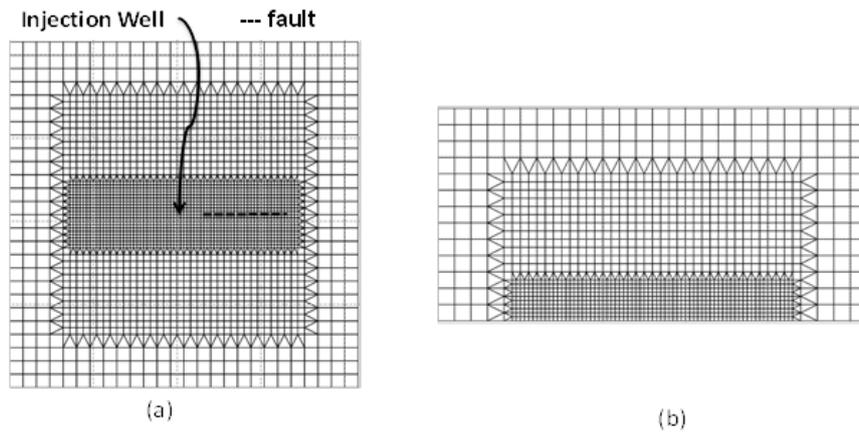


Figure 2. a) Model grid of two-parameter model; (b) upper half of the grid used in the simulation.

2.1.2 Forward Modelling and Pressure Output

TOUGH2 input file is prepared before performing the simulation of the model. In this study, PyTOUGH is used as the pre-processor to create the input file. A huge volume is assigned to the boundary blocks to assume an open boundary condition. Two rock types are added to the input file with isotropic permeability of  $1 \times 10^{-15} \text{ m}^2$  and  $1 \times 10^{-14} \text{ m}^2$ , which corresponds to the reservoir permeability and fault permeability, respectively. The porosity assigned for both rock types is 0.2. A generator injecting at a constant rate of 0.05 kg/s with a temperature of 50°C is placed at the bottom center of the grid block. The initial pressure used in the model is 0.1 MPa and 50°C. The initial pressure might be very low compared to actual pressure in geothermal fields, but this is already sufficient for the model as the study focuses only on changes of the fluid pressure. The initial temperature of the reservoir is the same as the injection temperature to eliminate thermal effects in the simulation and focus only on fluid pressure effects. The summary of the input parameters of the TOUGH2 input file is shown in Table 1. TOUGH2 simulation is performed with an end time of 5 years and time steps at every quarter of a year. Figure 3 presents the resulting pressure change within the reservoir at the end of 5 years.

Table 1. TOUGH2 Input data for the two-parameter synthetic model

Input Parameters	Values
Reservoir Permeability ( $\text{m}^2$ )	$1 \times 10^{-15}$
Fault Permeability ( $\text{m}^2$ )	$1 \times 10^{-14}$
Porosity (reservoir and fault)	0.20
Injection rate (kg/s)	0.05
Injection Temperature ( $^{\circ}\text{C}$ )	50
Run end time (yrs)	10

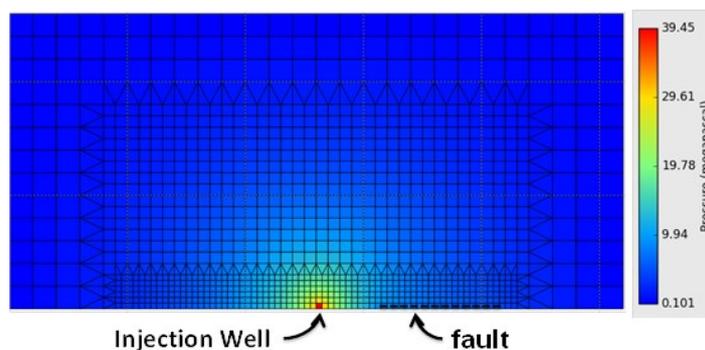


Figure 3. Pressure output of the fault and the reservoir of the two-parameter model after five years

The pressure output from the simulation is used to estimate the number of seismic events triggered along the fault. The location and the length of the fault are essential in estimating the pressure change along the fault. Table 2 summarizes the fault parameters used in the model. The fault was divided into 200 equal segments, and the coordinates of each point were computed given the fault parameters. The coordinates of the fault segments were then used to interpolate the pressure at each segment of the fault. The

procedure was done for each time step of the simulation to generate a pressure array for all the fault segments over time. This array of pressure at different fault segment over time is used to generate the average spatiotemporal seismicity rate of the fault.

**Table 2. Fault parameters of the two-parameter synthetic model**

Fault Parameters	Values
Fault Length (m)	2000
Fault Azimuth	90°
Distance from the injection (m)	1,000
Angle of Fault from Injection (from north)	90°

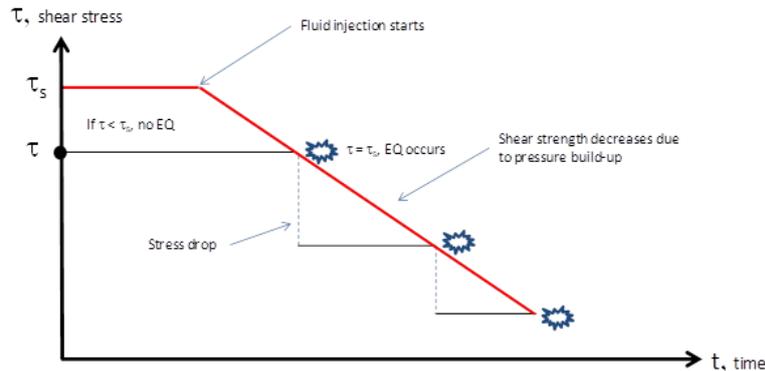
### 2.1.3 Forward Modelling and Pressure Output

The seismicity rate is estimated based on the pressure change on the fault, which affects the shear strength of the formation. Shear strength is the maximum shear stress the formation can handle before it results in an unstable slip, which could potentially trigger an earthquake. In induced seismicity, the shear strength changes depending on the fluid pressure in the fault. This relationship is described by the modified Mohr-Coulomb criterion equation given by:

$$\tau_s = f_s(\sigma_n - p) \tag{1}$$

Where  $\tau_s$ ,  $f_s$ ,  $\sigma_n$ , and  $p$  are the shear strength, coefficient of friction, normal stress, and fluid pore pressure, respectively. Based on the equation above, as the pressure build-up, the shear strength of the fault decreases thus promoting rock failure which in turn triggers a small earthquake. This is on the assumption that there is no significant change in the normal stress.

