MACHINE LEARNING INVESTIGATION OF INJECTION-SEISMICITY IN ROTOKAWA GEOTHERMAL FIELD

Pengliang Yu[1], David Dempsey[2], Aimee Calibugan[3], Rosalind Archer[1]

[1]Department of Engineering Science, University of Auckland, Auckland, New Zealand

[2] Department of Civil and Natural Resources Engineering, University of Canterbury, Christchurch, New Zealand

[3] Mercury NZ, Rotorua, New Zealand

pengliang.yu@auckland.ac.nz

Keywords: *Microseismicity, injection, time series feature engineering, seismicity rate, significant features, p-value*

ABSTRACT

Understanding the injection-seismicity relationship in geothermal reservoirs can provide insight into reservoir connectedness. One challenge is that, in real fields, fault and reservoir complexity make it difficult to apply simple analytical models to understand the data. Here, we use a machine learning technique called time-series feature engineering to study relationships between aspects of fluid injection and microearthquakes in Rotokawa geothermal field, New Zealand. We took four years of injection data between 2012 and 2016 and sliced it into smaller sub-windows. For each window, the average seismicity in a look-back period was computed, and then binary label of 1 was assigned if it exceeded a threshold. Automatic time series feature extraction from the raw and transformed injection data in each window was performed using Python package *tsfresh*. Significant features of the data were identified on the basis of distribution discrepancy between the two labels. The results show that the injection rate at some wells is a predictor of long-term (fortnightly) earthquake rates. At other wells, there is a poor correlation between injection rate and seismicity. We have been unable to find any link between rapid changes in injection rate and seismicity spikes, as suggested by some theoretical models.

1. INTRODUCTION

Microseismicity is a common phenomenon resulting from hydrothermal fluid circulation in geothermal operations (Giardini, 2009; Hopp et al., 2020; Majer et al., 2012). These small earthquakes are generally only detectable by sensitive networks, although occasionally felt events can occur. The locations, rates and magnitude frequency statistics of the microearthquakes carry information about pressure changes (Dempsey & Suckale, 2016) in a geothermal system and, by extension, its relative connectedness. However, this information is often underutilized, as it is difficult to interpret.

Numerous studies have investigated the mechanisms of induced seismicity and explored links with fluid injection. It is widely understood that injected fluids increase pore pressure and decrease the effective stress on critically stressed faults (Healy et al., 1968; King Hubbert & Rubey, 1959) and this can promote fault slip that is detectable as earthquakes. If well-located, the seismicity can reveal faults that may operate as fluid conduits or baffles in the system.

Segall and Lu (2015) used analytical and numerical models to explain the reasons for possible post-injection seismicity, and highlighted that a poroelastic surge may be the main reason for seismic events shortly after well shut-in. This was supported by the findings of Deng et al. (2020) who modelled wastewater disposal at multiple wells near the town of Cushing, Oklahoma and showed that poroelastic stress changes can affect the regimes on a preexisting fault where shear slip is promoted or inhibited. In contrast, Turuntaev (2018) used a rate-state model to estimate earthquake rates at the Basel geothermal project in Switzerland and showed only a small increase after shut-in. Dempsey and Riffault (2019) developed analytical and numerical models of seismicity rate changes when injection rates are reduced and found a decline, occasional quiescence, and eventual recovery of the seismicity rate to a new equilibrium after a rate reduction.

These studies suggest the relation between injection and seismicity is complex and that the underlying physics are hard to capture fully in current forecasting models. Notably, the prediction accuracy of these analytical models is strongly dependent on the input parameters (Eaton & Igonin, 2018), and numerical simulations might have problems with calibration when proper parameters are difficult to select.

Despite recent progress in understanding the mechanisms of induced seismicity, there remain several important questions that need to be resolved. Although the operational parameters of fluid injection rate, wellhead pressure, total injection volume, well injection location, geological, and geomechanical characteristics are considered important factors affecting induced seismicity rates (He et al., 2020; Hofmann et al., 2018), precise relationships are usually obscured by large and disorderly data. This is not necessarily a shortcoming of monitoring systems, but rather inherent randomness in the locations of earthquake generating fractures and their relative stress criticality that predisposes them to failure. All this makes it hard to extract useful information by conventional analytical or numerical methods (He et al., 2020).

Machine learning methods have proven to be a helpful tool for geoscientists in recent years due to their ability to identify obscured patterns in data (Holtzman et al., 2018; Wozniakowska & Eaton, 2020). This includes identification and forecasting of earthquakes (Corbi et al., 2019; DeVries et al., 2018; Limbeck et al., 2021; Lubbers et al., 2018; Wang et al., 2020; Zhang et al., 2020), classification of remote sensing images (Bialas et al., 2016; Frank et al., 2017), identification of changes in faulting processes

ISSN 2703-4275

(Holtzman et al., 2018), and factors controlling induced seismicity (Mehrabifard & Eberhardt, 2021; Wozniakowska & Eaton, 2020). However, studies that correlate microseismicity with injection well data using ML-based methods are rarely reported.

Rotokawa geothermal field is located in the central Taupo Volcanic Zone (TVZ) of New Zealand, a region of active rifting with associated volcanism, seismicity and widespread hydrothermal circulation. Natural seismicity in the TVZ usually occurs above depths of 6-8 km, which defines the depth of the brittle crust (Bryan et al., 2010). Much of the seismicity occurs in the Taupo Fault Belt, a central region of high-density faulting, however swarm activity is noted elsewhere, often associated with calderas (Bannister et al., 2016; Illsley-Kemp et al., 2021; Webb et al., 2012).Within geothermal fields of the TVZ, natural and anthropogenic seismicity associated with reservoir operations has been well documented (Sherburn et al., 2015; Clarke et al., 2009; Hopp et al., 2019; Sepulveda et al., 2016; Sherburn et al., 2013).

