Geophys. J. Int. (?) **?**, **??–??**

Wastewater-Induced Seismicity in California and Oklahoma

My Dinh, Michael Jetsupphasuk, Lisa Jian, Frank Mei

in original form 2018 April 3

SUMMARY
Given evidence of wastewater injection-induced seismicity in the central and eastern United States and the high levels of wastewater injection in California, we develop a generalized statistical method to test the association of wastewater injection with seismicity. Since previous studies have relied on some parametric model to fit the data and since earthquake modeling has been shown to be inaccurate, we propose a completely non-parametric statistical test. The proposed test statistic is an adjustment to the Spearman rank cross-correlation. Through simulation, the test statistic is shown to have moderate power but a potentially high false discovery rate. Therefore, caution must be taken when interpreting results. We apply the test to the California and Oklahoma and identify 10 regions out of 87 in California and 17 out of 84 in Oklahoma with suspected induced seismicity.

Key words: induced seismicity, earthquakes

1 INTRODUCTION

When oil is extracted from an oil well, large amounts of water are produced as a byproduct. Since the water is too saline to be used for consumption or irrigation, it is often injected into an injection well in the ground. In the central and eastern United States, there exists scientific evidence that wastewater injection has, in several instances, induced earthquakes, some with moment magnitudes greater than 4.0 [8, 10, 14, 17].

The process for how wastewater may cause seismic activity is approximately known [8]. Injected wastewater can cause an increase in pore pressure in the surrounding rock, which, if along a fault, can increase seismic hazard. Additionally, wastewater-induced seismicity can also occur if an injection well changes the loading conditions on a fault. However, measuring the changes in pore pressure, the stress along a fault, and the seismic hazard are difficult and are not generalized to all geological formations. Therefore, the observation of wastewater-induced seismicity in the central and eastern U.S. is not necessarily externally valid to other regions.

In California, wastewater injection volumes are higher than in Oklahoma [12]. Although there have been some attempts to define a general and casual link between injections and seismicity in California, these studies make several assumptions about the nature of induced seismicity that may not be necessarily accurate and proven, which we discuss below. Additionally, statistical methods performed in Oklahoma have similar problems on relying on strong assumptions.

Since the interaction between wastewater injection and the local geology on seismicity rates is not completely understood, it is difficult and perhaps inappropriate to compare regions with different geologies. Therefore, for a generalizable statistical test, one should analyze over the time dimension rather than the space dimension. For instance, in (Hough 2015) [13], statistical significance of the spatial proximity of earthquakes and injection wells was tested by Monte Carlo simulation where synthetic earthquakes were randomly shuffled within Oklahoma and the distance to fixed wells was measured. However, since earthquakes are not uniform in space, treating them as such would result in an anti-conservative p-value. There is a possibility of some confounding variable that influences injection wells to be correlated in space with earthquakes.

When modeling earthquakes, researchers often use a de-clustering method to separate main shocks from aftershocks in order to treat the remaining catalog as independent events from a Poisson distribution. For instance, (Goebel, et al. 2015) [11] implements the method proposed by (Gardner and Knopoff 1974) [2]. However, there is evidence that this method of de-clustering may not produce a Poisson distribution [6]. Other researchers have used the Epidemic Type Aftershock Sequence (ETAS) model to estimate background rates [4, 23]. Nevertheless, since we attempt to make as few assumptions as possible, we follow the example set by (McClure 2017) and (Langenbruch 2016) and do not de-cluster the catalog [16, 18]. And we include all earthquakes above our magnitude of completeness given that aftershocks are also just as socially significant as main shocks. Additionally, our test statistic is implicitly able to relate water injections to both main shocks and aftershocks since we incorporate information from several lags (see section 2.3.2 for more details).

In (Goebel, et al. 2015) [11], the ”Objective Induced Seismicity Correlation Model” is introduced. Although the paper applies it to Kern County in California, this statistical model is generalizable to any region. In essence, the model examines ”triggers” in wastewater injection rates and if these correlate with changes in seismicity over a certain time window and within a certain radius (both defined empirically), three statistical measures are taken: (1) a Poisson probability is taken based on the background rate; (2) a probability of the temporal significance; and (3) the ”R Statistic” [3, 4] which measures the relative change of time interval lengths between the first earthquake before and after the aforementioned trigger.

(Goebel, et al. 2015) [11] defines triggers in wastewater injection rates to be over 100 kbbl/month which is inspired by (Frohlich 2012) [7]. However, (Frohlich 2012) mentions the possible lack of external validity of the trigger threshold [7]. The trigger threshold could potentially be treated as another parameter to be computed per region analyzed - however, this may not be easily done so the ability to generalize is questionable.