The relationship between shear stress and shear strength provides information on seismicity (Rivera and Dempsey, 2017). Figure 4 shows a simple diagram of how fluid injection induced seismic events over time. The diagram shows a fault with initial shear stress  $\tau$  and shear strength  $\tau_s$ . A fault remains stable and will not trigger any earthquake if  $\tau < \tau_s$ . As the fluid is injection into the reservoir, the pressure increase will result in a decrease in the shear strength of the fault. Once  $\tau = \tau_s$ , the slip will cause rock failure and triggers an earthquake. The pressure required to trigger the first earthquake in the fault is called the critical pressure ( $P_{crit}$ ) of the fault. After a seismic event, the shear stress drops to a point where the force in the fault does not cause any slippage or motion. As the fluid is continuously injected into the reservoir, the shear strength continues to decrease thus seismicity continues to occur.



**Figure 4. Diagram of induced seismicity due to fluid injection**

The average number of events ( $N$ ) triggered within the fault length ( $L$ ) at a given time ( $t$ ) is computed first before the calculation of average seismicity rate using the following equation (Dempsey and Suckale, 2017).

$$N = k \int_0^t \int_0^L [\Delta\tau(x, t) - \Delta\tau_s(x, t)] \delta(\tau(x, t) - \tau_s(x, t)) dx dt \tag{2}$$

The  $k$  in the equation is a constant of proportionality that accounts for other relationships between stress and microseismicity and can also be used as a parameter to scale the integrated seismicity rate for the number of earthquakes being modelled. The first part of the equation shows the difference in the spatiotemporal stress and the strength evolution along the fault which are given by:

$$\tau(x, t) = \tau(x, 0) + \int_0^t \Delta\tau(x, t') u\left(\frac{\tau_s(x, t')}{\tau(x, t')} - 1\right) dt' \tag{3}$$

$$\tau_s(x, t) = \tau_s(x, 0) + \int_0^t \Delta\tau_s(x, t') dt' \tag{4}$$

where  $u(x)$  is the Heaviside step function. The second part of the equation  $\delta(x)$  is the Dirac delta function, which serves as a “switch” that makes the equation equal to zero (no earthquakes triggered) when the shear stress is not equal to the shear strength. The Dirac delta function is explained using the following equation:

$$\delta(x) = \begin{cases} 0, & x \neq 0 \\ \infty, & x = 0 \end{cases} \quad \text{and} \quad \int_{-\infty}^{\infty} \delta(x) dx = 1 \quad (5)$$

Given the relationship between the shear strength and pressure in equation (1), the number of events is computed based on the change in pressure in the fault in relation to its critical pressure. The critical pressure is set to 0.1 MPa equal to the initial block pressure set in the model grid to allow triggering of seismic events at low overpressure. The number of seismicity triggered is computed using the pressure output generated from the forward model. The seismicity rate of the fault is the number of seismic events triggered at a given time interval over time. It is defined as  $\lambda(t)$  and is given by (Dempsey and Suckale, 2017):

$$\lambda(t) = k \int_0^t [\dot{t} - \dot{t}_s] \delta(\tau - \tau_s) dx$$

$$\dot{t} = \frac{\partial \Delta \tau}{\partial t} u\left(\frac{\tau_s}{\tau} - 1\right), \quad \dot{t}_s = \frac{\partial \Delta \tau_s}{\partial t} \quad (6)$$

To simplify, the derivative of equation (2) at a given time interval generates the seismicity rate of the fault.

#### 2.1.4 Poisson Model and MEQ Generation

Given the technology we have today, the occurrence of an earthquake cannot be predicted ahead of time. There is a limitation in the understanding of the stress changes and rock characteristics that gives the element of randomness to the time and location of the seismic events. The inability of the existing technology to comprehend the heterogeneous environment underneath the surface of the Earth leads to the assumption of the randomness of earthquakes.

In this study, the uncertainty in the earthquake triggering is described by a Poisson point process because it satisfies the statistical model conditions. These are:

1. Earthquake occurrence is known to be random, and each occurrence is counted as one event.
2. There should exist an average number of events at a given time interval, which follows the stationary process.
3. The occurrence of each event should be independent of each other.

The induced earthquake triggering in this study follows a discrete Poisson distribution to generate random events at different time intervals. Figure 5 shows one of the generated MEQ datasets using the seismicity rate of the synthetic model. This MEQ data is used in the inverse modelling to estimate the permeability of the reservoir and the fault. Even though there is a single value of average seismicity rate calculated in the synthetic model, there is no single set of microearthquake data that accurately represents the computed average seismicity rate of the fault over time (Rivera and Dempsey, 2017). Each simulation run generates a different set of microearthquake data. Given this information, the inversion process presents a probability distribution for the reservoir and fault permeability instead of a single distinct permeability combination.

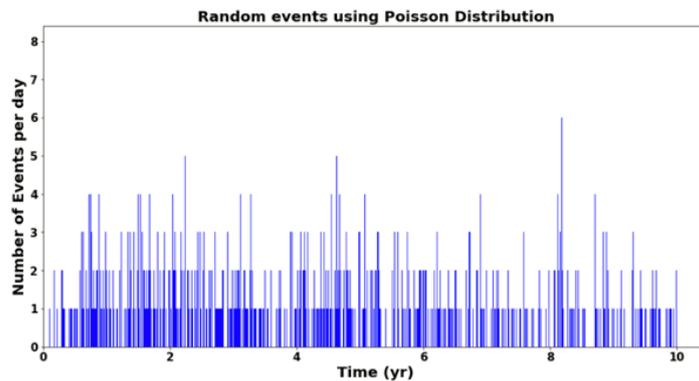


Figure 5. Synthetic MEQ data generated from the two-parameter synthetic model

## 2.2 Inverse Modelling

### 2.2.1 Model Calibration Process

Calibration is done by adjusting the permeability value of the reservoir and the fault to match approximately the occurrence times of earthquakes using the MEQ data generated from the synthetic model. This was done by running several forward models to generate possible values of reservoir and fault permeability. It is expected that no single set of reservoir parameters matches exactly the generated MEQ data due to the uncertainty in earthquake triggering. Thus, the permeability values presented in this study is in the form of probability distribution showing a spread of most likely values of the parameter.

The calibration starts with the preparation of the synthetic model with initial “guess” permeability for both reservoir and fault permeability while keeping other parameters constant. The values of the parameters are similar to the synthetic model to check the capability of the inversion method to estimate the permeability. The simulation is then performed using TOUGH2 to generate pressure evolution over time along the fault. The pressure is then used to calculate the average seismicity rate of the fault. The

seismicity rate is compared to the generated MEQ data and is scored using log-likelihood distribution (LLK) for nonhomogeneous Poisson process given by the equation (Lindqvist and Taraldsen, 2013):

$$LLK = \sum_{i=1}^n \log(\lambda(t_i, \theta)) - \int_0^{t_n} \lambda(u, \theta) du \quad (7)$$

In this equation,  $\theta$  is the parameter value in the seismicity rate model  $\lambda$  being estimated and constrained, which for this study is the permeability of the reservoir and the fault. This statistical method is used extensively as a model for events occurring in time (Dempsey and Suckale, 2017). The generated seismicity rate from the different models is then compared to the MEQ data, and the LLK values for each seismicity rate are plotted to come up with a likelihood distribution, and this summarizes the estimated permeability of the fault in a probabilistic sense.