The Rotokawa geothermal field has sustained geothermal energy production for more than 20 years (McNamara et al., 2016). Previous studies of microseismicity (Hopp et al., 2020; Sherburn et al., 2013) describe localization in a region south of the Waikato River and near to some injection wells. In 2010, a factor of four increase in reinjection and corresponding 3.7x increase in microseismicity, suggested a close relationship (Sherburn et al., 2013). More recently, Hopp et al.(2020) studied the rate and location of seismicity between 2012 and 2016 in the context of injection strategy and found the rate to be insensitive to major changes in well injectivity and change of the dominant injector.

Here, we use a machine learning technique to further study relationships between fluid injection and microearthquakes in the Rotokawa geothermal field. We used injection rate data from three wells at Rotokawa between 2012 and 2016, and the catalog of microearthquakes from Hopp et al.(2020). We applied systematic time series feature engineering to investigate the relationships between long-term and short-term injection rates, and long-term and short-term seismicity rates. We also investigated the sensitivity of seismicity rate to the individual injection wells, and the relative importance of different kinds of injection well data (injection rate, wellhead pressure, hydraulic energy). Finally, based on the hypothesis of Segall & Lu, (2015), we tested the link between rapid change in injection rate and seismicity rate spikes, and found no evidence for this phenomenon.

2. METHODOLOGY

The time series feature engineering techniques used in this paper were initially developed for anomaly detection in industrial steel casting (Christ et al., 2016). Later, Rouet-Leduc et al. (2018) and Dempsey et al., (2020) introduced the technique to study links between seismic tremor signals and slip on subduction zones, and automatic recognition of volcanic eruption at Whakaari, New Zealand, respectively. Here, we are exploring whether they may be useful in a geothermal setting. The Python package *tsfresh* was deployed to implement the automatic time series feature extraction and selection on the basis of the FRESH algorithm (Christ et al., 2016). The result is over 700 extracted features from raw data (e.g., injection rate), which can then be evaluated for statistical significance against a label quantity of interest (e.g., seismicity rate).

2.1 Data Compilation and Preparation

The raw earthquake data for Rotokawa geothermal field were obtained from Hopp et al., (2020) shared at DOI https://doi.org/10.17605/OSF.IO/C2M6U. The catalog data were filtered to retain events shallower than 6 km, near the injection area of Rotokawa, and above an estimated magnitude of completeness of 0.5. The hourly injection data (rate and wellhead pressure) for selected wells (well #1, #2, #3), from 2012 to 2016, were first processed using data interpolation to fill any gaps. The total injection rate was calculated as a sum of these three wells.

2.2 Time series feature engineering method

2.2.1 Seismicity label vector calculation

The label vector is a binary classification metric used to denote whether the future earthquake rate, λ, is high (1) or low (0). It is obtained by the following steps: (a) build a continuous estimate of the earthquake rate averaged over some time period Δt_i, which provides for different levels of smoothing, $\lambda(t, \Delta t_i) = Eqs_{\Delta t_i}/\Delta t_i$. The term $Eqs_{\Delta t_i}$ is the number of events in the time period $[t - \Delta t_i, t]$, with larger Δt_i representing a longer-term earthquake rate, and smaller Δt_i representing a shorter-term earthquake rate. By varying Δt_i we can investigate injection features that affect both short and long-term seismicity rates; (b) After computing the average earthquake rate λ, an initial binary classification is computed by comparison against a threshold earthquake rate λ_{th}. If $\lambda > \lambda_{th}$, a value $Y_{01} = 1$ is assigned, otherwise, $Y_{01} = 0$. The vector Y_{01} classifies the instantaneous seismicity as high or low relative to an arbitrary threshold. (c) In this study, we are interested in whether seismicity is high or low related to prior aspects of the injection. Therefore, we defined a look back time gap Δt_j that defines the periods *preceding* elevated seismicity, as defined by Y_{01}. We calculate the label vector $Y_{\Delta t_j}$ for each time gap of Δt_j, which is expressed as $Y_{\Delta t_j} = [Y_{01}(t_i - \Delta t_j) == 1]$. Thus, the label vector $Y_{\Delta t_j}$ classifies whether future seismicity (looking forward a distance Δt_j) will be elevated. This vector helps us identify the time periods preceding earthquake rates of interest, and focuses the search for injection relationships in these periods. Figs.1 and 2 show the seismicity rate λ and label vector $Y_{\Delta t_j}$ when Δt_i=1 week, Δt_j=2 weeks, and λ_{th}=70 events per week, which is an arbitrary threshold for illustration purposes.