Also, as (McClure 2017) [18] mentions, there is also a problem with how statistical significance is treated. The test was performed by generating synthetic earthquake data with the assumption of uniformity in space. As (McClure 2017) [18] points out, earthquakes and wells are spatially variable so synthesizing data from a spatially uniform distribution would not be appropriate.

Another approach to a general, statistical method to test induced seismicity is found in (McClure 2017) [18]. Again, there is a reliance on a Poisson model, this time on a catalog that has not been de-clustered. Since the raw catalog contains dependent events (main shocks and after shocks) it would traditionally be inappropriate to model the events with a Poisson distribution. To induce some sort of dependence on previous events, (McClure 2017) includes past earthquakes as a parameter in the model. However, given the lack of research on modeling earthquakes in this manner, we express trepidation assuming this method for a generalized test.

There is also a problem with how (McClure 2017) [18] assess statistical significance. Their study divides California into a grid and then finds a p-value for each block in that grid. At the end of the study, they combine the p-values using Fisher's method. This is potentially problematic since the p-values are not obviously independent - injections in one block can, for instance, cause an earthquake in an adjacent block. The result would be an anti-conservative overall p-value [1]. Additionally, we question the utility of reporting an overall state-wide p-value. Given that the nature of induced seismicity is highly region-specific, it may not be useful to speak in terms of an entire state.

This paper draws inspiration from these studies in an effort to propose a generalized, non-parametric method to measure wastewater-induced seismicity while making as few assumptions as possible about the underlying physics.

The paper is organized as follows: first, in section 2, we outline our methodology and test statistic; in section 3, we test our test statistic on simulated data to check the validity of our statistic and report our results; in section 4 we apply our test on the real data in California and Oklahoma and report our results; in section 5, we acknowledge the limitations of our design and discuss potential future improvements.

2 METHODOLOGY

2.1 General Framework

Like (McClure 2017) [18], our design is a longitudinal study rather than a cross-sectional study. With the assumption that confounding variables relating to space are stable over time, the remaining relevant variable is wastewater injection.

Below is a rough outline of our study design:

(i) We set up our design according to the method described by (McClure 2017) [18]. We overlay a grid with blocks of size 0.2° latitude by 0.2° longitude over the states of California and Oklahoma. The size of the blocks is chosen to be consistent with the spatial scales used in (McClure 2017) and in other investigations [18]. The purpose of this is to uniquely associate wells and earthquakes so blocks can be analyzed independently. The drawback of this layout, as mentioned in the introduction, is that earthquakes and wells on the borders of the blocks do not relate to that of neighboring blocks even if it may make more sense for them to do so. To examine the sensitivity of our design to this issue, we also run our test on a grid shifted by 0.1°.

(ii) We filter the blocks to analyze only the ones with both non-zero water injection and non-zero seismicity.

(iii) Within each block, we compute our test statistic.

(iv) Finally, for each block, we run a permutation test by permuting earthquake data with a moving-block method to generate a null distribution of our test statistic, from which we obtain a p-value that measures the significance of our observed test statistic.

2.2 Data

The time periods analyzed for both states are different due to differences in the availability of data. For California, the time analyzed is between January 1980 and June 2017. For Oklahoma, it is between January 2011 and December 2016. Unlike (McClure 2017) [18], we use monthly data to maintain as much detail as possible, since the time scale of wastewater-induced seismicity is unknown and variable.

Earthquake data from both California and Oklahoma are collected from the United States Geological Survey's earthquake catalog [20]. The minimum magnitude used is 2.5 for California and 3.0 for Oklahoma [5, 18]. Since the available earthquake data are in different magnitude scales and there exists no general way to convert magnitudes from one scale to another, we cannot compare the relative sizes of earthquake events. Thus, our analysis only looks at counts of earthquakes and disregards magnitude. We do not decluster the catalog for the reasons mentioned in the introduction.

Injection well data for California were collected from California's Department of Conservation - Division of Oil, Gas, and Geothermal Resources (DOGGR) [21]. For Oklahoma, they were collected from the Oklahoma Corporate Commission [22].

Geographical map data were collected using the `get_map` function in the R package `ggmap` [9].

2.3 Spearman Rank Cross-Correlation

In order to measure the positive dependence structure between seismicity and wastewater injection within a block, we want to construct a statistic that incorporates the following objectives:

(i) explicitly defines our desired dependence structure between seismicity and wastewater;

(ii) adjusts for the possibility that the induced seismicity depends on wastewater injection, with an unknown time lag.

For more details about this section, please see the Appendix.