**2.2.2 Parameter Estimation using Markov Chain Monte Carlo (MCMC) Method**

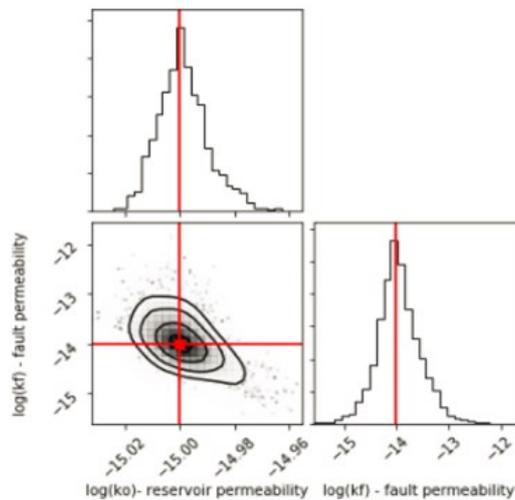
In the previous study done by Rivera and Dempsey (2017), in which one-parameter model is used as a synthetic model, the permeability values used as “guess” permeability for each simulation comes from a range of permeability values. This method, known as “grid search” method, performs all combination from a set of parameter values, and more sample size creates a better resolution of the probability distribution. However, this process is computationally expensive and time-consuming, especially when dealing with models with multiple unknown values to estimate. Thus, this study carries out a better sampling method with the use of Markov Chain Monte Carlo (MCMC). MCMC is considered a better sampling method than a grid search method as it reduces the computational expense, especially for a multi-dimensional probability distribution, since a smaller number of iterations are necessary to attain the likelihood distribution with the same level of accuracy (Foreman-Mackey, Hogg, Lang, & Goodman, 2013).

MCMC is a computer-driven sampling method which is a combination of Monte Carlo method and Markov Chain property. Monte Carlo obtains numerical results and estimates distribution properties by repeated random sampling from the distribution. The Markov Chain property is integrated into MCMC to include a unique sequential process in which the results of each random sample are used to generate the next random sample, thus creating a “chain” (van Ravenzwaaij, Cassey, & Brown, 2016). MCMC is useful in probabilistic data analysis, especially in Bayesian inference, and several studies have applied the MCMC method to calibrate models of earthquakes and seismicity (Debski, 2008; Stuart, Yang, Minkoff, & Pereira, 2016; Dempsey & Suckale, 2017).

MCMC sampling method is performed in this study using the Python module “emcee” developed by Foreman-Mackey, Hogg, et al. (2013). MCMC, in its most general terms, draws samples from the posterior probability density function, uses these samples to generate a new posterior then compare the latest value to the previous. If the value is accepted, it is used to the next iteration. If the value is rejected, it retains the last value and uses it again on the iteration process. The sampling continues until it converges to a stationary set of samples from the distribution (Foreman-Mackey, Hogg, Lang, & Goodman, 2013; van Ravenzwaaij, Cassey, & Brown, 2016). These values are like “walking” randomly from the initial guess value towards the target value.

**3. RESULTS AND DISCUSSIONS**

The inversion process performed in this study uses the Python “emcee” module with two-dimensional vectors representing the reservoir and fault permeability, respectively. Ten initial values (called “walkers”) were used in this synthetic inversion and are set to run for 1,000 iterations or until convergence is achieved. The initial values for each walker came from  $10^{-14}$  to  $10^{-16}$  m<sup>2</sup> for reservoir permeability and from  $10^{-13}$  to  $10^{-15}$  m<sup>2</sup> for fault permeability (Wallis, Moon, Clearwater, Azwar, and Barnes, 2015). Moreover, these priors assume a uniform distribution. A corner plot using Python “corner” module, which shows the likelihood distribution of each parameter together with the contour plot of the two-dimensional vectors, is shown in Figure 6.



**Figure 6. Corner plot for the likelihood distribution of the two-parameter model**

It can be observed from the plot that the actual permeability values used in the synthetic model (shown by the red marker) lie within the range of most likely values of permeability for both parameters which suggest a well-calibrated model. Also, the maximum likelihood estimates for both reservoir and fault permeability are accurately estimated by the inversion process. The range of most likely values for the reservoir permeability is quite narrow, which suggest minimal uncertainty in the permeability estimation. One

good reason for this is the number of MEQ data used in the calibration. A large number of MEQ data used during calibration provides a better match between the model and the actual data (Rivera and Dempsey, 2017). Also, the reservoir permeability controls the “delay” or the “arrival” of the earthquake triggered along the fault, and it has been entirely constrained by the inverse method presented in this study.

A wide range of likelihood distribution, as shown in the corner plot, can be observed for the fault permeability, which gives more considerable uncertainty in the estimation of the permeability. The actual value still lies between the range of values. A good explanation for a broader range of most likely values for the fault permeability is that the inverse method cannot restrict the fault permeability value since more realizations match the pressure change along the fault. The length of the fault and the location of seismic events contribute to the number of permeability combination that matches the synthetic MEQ data. Increasing the fault length also increases the occurrence of random earthquakes, thus adding more uncertainty in the estimation.

### 3.1 Effect of Fault Segmentation in Parameter Estimation

Based on the result of the two-parameter model presented above, a wide range of most likely values is observed when many MEQ data is used in estimating the permeability of a relatively long fault. Limiting the number of earthquakes calibrated in a given fault length improves the likelihood distribution, which can be done by discretizing the fault into smaller segments and performing inversion at each segment. In this section, additional simulations using MCMC were performed using the same model to demonstrate the effect of discretizing a fault into segments in the inversion. The fault is then divided into equal segments, and the seismicity rate on each segment is then calculated based on the pressure evolution from the reservoir simulation. Each segment is calibrated by matching the generated seismicity rate that corresponds to the MEQ data at a given segment. The log-likelihood values for each segment are then added to come up with the total log-likelihood for the whole fault. The inversion is performed using 1,000 iterations or until convergence is achieved. This will generate the likelihood distribution of the permeability for both the formation and the fault. Figure 7 shows the resulting likelihood distribution by dividing the fault into 2, 4, 8, and 16 segments and are plotted together with the likelihood distribution using the entire fault.

It is clearly shown in the plot that there is a significant improvement in the distribution of the fault permeability when inversion is done at each fault segment. The figure gives higher confidence that the fault segmentation obtained a well-calibrated model given that the true value lies within the range of most likely value and that the distribution is narrower compared to the likelihood distribution using the entire fault. It also shows that as the number of segments used in the inversion is increased, the precision in estimating the permeability also increased. The inversion process this time captures the true permeability value as it calibrates the possible MEQ occurrences at each portion of the fault. It is also noticeable during the simulation that the permeability value converges smoothly, and convergence is attained with less iteration when more segments are used in the inversion process.