ISSN 2703-4275

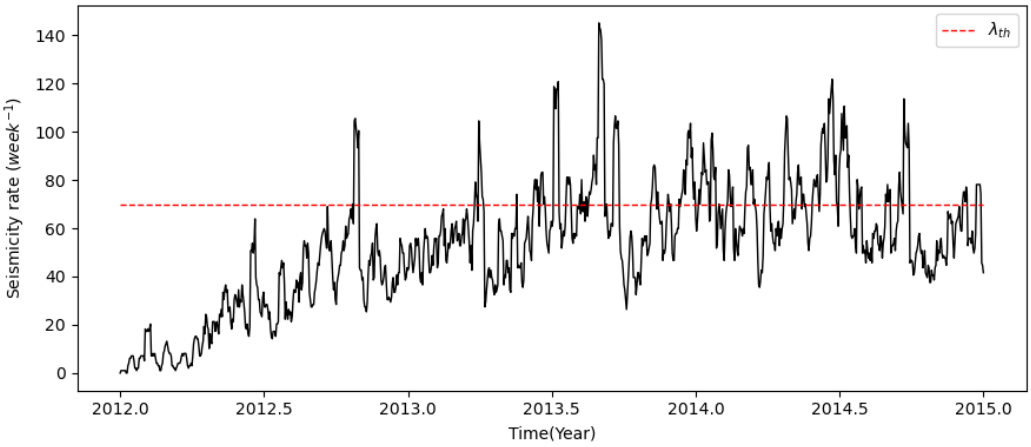


Figure 1: Seismicity rate change over time

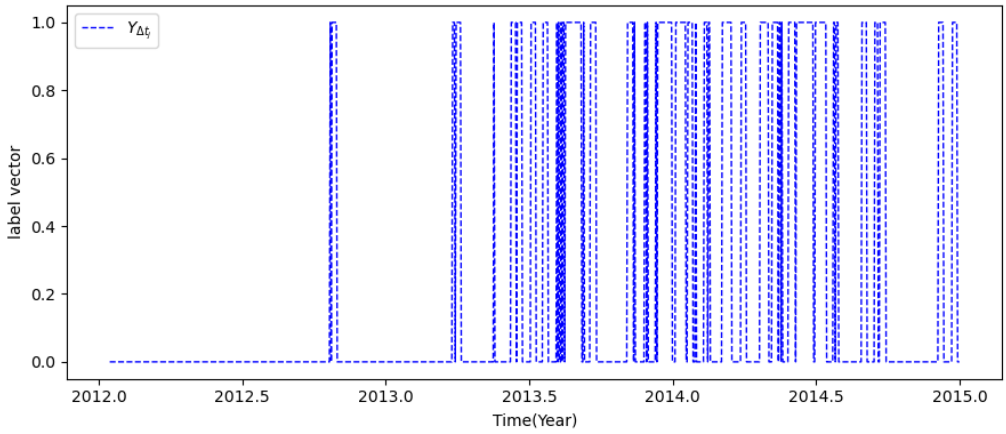


Figure 2: Label vector when seismicity rate threshold $\lambda_{th} = 70$ events per week.

2.2.2 Feature matrix extraction and statistical tests

For each of Well #1, Well #2 and Well #3, automatic time series feature extraction was performed for injection rate (Q), well head pressure (WHP), hydraulic energy (defined as $Q \times WHP$), and corresponding time derivatives (dQ/dt, $dWHP/dt$ and $d(Q \times WHP)/dt$). Feature extraction was also performed for Q_{total} and dQ_{total}/dt. Each of these raw and transformed data streams was subdivided into overlapping time windows of length $\Delta t_w = 2$ days and with overlapping factor of 0.75. The Python package *tsfresh* was used to calculate 759 different times series features for each time window. Examples of time series features include distributional parameters (mean, median, std. dev. of windowed data), measures of autocorrelation (FFT and lag coefficients), information (energy, entropy), and parameters of linear and non-linear regressors. For example, for a window containing n samples, the feature quantile *(n, q)* is calculated as the value of n greater than q% of the ordered values from n in each time window.

Statistical tests are used to determine which of the many possible features are relevant to the seismicity periods of interest. For each feature, *tsfresh* evaluates its significance by computing a Mann-Whitney U test that assesses whether the distribution of feature values corresponding to label $Y_{\Delta t_j} = 0$ and the distribution corresponding to label $Y_{\Delta t_j} = 1$ have different medians. Each test yields a p-value, and this allows us to sort features from smallest (most likely to be significant) to largest (least likely).

Several steps were taken to control for spurious associations in the p-value tests. First, to ensure a balanced testing set, we performed random undersampling of the windows so that the number of 1 and 0 labels in $Y_{\Delta t_j}$ was equal. As this procedure could generate multiple possible p-values depending on the inherent randomness of the undersampling, we performed it 100 times and took the median p-value. Second, when comparing across different values of λ_{th} or for different wells, we cannot guarantee that the undersampled feature sets, although balanced, were of a similar size. Therefore, the undersampled features were further downsampled so that they are all contained same numbers of values, corresponding to the minimum size among all undersampled feature sets for all values of λ_{th} and different wells.

3. RESULTS AND DISCUSSION

There are several ways to qualitatively illustrate and rank the significance of the relationships between injection rate and seismicity. Most directly, we can study how p-value varies as a function of parameters controlling the injection rate (averaging scale) or the label

vector (averaging scale and threshold). For individual features with particularly low p-values, we can examine their corresponding label-1 and label-0 distributions to understand the degree of overlap. Finally, to obtain a bulk measure of significance for individual data streams, e.g. injection rate into Well #3, $dWHP/dt$ for Well #1, we calculate the sum of the logarithm of all p-values. This measure aggregates all features together, and is only useful for comparing relative associations.

We first assessed, at a coarse-scale, associations between injection and seismicity rates when those quantities are calculated for different time-scales. For example, we checked whether the daily or fortnightly seismicity rate had a stronger link with injection. This is described in Section 3.1. In Section 3.2, we explore the distributions of feature values for total and individual well injection rates. In Section 3.3, we analyze injection rates for the different wells to investigate which of these has a closer association with microseismicity. Finally, in the Section 3.4, we rank the relative importance of different well data (injection rate, wellhead pressure, hydraulic energy) in controlling the level of microseismicity. As the relationship between injection rate reduction and induced seismicity remains largely enigmatic, we also explored the evidence of throttling – rapid changes in injection – as a trigger of microseismicity, by comparing bulk associations with Q and dQ/dt with the short-term seismicity rate.