2.3.1 *Dependence*

Positive dependence between seismicity and wastewater injection is difficult to define, as many different dependence structures exist. In this paper, we look for a reasonable dependence structure, that is, a positive, monotonic relationship between wastewater injection and seismicity, measured by Spearman's rank correlation.

Spearman's correlation measures the monotonic relationship between two variables. It is a more relaxed measure of correlation than Pearson's correlation, which assesses linear relationships. It is also less sensitive to outliers. We adjust the statistic to compute cross-correlations; that is, correlation is assessed after shifting one of the series by an appropriate lag.

2.3.2 *Time Lag Effect of Induced Seismicity*

We know injection can cause earthquakes some time later, but the exact time lag of induced seismicity is unknown. Therefore, we want to search for correlation at multiple time lags.

Therefore, let W be a time series of wastewater injection, and S be a time series of seismicity. We want to search for Spearman's correlation $r_t(W, S)$ at some unknown lag t.

Using the time scales by previous studies like (McClure 2017) [18] as a guideline, we search at time lags of length up to one year. That is, at lag $t \in [1, 12]$, we compute value $T_t = \max(0, r_t(W, S))$.

Then, we incorporate the correlations at different lags into one statistic by taking the 2-norm of all the correlations to balance between weighting values at each time lag equally. Since the norm takes the absolute value of each correlation and we want our statistic to be able to capture only the positive effect of water injection on induced seismicity, we discard the negative correlations by taking values at each lag to be equal to the maximum of 0 and the correlation.

Additionally, incorporating a range of time lags allows our test to be robust to aftershocks. For example, if there is high water injection volume one month, then a main shock next month, and then an aftershock the month after; both the main shock and aftershock will capture the correlation to the water injection.

2.3.3 *Test Statistic*

Let W be a time series of wastewater injection, and S be a time series of seismicity. As a result, we construct our test statistic as:

$$T(W, S) = \|(T_t, t \in [1, 12])\|_2. \tag{1}$$

2.4 Permutation Test

In the discussion of Spearman's correlation, if at least one of the series is exchangeable given the other, then under the null hypothesis, all permutations of one variable are equally likely to be paired with the values of the other. There is no reason to believe wastewater injection is not autocorrelated. Main shocks in the earthquake data, on the other hand, are generally known to be stationary. However, since we do not de-cluster the catalog for main shocks and after-shocks, there is temporal clustering in the data. By permuting moving-blocks, we can control for this clustering pattern. We group the seismicity data by time cells of six months and permute the cells instead. In doing so, the clustering is preserved within a cell and the time cells are exchangeable. We take 10,000 random permutations and calculate a p-value. For more details, please see the Appendix.

3 SIMULATION

To examine the power of the test, we test on simulated data to see how it performs.

The simulation has the following steps:

(i) choose and fix actual water injection data from one block;
(ii) choose a set of values for the parameters based on the Poisson model in (McClure 2017) [18];
(iii) simulate a time series of seismicity based on that model;
(iv) perform the test on the above two time series 500 times to generate a distribution of p-values;
(v) repeat 1-4 until all the desired combinations of values for the parameters to the Poisson model are exhausted;
(vi) report summary statistics for the p-values for each combination of parameters and discuss results.

(McClure 2017) constructed a statistical model for natural and induced seismicity [18]:

$$y_{ij} = \text{Poi}(e^{\mathcal{N}(0,\sigma_i)}(\mu_i + \beta_i x_{ij}) + a y_{i,j-1} e^{\mathcal{N}(0,\sigma_{II})}), \tag{2}$$

where y_{ij} is the number of earthquakes in block i and month j; x_{ij} is the cumulative water injected; σ_i is a measure for the variance of seismicity; μ_{ij} is the rate of natural seismicity; β_i is the dependence on water volume; a is the degree to which earthquakes cluster; and σ_{II} is the variability of the clustering.

Though we criticized this model earlier in the paper, we use it here for the purpose of simulation since we have only five requirements for generating synthetic earthquake data:

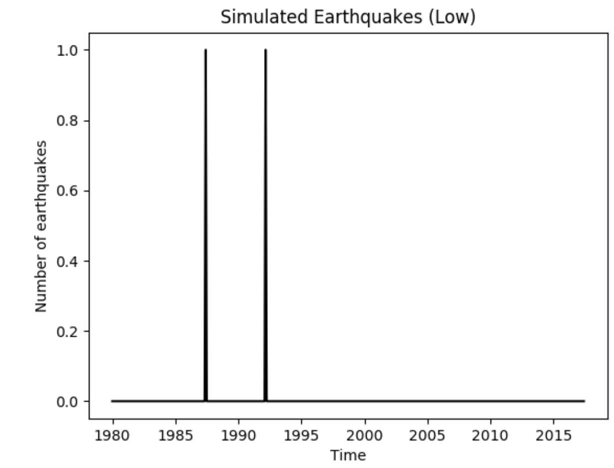
(i) There must be some, implicit or explicit, dependence on wastewater injection data.
(ii) The synthetic data must be similar to the real data. We determine "similarity" here heuristically by visually examining the data.
(iii) There should be some randomness involved.
(iv) The earthquake generation function should increase monotonically with the parameters (i.e. higher water contribution should consistently result in more earthquakes).
(v) The function must be applicable to raw catalogs.