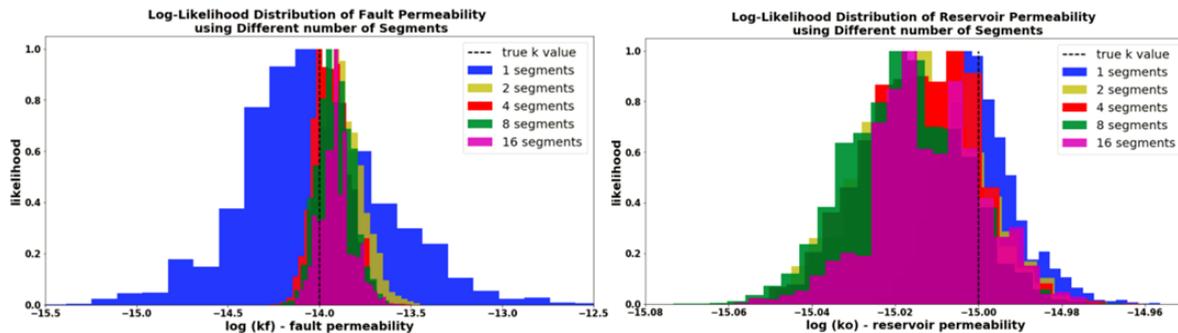


Figure 7. Likelihood distribution of fault permeability at different number of fault segments used in the inversion

### 3.2 Effect of Uncertainty in Reservoir Parameters

This study uses the MEQ data generated from the synthetic model, whose reservoir parameters are known, to check the accuracy and capability of the presented reservoir development workflow in estimating the permeability of the formation and the fault. However, in a real inverse modelling scenario, the model is not yet present, and the parameters are initially unknown. The reservoir model is constructed based on modeller’s understanding of the reservoir, and the reservoir parameters are acquired either from field measurement, from a calculation using the measured parameters, or from the results of experiments or previous studies. These parameters play an essential role in the simulation and contribute to the uncertainty in predicting permeability.

This part of the sensitivity analysis focuses on the effect of uncertainty in porosity on the estimation of permeability. Porosity is one of the reservoir parameters in the model that is difficult to quantify due to the heterogeneity of a geothermal system. Porosity is affected by any changes in the reservoir and can be modified by deformation, hydrothermal alteration, and metamorphism. (Anovitz & Cole, 2015). Though there are several techniques in quantifying total and effective porosity of rocks (e.g., Manger, 1963) and there are methods of porosity determination through density, sonic, and neutron logs (Tiab & Donaldson, 2013), porosity cannot be established precisely in any particular region.

Two additional simulations using the two-parameter model were performed to demonstrate the effect of porosity by modifying its value to 0.15 and 0.25 while keeping other reservoir parameters constant. The resulting likelihood distributions of both simulations are then compared to the likelihood distribution of the synthetic model with porosity value equal the value of 0.2. Figure 8 presents how different porosity values affect the likelihood distribution of fault permeability. It can be observed from the plot that the distribution of permeability values for the additional simulations are wide and scattered compared to the distribution of the model with correct porosity value. Though the most likely value lies within the true value, the plot implies that uncertainty in estimation may substantially increase if the assigned porosity differs from the correct porosity value.

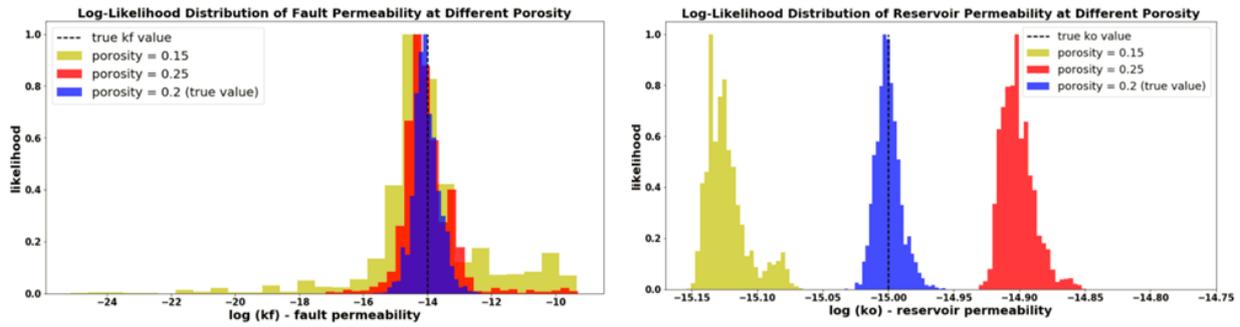


Figure 8. Likelihood distribution of fault and reservoir permeability at different porosity values

The resulting probability distribution of reservoir permeability also follows similar behaviour. The distribution shows a more extensive range if the different porosity value is used in the simulation. One notable observation for the distribution in reservoir permeability value is that the true value of permeability does not lie within the range of most likely values for the two additional simulations. The distribution is shifted away from the correct value. The deviation of the distribution from the actual value affects the accuracy of estimated permeability. The distribution of reservoir permeability in the model is considered tight. Thus minimal differences in the permeability values are observed among the simulations.

### 3.3 Effect of Informative Priors

The prior distribution plays a vital role in Bayesian parameter estimation. It strongly influences the values of the target posterior distribution and can affect the quality of the approximation. The most common type of prior is called the “non-informative” prior. It is used when there is no solid or there is insufficient information about the parameter being estimated. The initial values are taken from a range of values which assume a uniform distribution. This means that the values of the prior used in the inversion are equally likely to occur. This approach of selecting priors is effective, especially if there is no preference in any prior values. The inversion process will control the parameter estimation to come up with the best target distribution. However, for inversion using MCMC, it raises issues with convergence and burn-ins, especially if those proposed values are very far from the target distribution. However, in most cases, a modeller has an idea of how the target distribution looks like or has information on the prior values needed in the inversion. This is called “expert” or “informative” prior. An informative prior is usually used if there is available information about the parameter, say from experiments or previous studies.

The synthetic inversion done with the two-parameter model uses a “non-informative” prior, in which the set of prior values in each chain of the MCMC are given an equal probability. In this section, informative priors are used in the synthetic model and are compared to the non-informative prior used in the previous simulation to demonstrate the effect of assigning biases to the prior values in the inversion. Normal distribution prior with a mean value set at -15 (in log space) is assigned for the reservoir permeability, which is equal to the actual value in the synthetic model. Similarly, normal distribution prior is also used for the fault permeability. However, the mean value is set to -13.5 (in log space), different from the actual values of -14 in the model to demonstrate the effect of modeller’s bias. Both prior values use two standard deviation values of 0.5 and 0.1 to show the effect of restricting the spread of most likely values in the prior distribution. The simulation is then set to run to a maximum of 1,000 iterations or until convergence is achieved. The resulting likelihood distribution of the fault and reservoir permeability estimation is shown in Figure 9.