3.1 Seismicity associations with total and individual well injection rates across different time scales

In this section, we test for an association between long-term (Δt_i=20 days) or short-term (Δt_i=2 days) seismicity rate with the total injection rate, as well as rates into Well #2 and Well #3 wells individually. Injection rate was calculated as the average over different periods of length Δt_q which ranged between 2 and 20 days. The seismicity rate was calculated for an averaging period Δt_i between 2 and 20 days where for each time-scale a threshold, λ_{th}, was used at the 50th percentile. Figure 3 plots $\log_{10} p$-value under different combinations of Δt_i, Δt_q, and λ_{th}, taking imbalance and size effects into account by under-sampling and down-sampling.

Focusing on total injection rate, Fig.3 (a) shows an elevated association (smaller p value) between long-term seismicity and the long-term total injection rate. Separating individual well contributions shows much stronger associations for both Well #2 and Well #3. This association is between long-term seismicity, and the injection rates averaged over all time scales. The reduced association with microseismicity measured over periods of only a few days can be explained by relative dominance of aleatoric variability. Short-term earthquake rates, by construction, have fewer events from which to infer an underlying seismicity rate. This results in a large degree of variability in the day-to-day seismicity rate and this in turn obscures any relationships with the injection. In comparison, by averaging over several weeks, we smooth out short-term stochasticity, and an association with the injection rate becomes apparent.

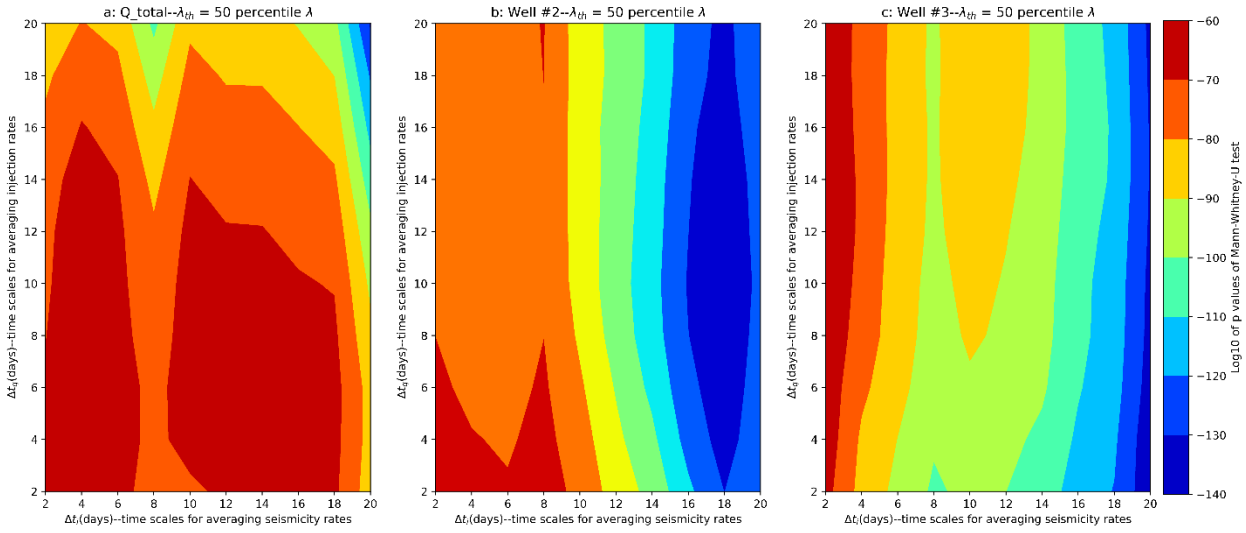


Figure 3: $\log_{10} p$-values of Mann-Whitney-U test for seismicity and injection rates averaged under different Δt_i and Δt_q taking λ_{th} at the 50th percentile.

3.2 Time series feature analysis of total and individual well injection rates

On the basis of Section 3.1, we decided to investigate the long-term seismicity rate (Δt_i = 20 days) and how it depends on different aspects of the short-term injection rate (Δt_w = 2 days). Fig. 4 plots distributions of the median injection rate for all time windows. We plot the distributions corresponding to label vectors 1 and 0 separately to emphasize how different injection styles are followed later by different levels of microseismicity. Note that the distributions in Fig. 4 have been scaled to emphasize the relative position of their centre and spread.

Fig.4 (a) exhibits the weak association between total injection rate and seismicity as a lack of clear separation between distributions for label vectors 1 and 0. In contrast, Fig.4 (b) shows that higher injection rates precede higher rates of microseismicity for injection into the Well #2. Finally, Fig.4 (c) shows a negative association for the Well #3 between higher injection rate and high rates of microseismicity.

ISSN 2703-4275

These trends reflect a shift in field operation when a portion of reinjection was switched from Well #3 to Well #2 due to an injectivity decline in Well #3 (Hopp et al., 2020). At the same time, there was an increase in the seismicity rate, which accounts for the positive and negative associations with Well #2 and Well #3, respectively. This indicates that, relationships between injection rate and seismicity are potentially sensitive to reservoir operations. The relationships are complex and forecasting seismic response to future injection will be neither easy nor completely accurate.