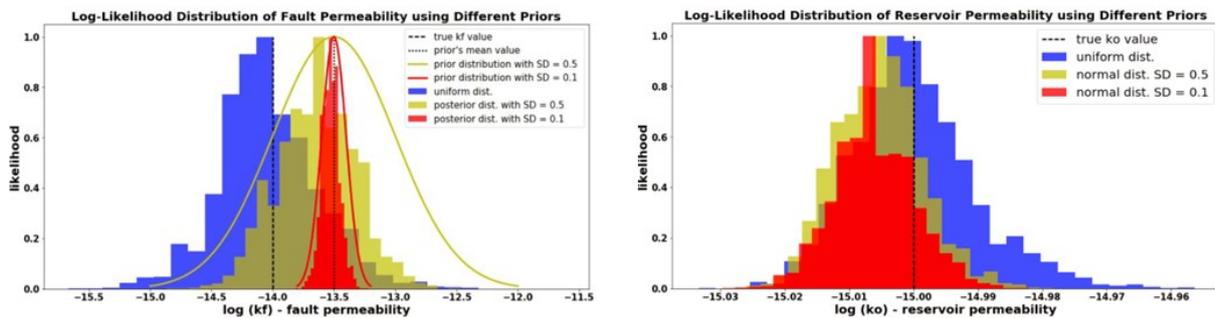


Figure 9. Likelihood distribution of fault and reservoir permeability at a different prior distribution

The plot above presents the influence of informative prior to parameter estimation, which in this study, the permeability of the reservoir and the fault. It can be observed that a decrease in the standard deviation of a normally distributed prior leads to a decrease in the width of the posterior distribution. This is evident in the likelihood distribution of both reservoir and fault permeability. An informative prior influences the resulting posterior distribution by giving more weight to permeability values within the prior’s normal distribution curve. Also, these priors assign lower weights to the permeability values outside the distribution curve. This way, the posterior distribution efficiently converges to the values assigned in the prior.

Additionally, the plot provides is that the maximum likelihood estimate of the fault permeability deviates from the actual value. This is because the posterior distribution is drawn closer to the mean value assigned to the prior distribution used in the synthetic inversion. This is one of the common issues when using biases in the prior distribution. There is a risk of being stuck in the local maxima – a range of values with higher probability – given that the target distribution is drawn near these values while neglecting the neighbouring values with lower probability. Generally, the standard deviation of the informative prior dictates the width while the mean value of the normally distributed prior affects the maximum likelihood estimate of the posterior distribution.

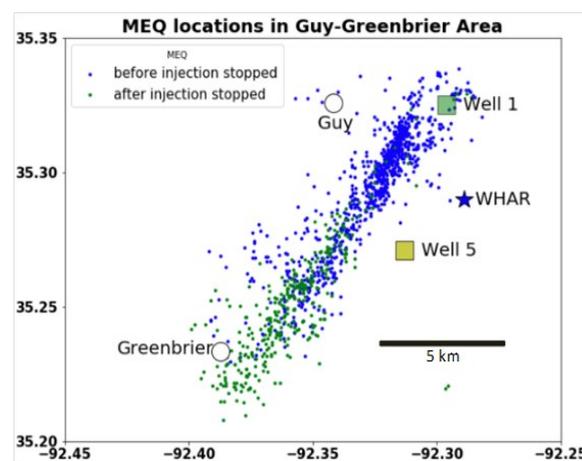
## 4. FLUID INJECTION-INDUCED SEISMICITY CASE STUDY

### 4.1 Background

One of the recent well-documented cases of wastewater injection-induced seismicity is the swarm of earthquakes triggered between the towns of Guy and Greenbrier in Central Arkansas (Horton, 2012). Eight injection wells within the Fayetteville Shale gas play were drilled and were used as disposal of the wastewater from hydraulic fracturing activities in north-central Arkansas. Increased seismic events were experienced in the region after the operation of the first injection well in April 2009. Since then, seismic events continue to occur, clustering at the Guy-Greenbrier area, until the year of 2011. The background seismicity rate of Arkansas as reported in the Advanced National Seismic System (ANSS) catalog from 1967-2008 is 1.4 events/yr for earthquakes with magnitude  $>3.0$ . However, after the start of injection in the area, the rate increased to 16 events/yr from 2009-2011, and this was mostly due to events occurring in the Guy-Greenbrier area (Ogwari, Horton, & Ausbrooks, 2016).

The geohydrologic system in which most of the seismic events occur is called Ozark Plateau Aquifer System, which is located in the eastern part of Arkoma Basin. This system is bounded by the upper system called the Western Interior plains confining system at the top and the Precambrian confining unit at the bottom; both are composed of low permeability rocks (Horton, 2012). The Ozark Plateau Aquifer system has alternating aquifer and confining unit, with faults and fractures that provide pathways for fluid to flow in the area. The shallowest part of the Aquifer is called Springfield Aquifer which is composed mostly of limestone with low intrinsic porosity yet remains permeable due to the dissolution of limestone along with the fractures and bedding. The bottom part of the Aquifer is called Ozark aquifer -- the thickest part of Aquifer -- which is composed mostly of dolostone with some sandstone and limestone, and is characterized as having relatively low intrinsic porosity of 4-6% (Ogwari, Horton, & Ausbrooks, 2016). Chattanooga Shale is a confining unit which is about 14m thick, serves as a confining unit that separates the Ozark Aquifer and the overlying Springfield Aquifer. The confining unit, which is 25% shale and 75% sandstone, is slightly permeable but the hydraulic connections between the two aquifers vary significantly with the local lithologic and structure differences (Horton, 2012; Ogwari, Horton, & Ausbrooks, 2016).

Wells 1 and 5 were assumed to be responsible for the swarm of earthquakes occurred in the area. Well 1 started its operation on 7 July 2010, with the fluid injecting both at Springfield aquifer and Ozark aquifer. Well 5 which is situated near Well 1 started its operation on 18 August 2010, injecting at the deeper part of the Ozark aquifer. After the operation of the two wells, hundreds of small to moderate earthquakes were detected near the wells, and the seismicity migrated towards the southwest, which happens until the end of the year 2010. The events were located at depths between 3 and 7 km, and forming a linear structure of 5 to 6 km length striking at N30E. This structure, which is called Guy-Greenbrier fault, is consistent with the NE-trending of the focal mechanism of the magnitude 4.0 earthquake that occurred on 11 October 2010 in the same area (Horton, 2012). The seismic activity continued to occur several kilometres to the south, which makes the fault more evident as it became 13 km in length. On 27 February 2011, a magnitude 4.7 earthquake was triggered in the area, which caused an emergency shutdown order from the Arkansas Oil and Gas Commission (AOGC) that halted the operation of the two wells on 4 March 2011. Seismic events continued to occur for the next seven months after the termination of injection, but the rate of earthquake occurrence gradually decreased (Horton, 2012). Figure 10 shows the map of microseismicity occurred within the Guy-Greenbrier fault, in reference to the ANSS catalog, together with the location of the wells (Huang and Beroza, 2015). This case study aims to estimate the permeability of the Guy-Greenbrier fault and the Ozark Aquifer using the seismic events that occurred from July 2010 to October 2011. This covers the events triggered after injection started in Well 1, and it also includes events that occurred several months after injection was halted.



**Figure 10. Map of the location of microseismic events within the Guy-Greenbrier area based on ANSS catalog including the location of Well 1 and Well 5. Source: (Dempsey, Suckale, & Huang, 2016)**

### 4.2 Model Setup

A simple numerical model is constructed to represent the wastewater injection in the Guy-Greenbrier area where the swarm of earthquakes is located. The model covers both the location of the observed microseismic events together with the two wells where the fluid is being injected. The rectangular grid is created using PyTOUGH with a dimension of 20 km x 30 km x 1.03 km, representing the Ozark aquifer. The model assumes a no-flow boundary both at the top and bottom of the grid as the aquifer is covered with impermeable cap rock and sealed by a basement confining unit below. The boundary blocks have a dimension of 1,000 m x 1,000m, which assumes a constant pressure boundary condition. The grid block size decreases to 800 m, 700 m, and 500 m towards the centre of the grid. A local grid refinement with a dimension of 250 m x 250 m is created in the centre of the numerical model grid where the fault and the wells are situated. This is done to enhance the accuracy of the simulated pressure

change within the area. The model is rotated with an angle of 35° from the north to align with the inferred Guy-Greenbrier fault. Figure 11 shows the model grid used in this case study. The figure also shows the location of the wells and the grid blocks where the assumed fault is located.

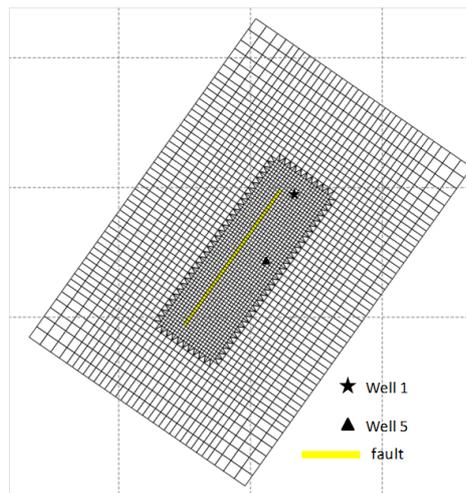


Figure 11. Numerical model grid of wastewater injection area in Arkansas

The model grid is divided into two layers. The top layer has a thickness of 70 m and represents the combined injection interval of Well 1 in Ozark aquifer and Springfield aquifer. The model assumes that the two aquifers are well-connected and all of the injected fluids in Well 1 flow through the first layer. The second layer has a thickness of 960 m and represents the injection interval of Well 5 in the deep portion of the Ozark aquifer. The fluid injected in Well 5 is assumed to flow through this layer. The injection rate used in the model is shown in Figure 12. It covers the monthly injection rate of the two wells from July 2010 to October 2011, the time when the seismic events are triggered along the fault (Huang and Beroza, 2015). The temperature of the injecting fluid is set at 30°C, which is the same as the initial temperature assigned in the grid blocks of the model to eliminate thermal effects. Two rock types are created in this model, which corresponds to the Ozark Aquifer and the Guy-Greenbrier fault. The swarm of microseismic events is represented by a line segment, 13 km in length, and is situated at the bottom layer of the model, in the assumption that the formations in which most of the events are triggered are very well connected. The rock type assigned to all blocks in the model represents the Ozark aquifer except for those blocks traversed by the assumed fault, in which the Guy-Greenbrier fault rock type is assigned. The rock properties use the default values set in AUTOUGH2, except for the porosity in which the value assigned is 5% for both rock types.

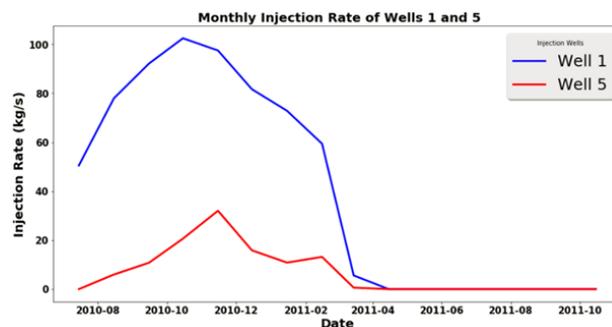


Figure 12. Monthly Injection rates (in kg/s) of Well 1 and Well 5

### 4.3 Inverse Modelling and Result

This case study aims to estimate the permeability of the Ozark aquifer and Guy-Greenbrier fault of the wastewater injection area in Arkansas using the induced seismicity model and the calibration method presented in this paper. The simulation is conducted using TOUGH2 with EOS1 as the equation of state and runs for 1.25 years (the period similar to the injection data). The model uses a critical pressure of 0.10 MPa, which is equal to the initial block pressure set in the model grid to allow triggering of seismic events at low overpressure. The initial pressure assigned on the grid blocks does not affect the calibration since the earthquake triggering presented in this study relies only on pressure changes in the grid blocks. The average seismicity rate along the fault is then generated by integrating the pressure output at every month along the fault. The seismicity rate is then compared to the actual MEQ data along the Guy-Greenbrier fault and is scored using the log-likelihood for a nonhomogeneous Poisson process. The MEQ data used in the calibration is presented in Figure 13 (Huang and Beroza, 2015), which is obtained from the ANSS catalog and only covers earthquakes with magnitude  $\geq 2.0$ .

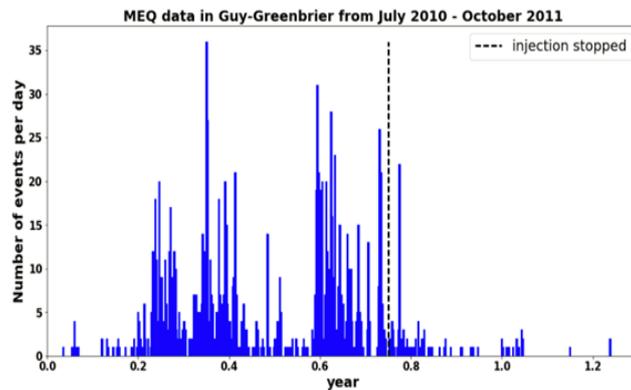


Figure 13. MEQ catalog of Guy-Greenbrier fault from July 2010 to October 2011

Inverse modelling is performed to estimate the permeability of the aquifer and the fault by running several forward models of different permeability values and plotting all likelihood values to generate a likelihood distribution. The inversion is done using MCMC by running Python “emcee” module with two-dimensional vectors which correspond to aquifer and fault permeability. The initial values for the aquifer permeability range from  $10^{-12}$  to  $10^{-16}$  m<sup>2</sup> while the initial values for fault permeability range from  $10^{-12}$  to  $10^{-14}$  m<sup>2</sup>. The permeability of the blocks assumed to be isotropic, and the priors are uniformly distributed. The inversion is set to run for 1,000 iterations or until it reaches convergence. Figure 14 shows the resulting likelihood distribution of the aquifer and fault permeability together with the contour of possible permeability combinations.