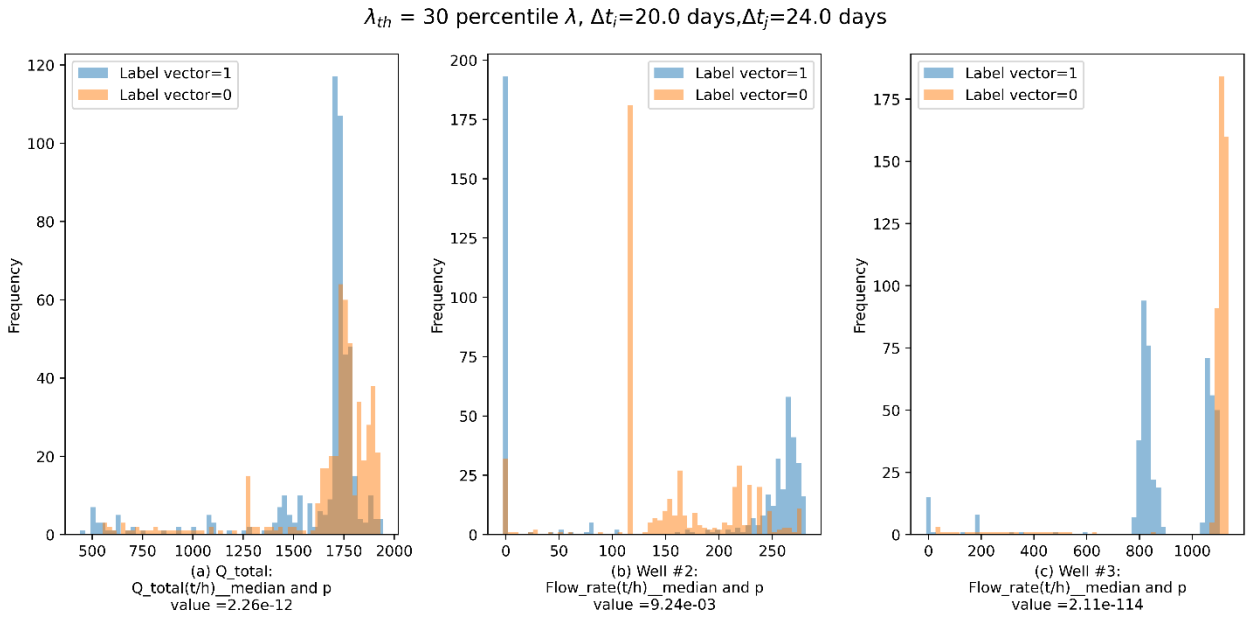


Figure 4: Median injection distributions for (left) total injection rate, as well as (middle) Well #2 and (right) Well #3.

3.3 Time series feature analysis of injection rate of selected Rotokawa wells

Sections 3.1 and 3.2 investigated the effect of total injection rate, and the injection rates for Well #2 and Well #3 separately. We can use feature analysis and significance metrics to decompose this further and attribute relative effects to the three individual wells. Figure 5 plots the summed log p-values for seismicity across different time-scales and using features computed only using data from each of the three wells. Larger associations between injection and seismicity plot as larger negative numbers.

The vertical separation between the curves indicate that injection into Well #3 has a much larger association with microseismicity than the other two wells when a 50th percentile threshold is used to separate high and low rates. As previously indicated, this is a negative association attributed to the shift of injection to Well #2. For λ_{th} = 80th percentile, Well #3 has the stronger association with microseismicity, particularly for Δt_i larger than 6 days (i.e., seismicity rates averaged over about a week or longer). In contrast, there are no clear associations between Well #1 injection and seismicity. As also described in Section 3.1, there is a reduced association for short-term seismicity rates, and this is likely due to inherent randomness when inferring rates from small event counts.

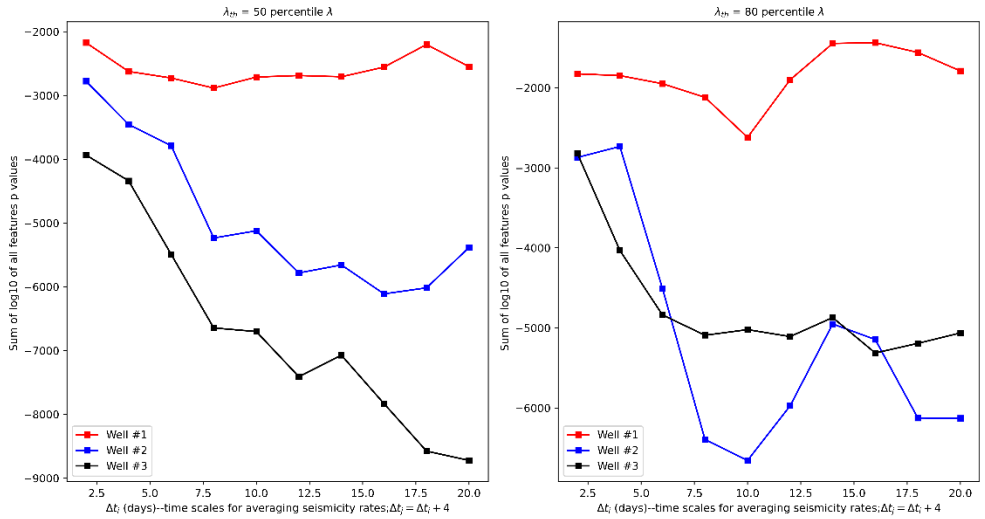


Figure 5: Sum of log_{10} p-values for all features under different Δt_i for different wells

ISSN 2703-4275

3.4 Features comparison for different data types

Having identified Well #3 as one well of interest, in this section, we compare the relative importance of the different data streams derived from this well. Following the approach of Section 3.3, we compute the sum of log p-values for the different data types and compare their values across different earthquake averaging rates.