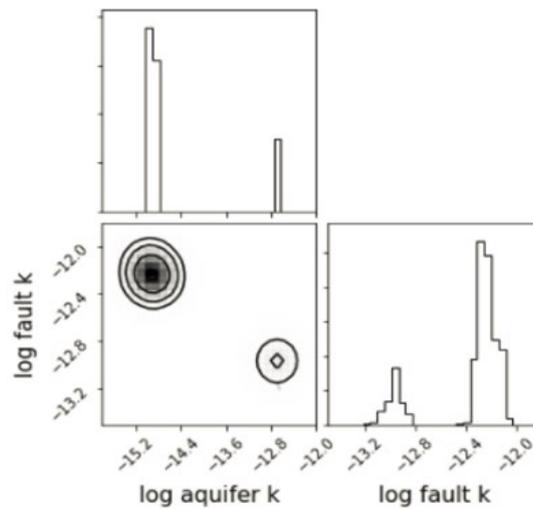


Figure 14. Corner plot for the likelihood distribution of permeability of Ozark aquifer and Guy-Greenbrier fault

The plot shows a bimodal distribution – a distribution having two maxima or modes – for both the aquifer and fault permeability. Most of the likelihood values concentrate on the upper left portion of the corner plot with aquifer permeability value between  $1.0 \times 10^{-15}$  and  $2.1 \times 10^{-15}$  m<sup>2</sup> and fault permeability value between  $3.2 \times 10^{-13}$  and  $1.0 \times 10^{-12}$  m<sup>2</sup>. Some of the likelihood value merge on the lower right corner of the plot has a range of aquifer permeability between  $2.0 \times 10^{-13}$  and  $2.5 \times 10^{-13}$  m<sup>2</sup> and fault permeability between  $6.3 \times 10^{-14}$  and  $1.6 \times 10^{-13}$  m<sup>2</sup>. Both combinations of aquifer and fault permeability values match the actual MEQ data (Figure 15). However, the smaller contour is highly unlikely in a geothermal system since the aquifer permeability is higher than the fault permeability. Though the resulting permeability values match the actual MEQ data, this result will not be the focus of this paper.

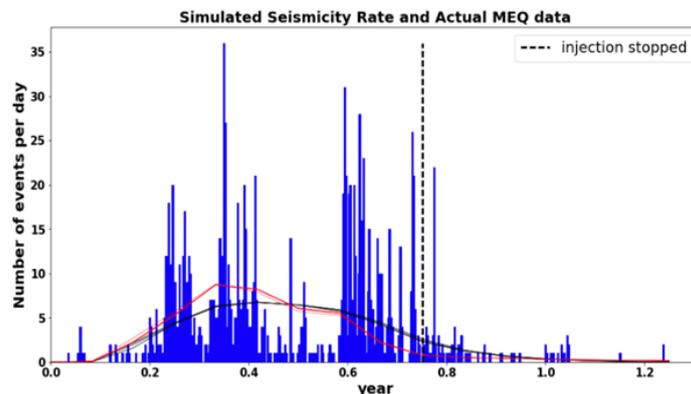
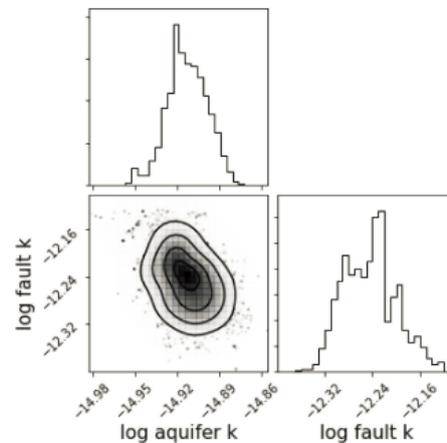


Figure 15. Simulated seismicity rate and observed seismicity rate of Guy-Greenbrier area. The black and red lines correspond to the larger and the smaller contour, respectively.

The larger contour of the resulting corner plot is considered as the more relevant result for the permeability estimate since the fault permeability is higher than that of the aquifer. Figure 16 shows a more detailed corner plot of the permeability estimate by discarding the probability values associated with the smaller contour. Similar to the result of the synthetic model, the most likely values of the aquifer permeability remains very tight. This might be because the permeability of the blocks in the model is mostly aquifer permeability. It is noted that the onset time of MEQ triggering is sensitive to aquifer permeability (Dempsey, Suckale, & Huang, 2016). On the other hand, the very tight probability distribution seems to occur regularly in all of the simulations of the study which might be caused by some uncertainty in the simulation or might be due to the limitation of the methods developed in this study. The fault permeability, similar to the synthetic model, has been effectively estimated by the inversion. The pressure change as the fluid flows along the fault blocks captures the permeability of the fault as the calibration matches the generated seismicity rate with the actual MEQ data. The narrow range of most likely values gives high confidence of a well-calibrated model.



**Figure 16. Corner plot for the likelihood distribution of aquifer and fault permeability discarding data of smaller contour**

## 5. CONCLUSION AND RECOMMENDATION

The study introduced a theoretical approach in integrating MEQ data in reservoir model development to estimate reservoir parameters, specifically the permeability of the reservoir and the fault. The method presented in this study started by creating a numerical model and running it using TOUGH2 to estimate the pressure changes along the fault and in the reservoir. The calibration for the induced seismicity model was done by adjusting reservoir parameters to match approximately the simulated seismicity rate to the generated MEQ from the synthetic model. Inverse modelling was performed to estimate the permeability of the reservoir and the fault using MCMC as its sampling method. The probability distribution of the most likely values was then generated, given that there is uncertainty in MEQ observations due to the randomness of the occurrence of earthquakes.

The result of the two-parameter model suggests that it is possible to identify reservoir parameters provided that there is a sufficient MEQ dataset used for calibration. The result also pointed out the effect of the other uncertainty in other reservoir parameters (e.g., porosity), accurate estimates of which the inversion relies on. The study shows that prior knowledge can influence the output of the probability distribution. Also, dividing the fault into smaller segments and performing calibration at each segment reduces the uncertainty in the probability distribution of the reservoir and fault permeability that gives higher confidence of a well-calibrated model.

The case study on the wastewater injection in Arkansas shows that the theoretical method presented in this paper can be applied to an actual MEQ dataset. The result shows a bimodal distribution giving two combinations of the aquifer and fault permeability that match the MEQ dataset. The value in which the distribution is concentrated shows higher permeability of the fault than that of the aquifer. This is consistent with the knowledge that the fault serves as the principal conduit of fluid flow in the reservoir; thus, it is expected to be more permeable than the aquifer. The inversion method captures the fault permeability from the pressure change due to fluid flow within the surrounding aquifer blocks.

Further improvements can be made in the model development and calibration, which include the following:

1. Earthquake declustering to improve model calibration by removing the recorded aftershocks.
2. Integration of thermal stress in earthquake triggering.
3. Refinement of a numerical model for the case study and validating the estimated permeability through other studies that also estimates the permeability of the fault using MEQ.

## REFERENCES

- Anovitz, L., & Cole, D. (2015). Characterization and Analysis of Porosity and Pore Structures. *Reviews in Mineralogy & Geochemistry*, 80, 61-164. doi:<https://doi.org/10.2138/rmg.2015.80.04>
- Croucher, A. (2016). *PyTOUGH User's Guide*. Auckland, New Zealand: Department of Engineering Science, The University of Auckland.
- Croucher, A., & O' Sullivan, M. (2013). Approaches to Local Grid Refinement in TOUGH2 Models. 35th New Zealand Geothermal Workshop: 2013 Proceedings. Rotorua, New Zealand: New Zealand Geothermal Workshop.
- Debski, W. (2008). Estimating the Earthquake Source Time Function by Markov Chain Monte Carlo Sampling. *Pure and Applied Physics*, 165(7), 1263-1287. doi:<https://doi.org/10.1007/s00024-008-0357-1>

- Dempsey, D., & Suckale, J. (2017). Physics-based forecasting of induced seismicity at Groningen gas fields, The Netherlands. *Geophysical Research Letters*, 44. doi:10.1002/2017GL073878
- Dempsey, D., Suckale, J., & Huang, Y. (2016). Collective properties of injection-induced earthquake sequences: 2. Spatiotemporal evolution and magnitude frequency distribution. *Journal of Geophysical Research: Solid Earth*, 121. doi:10.1002/2015JB012551.
- Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. (2013). emcee: The MCMC Hammer. *Publications of the Astronomical Society of the Pacific*, 125, 306-312.
- Horton, S. (2012). Disposal of hydrofracking waste fluid by injection into subsurface aquifers triggers earthquake swarm in Central Arkansas. *Seismological Research Letters*, 83(2), 250-260. doi:10.1785/gssrl.83.2.250
- Huang, Y., & Beroza, G. (2015). Temporal variation in the magnitude-frequency distribution during the Guy-Greenbrier earthquake sequence. *Geophysical Research Letters*, 42, 6639-6646. doi:10.1002/2015GL065170
- Lindqvist, B. H., & Taraldsen, G. (2013). Exact Statistical Inference for Some Parametric Nonhomogeneous Poisson Process. *Journal of Iranian Statistical Society*, 12(1), 113-126. Retrieved from <http://jirss.iostat.ir/article-1-211-en.html>
- Manger, E. G. (1963). Porosity and Bulk Density of Sedimentary Rocks. Washington D.C.: Geological Survey Bulletin 1114-E.
- Ogwari, P., Horton, S., & Ausbrooks, S. (2016). Characteristics of Induced/Triggered Earthquakes during the Startup Phase of the Guy-Greenbrier Earthquake sequence in North-Central Arkansas. *Seismological Research Letters*, 87(3). doi:10.1785/0220150252
- O'Sullivan, J., Dempsey, D., Croucher, A., Yeh, A., & O' Sullivan, M. (2013). Controlling Complex Geothermal Simulations Using PyTOUGH. Proceedings, 38th Workshop on Geothermal Reservoir Engineering. Stanford, California: Stanford University.
- O'Sullivan, M., & Pruess, K. (2000). Geothermal Reservoir Simulation: The State-of-Practice and Emerging Trends. World Geothermal Congress.
- Pramono, B., & Colombo, D. (2005). Microearthquake Characteristics in Darajat Geothermal Field, Indonesia. Proceedings World Geothermal Congress. Antalya, Turkey.
- Pruess, K., Oldenburg, C., & Moridis, G. (1999). TOUGH2 User's Guide, VERSION 2.0. Berkeley: Lawrence Berkeley National Laboratory.
- Rivera, J. M., & Dempsey, D. (2017). Forward and Inverse Modelling of Geothermal Microseismicity using TOUGH2 Coupled with an Earthquake Simulator. 39th New Zealand Geothermal Workshop: 2017 Proceedings. Rotorua, New Zealand: New Zealand Geothermal Workshop.
- Rivera, J. M., & Dempsey, D. (2018). Forward and Inverse Modelling of Induced Seismicity using TOUGH2 Coupled with an Earthquake Simulator (Master's Thesis). Retrieved from <http://hdl.handle.net/2292/37430>
- Sepulveda, F., Andrews, J., Alvarez, M., Montague, T., & Mannington, W. (2013). Overview of Deep Structure Using Microseismicity at Wairakei. Proceedings: 35th New Zealand Geothermal Workshop. Rotorua, New Zealand.
- Sewell, S., Cumming, W., Bardsley, C., Winick, J., Quinao, J., Wallis, I., . . . Bannister, S. (2013). Interpretation of Microearthquakes at the Rotokawa Geothermal Field, 2008 to 2012. Proceedings: 35th New Zealand Geothermal Workshop. Rotorua, New Zealand.
- Sherburn, S., Bromley, C., Bannister, S., Sewell, S., & Bourguignon, S. (2015). New Zealand Geothermal Induced Seismicity: an overview. Proceedings World Geothermal Congress. Melbourne, Australia.
- Stuart, G., Yang, W., Minkoff, S., & Pereira, F. (2016). A two-stage Markov chain Monte Carlo method for velocity estimation and uncertainty quantification. Conference: Society for Exploration Geophysicists Annual Meeting 2016, (pp. 3682-3687). doi:<https://doi.org/10.1190/segam2016-13865449.1>
- Tiab, D., & Donaldson, E. (2015). *Petrophysics: Theory and Practice of Measuring Reservoir Rock and Fluid Transport Properties* (4th ed.). Elsevier Inc.
- Wallis, I., Moon, H., Clearwater, J., Azwar, L., & Barnes, M. (2015). Perspective on Geothermal Permeability. Proceedings: 37th New Zealand Geothermal Workshop. Taupo, New Zealand.
- van Ravenzwaaij, D., Cassey, P., & Brown, S. (2016). A simple introduction to Markov Chain Monte-Carlo sampling. *Psychonomic and Bulletin Review*, 1-12. doi:<https://doi.org/10.3758/s13423-016-1015-8>