3.4.1 Comparison of features of Q and dQ/dt at short-time seismicity rates

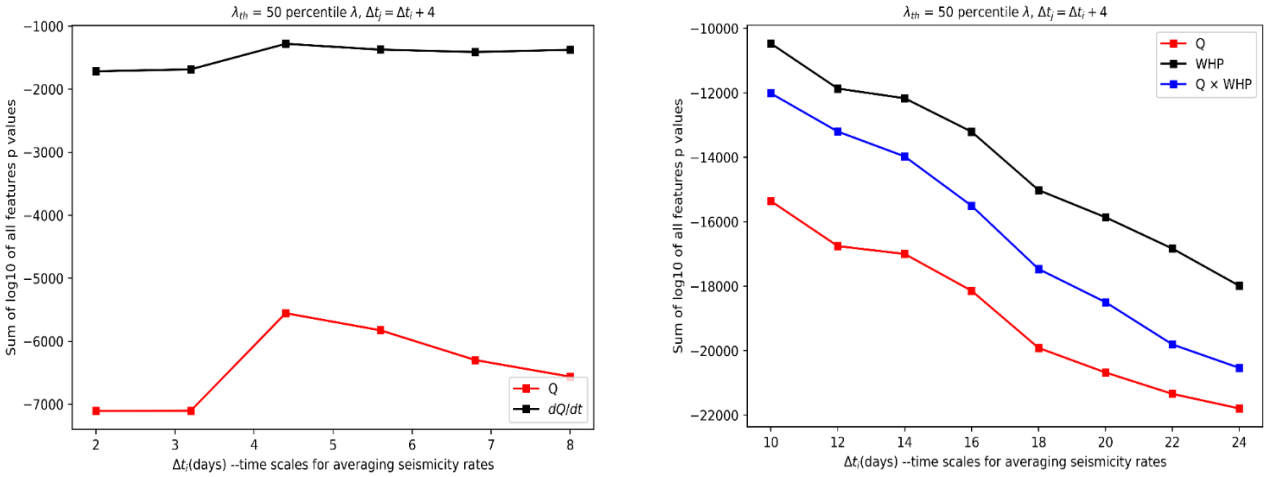
Segall and Lu (2015) proposed that rapid cycling of injection rates can induce transient poroelastic stresses in reservoirs that enhance earthquake triggering. The seismicity transient is proposed to be short-lived, occurring shortly after the injection rate change. We test this hypothesis by looking for overall association between short-term seismicity rate (focusing on averages between 2 and 8 days) and the derivative of injection rate, which will have large values for rapid throttling.

Fig.6 (a) shows that the sum of log p-values for dQ/dt are much larger than Q, indicating no elevated association with injection rate changes. Indeed, this data type appears to be a poorer predictor than Q on its own. In summary, we find no clear evidence that well shut-in or restarts produces statistically elevated earthquake rates.

We acknowledge that absence of evidence is not evidence of absence. It may be that the data we have used are not sensitive enough to detect a poroelastic transient, particularly given that we had earlier indicated that short-term earthquake rates are contaminated by stochasticity. The corollary is that poroelastic transients, if present, are too small to be of concern. It may also be the case that Rotokawa is not as poroelastically sensitive as other reservoirs where this mechanism may prevail.

3.4.2 Feature analysis of Q, WHP and $WHP \times Q$ for long-term earthquake rate

Wellhead pressure can be a good predictor of microseismicity rates, being directly linked with the stability conditions on faults that promote earthquake triggering. We computed summed log p-values for WHP and hydraulic energy (product of pressure and injection rate) for Well #3 and compared these to Q across similar seismicity rates in Sections 3.1 to 3.3. As shown in Fig.6 (b), we found some improved association with WHP and hydraulic energy, but they were not as large as with injection rate alone. Further work is required to make full use of wellhead pressure data, particularly as it does not necessarily correlate with downhole pressure conditions if the well is not in a liquid condition.



(a) **(b)**

Figure 6: Sum of $\log_{10} p$-values comparison of all features under different rata type of Well #3, (a). under different short-term Δt_i for Q and dQ/dt, (b). under different long-term Δt_i for Q, WHP and $Q \times WHP$.

4. CONCLUSION

We have undertaken a machine-learning study of injection and microseismicity data in the Rotokawa geothermal field. Our goal was to infer associations and patterns between the complex injection schedule involving multiple wells and the time-varying seismicity. Our main conclusions are summarized as:

1. There is an observed correlation between long-term total injection and long-term averaged microseismicity rates in the field. The short-term earthquake rate is likely dominated by randomness.
2. In general, larger injection rates into Well #2 are associated with elevated microseismicity. However, it is not clear whether this is a reliable relationship or an incidental association between microseismicity and reservoir operations.
3. Negative and positive associations of Well #3 and Well #2 with microseismicity could reflect underlying connectivity of these wells to seismogenic structures in the reservoir.
4. We find no evidence that rapid injection rate changes trigger elevated seismicity. Further, injection rate data appears to be the most informative parameter for microseismicity.

Future applications of this work include taking a higher-resolution analysis of individual sectors within the field, and focusing on subsections of the injection schedule. The analysis could be improved by further filtering of the data to exclude earthquakes that might be associated with the production area. Finally, we still need to test whether there are sufficient associations in the data to develop classification models with predictive capability. In the first instance, these could be useful to define injection rates and schedules that limit microseismicity to desired levels. Such applications could be useful in other geothermal fields where induced seismicity has been of a nuisance magnitude. Most seismicity at Rotokawa is smaller than magnitude 2 and these small events are unlikely to be felt outside the immediate vicinity.

ACKNOWLEDGEMENTS

We thank Mercury, NZ. for providing injection data for the Rotokawa geothermal field that made this research possible. We also thank Manuel Rivera for discussions that improved our understanding of the Rotokawa field. This research was supported by the MBIE Endeavour "Empowering Geothermal" research program.

REFERENCES

Bannister, S., Sherburn, S., & Bourguignon, S. (2016). Earthquake swarm activity highlights crustal faulting associated with the Waimangu–Rotomahana–Mt Tarawera geothermal field, Taupo Volcanic Zone. *Journal of Volcanology and Geothermal Research*, *314*, 49–56. https://doi.org/10.1016/J.JVOLGEORES.2015.07.024

Bialas, J., Oommen, T., Rebbapragada, U., & Levin, E. (2016). Object-based classification of earthquake damage from high-resolution optical imagery using machine learning. *Journal of Applied Remote Sensing*, *10*(3), 036025. https://doi.org/10.1117/1.JRS.10.036025

Bryan, C. J., Sherburn, S., Bibby, H. M., Bannister, S. C., & Hurst, A. W. (2010). Shallow seismicity of the central Taupo Volcanic Zone, New Zealand: Its distribution and nature. *Http://Dx.Doi.Org/10.1080/00288306.1999.9514859*, *42*(4), 533–542.

Christ, M., Kempa-Liehr, A. W., & Feindt, M. (2016). *Distributed and parallel time series feature extraction for industrial big data applications*. https://arxiv.org/abs/1610.07717v3

Clarke, D., Townend, J., Savage, M. K., & Bannister, S. (2009). Seismicity in the Rotorua and Kawerau geothermal systems, Taupo Volcanic Zone, New Zealand, based on improved velocity models and cross-correlation measurements. *Journal of Volcanology and Geothermal Research*, *180*(1), 50–66. https://doi.org/10.1016/j.jvolgeores.2008.11.004

Corbi, F., Sandri, L., Bedford, J., Funiciello, F., Brizzi, S., Rosenau, M., & Lallemand, S. (2019). Machine Learning Can Predict the Timing and Size of Analog Earthquakes. *Geophysical Research Letters*, *46*(3), 1303–1311. https://doi.org/10.1029/2018GL081251

Dempsey, D. E., Cronin, S. J., Mei, S., & Kempa-Liehr, A. W. (2020). Automatic precursor recognition and real-time forecasting of sudden explosive volcanic eruptions at Whakaari, New Zealand. *Nature Communications 2020 11:1*, *11*(1), 1–8. https://doi.org/10.1038/s41467-020-17375-2

Dempsey, D., & Riffault, J. (2019). Response of Induced Seismicity to Injection Rate Reduction: Models of Delay, Decay, Quiescence, Recovery, and Oklahoma. *Water Resources Research*, *55*(1), 656–681. https://doi.org/10.1029/2018WR023587

Dempsey, D., & Suckale, J. (2016). Collective properties of injection-induced earthquake sequences: 1. Model description and directivity bias. *Journal of Geophysical Research: Solid Earth*, *121*(5), 3609–3637. https://doi.org/10.1002/2015JB012550

Deng, K., Liu, Y., & Chen, X. (2020). Correlation Between Poroelastic Stress Perturbation and Multidisposal Wells Induced Earthquake Sequence in Cushing, Oklahoma. *Geophysical Research Letters*, *47*(20), e2020GL089366. https://doi.org/10.1029/2020GL089366

DeVries, P. M. R., Viégas, F., Wattenberg, M., & Meade, B. J. (2018). Deep learning of aftershock patterns following large earthquakes. *Nature 2018 560:7720*, *560*(7720), 632–634. https://doi.org/10.1038/s41586-018-0438-y

Eaton, D. W., & Igonin, N. (2018). What controls the maximum magnitude of injection-induced earthquakes? *The Leading Edge*, *37*(2), 135–140. https://doi.org/10.1190/TLE37020135.1

Frank, J., Rebbapragada, U., Bialas, J., Oommen, T., & Havens, T. C. (2017). Effect of Label Noise on the Machine-Learned Classification of Earthquake Damage. *Remote Sensing 2017, Vol. 9, Page 803*, *9*(8), 803. https://doi.org/10.3390/RS9080803

Giardini, D. (2009). Geothermal quake risks must be faced. *Nature 2009 462:7275*, *462*(7275), 848–849. https://doi.org/10.1038/462848a

He, M., Li, Q., & Li, X. (2020). Injection-Induced Seismic Risk Management Using Machine Learning Methodology – A Perspective Study. *Frontiers in Earth Science*, *0*, 227. https://doi.org/10.3389/FEART.2020.00227

Healy, J. H., Rubey, W. W., Griggs, D. T., & Raleigh, C. B. (1968). The denver earthquakes. *Science*, *161*(3848), 1301–1310.

Hofmann, H., Zimmermann, G., Zang, A., & Min, K. B. (2018). Cyclic soft stimulation (CSS): a new fluid injection protocol and

traffic light system to mitigate seismic risks of hydraulic stimulation treatments. *Geothermal Energy*, *6*(1). https://doi.org/10.1186/s40517-018-0114-3

Holtzman, B. K., Paté, A., Paisley, J., Waldhauser, F., & Repetto, D. (2018). Machine learning reveals cyclic changes in seismic source spectra in Geysers geothermal field. *Science Advances*, *4*(5), eaao2929. https://doi.org/10.1126/SCIADV.AAO2929

Hopp, C., Sewell, S., Mroczek, S., Savage, M., & Townend, J. (2019). Seismic Response to Injection Well Stimulation in a High-Temperature, High-Permeability Reservoir. *Geochemistry, Geophysics, Geosystems*, *20*(6), 2848–2871. https://doi.org/10.1029/2019GC008243

Hopp, C., Sewell, S., Mroczek, S., Savage, M., & Townend, J. (2020). Seismic response to evolving injection at the Rotokawa geothermal field, New Zealand. *Geothermics*, *85*, 101750. https://doi.org/10.1016/J.GEOTHERMICS.2019.101750

Illsley-Kemp, F., Barker, S. J., Wilson, C. J. N., Chamberlain, C. J., Hreinsdóttir, S., Ellis, S., Hamling, I. J., Savage, M. K., Mestel, E. R. H., & Wadsworth, F. B. (2021). Volcanic Unrest at Taupō Volcano in 2019: Causes, Mechanisms and Implications. *Geochemistry, Geophysics, Geosystems*, *22*(6), e2021GC009803. https://doi.org/10.1029/2021GC009803

King Hubbert, M., & Rubey, W. W. (1959). Role of fluid pressure in mechanics of overthrust faulting: I. Mechanics of fluid-filled porous solids and its application to overthrust faulting. *Geological Society of America Bulletin*, *70*(2), 115–166.

Limbeck, J., Bisdom, K., Lanz, F., Park, T., Barbaro, E., Bourne, S., Kiraly, F., Bierman, S., Harris, C., Nevenzeel, K., Bezemer, T. den, & Elk, J. van. (2021). Using machine learning for model benchmarking and forecasting of depletion-induced seismicity in the Groningen gas field. *Computational Geosciences 2021 25:1*, *25*(1), 529–551. https://doi.org/10.1007/S10596-020-10023-0

Lubbers, N., Bolton, D. C., Mohd-Yusof, J., Marone, C., Barros, K., & Johnson, P. A. (2018). Earthquake Catalog-Based Machine Learning Identification of Laboratory Fault States and the Effects of Magnitude of Completeness. *Geophysical Research Letters*, *45*(24), 13,269-13,276. https://doi.org/10.1029/2018GL079712

Majer, E., Nelson, J., Robertson-Tait, A., Savy, J., & Wong, I. (2012). Protocol for addressing induced seismicity associated with enhanced geothermal systems. *US Department of Energy*, *52*.

McNamara, D. D., Sewell, S., Buscarlet, E., & Wallis, I. C. (2016). A review of the Rotokawa Geothermal Field, New Zealand. *Geothermics*, *59*, 281–293. https://doi.org/10.1016/J.GEOTHERMICS.2015.07.007

Mehrabifard, A., & Eberhardt, E. (n.d.). *Investigation of the Dependence of Induced Seismicity Magnitudes on Differential Stress and Pore Pressure Using Supervised Machine Learning, Northeastern British Columbia (NTS 093, 094A, B, G, H) and Globally*.

Rouet-Leduc, B., Hulbert, C., & Johnson, P. A. (2018). Continuous chatter of the Cascadia subduction zone revealed by machine learning. *Nature Geoscience 2018 12:1*, *12*(1), 75–79. https://doi.org/10.1038/s41561-018-0274-6

Segall, P., & Lu, S. (2015). Injection-induced seismicity: Poroelastic and earthquake nucleation effects. *Journal of Geophysical Research: Solid Earth*, *120*(7), 5082–5103. https://doi.org/10.1002/2015JB012060

Sepulveda, F., Siega, C., Lim, Y. W., Urgel, A., & Boese, C. (2016). Oservations of deep and shallow seismicity within the Wairakei-Tauhara geothermal system. *Proceedings 38th New Zealand Geothermal Workshop*.

Sherburn, S., Bourguignon, S., Bannister, S., Sewell, S., Cumming, B., Bardsley, C., Quinao, J., & Wallis, I. (2013). Microseismicity At Rotokawa Geothermal Field, 2008 TO 2012. *35th New Zealand Geothermal Workshop: 2013 Proceedings*. https://www.researchgate.net/publication/258764853

Turuntaev, S. B., & Riga, V. Y. (2018). Rate-state based simulation of laboratory and natural-induced seismicity. In *SEG Technical Program Expanded Abstracts 2018* (pp. 5002–5006). Society of Exploration Geophysicists.

Wang, R., Schmandt, B., Zhang, M., Glasgow, M., Kiser, E., Rysanek, S., & Stairs, R. (2020). Injection-Induced Earthquakes on Complex Fault Zones of the Raton Basin Illuminated by Machine-Learning Phase Picker and Dense Nodal Array. *Geophysical Research Letters*, *47*(14), e2020GL088168. https://doi.org/10.1029/2020GL088168

Webb, T. H., Ferris, B. G., & Harris, J. S. (2012). The Lake Taupo, New Zealand, earthquake swarms of 1983. *Http://Dx.Doi.Org/10.1080/00288306.1986.10422160*, *29*(4), 377–389.

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. https://doi.org/10.1029/2020GL089651

Zhang, X., Zhang, J., Yuan, C., Liu, S., Chen, Z., & Li, W. (2020). Locating induced earthquakes with a network of seismic stations in Oklahoma via a deep learning method. *Scientific Reports 2020 10:1*, *10*(1), 1–12. https://doi.org/10.1038/s41598-020-58908-5