Though we outline our concerns with the above model in the introduction, we find that it meets the above 5 conditions and proceed. We discuss the limitations resulting from using this model in further detail in section 5.

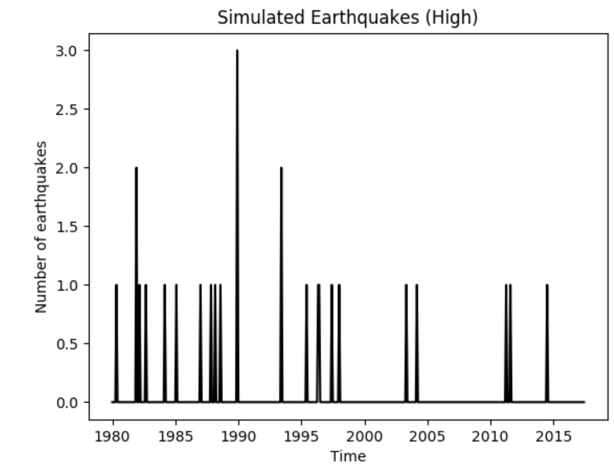
Using this model, we simulate the seismicity data using a slightly adjusted one, which fixed some variables and added multiple historical wastewater terms for a multiple lag effect. Specifically, we fix both sigmas to be 1 and alpha as 0.01. We test combinations of the following parameters: $\mu = 0.003, 0.015$ and β as equal to $7.5e^8$ and $3.75e^7$ split into 3 and 9 parts according to an exponential distribution. The exact values of beta can be seen in the Appendix.

After fixing the water injection sample, these values were chosen to test combinations of low/high background seismicity, low/high water injection contribution, and short/long time effects. Figure 1 below shows an example of simulated "low" seismicity and "high" seismicity as well as the water injection data used in the simulation.

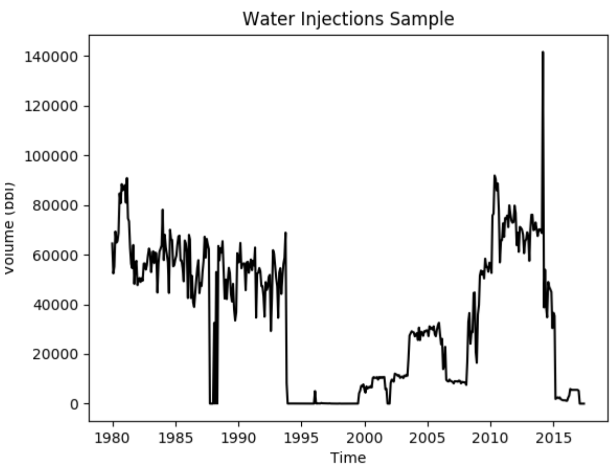
In addition to the 8 cases described above, we also test no background seismicity with low/high water injection contribution and low/high background seismicity with no water injection contribution. For these 4 additional cases, we fix the time effect to be short since from the first 8 simulations we see little difference between the short/long time effects. Results from these 12 cases are reported below in Table 1. "Mean lower" and "Mean upper" refer



(a) simulated seismicity with low background rate, low water contribution, 3 months



(b) simulated seismicity with high background rate, high water contribution, 3 months



(c) water injection used for simulation

Figure 1. Simulation Example

to the average lower and upper 95% confidence intervals for the p-value estimate. "% sig" refers to the percent of p-values out of the 500 simulations equal to or less than the 0.05 threshold.

The results in Table 1 show that under this model, our test statistic is only useful in cases when there is high water contribution and low background seismicity under the Poisson model used. If there is high background seismicity, then even in the case of similarly high water contribution, the test is not useful. We also find that our test is robust to different length time lags as there is little difference between similar cases of 3 months and 9 months. Finally, we test the uncorrelated case where there is no water contribution in the model. We find that the false discovery rate is higher than anticipated. Under a significance level with alpha = 0.05, we would expect the percent of p-values under 0.05 to be 5%. Therefore, our test is anti-conservative and so there should be caution when interpreting results from this test.

4 RESULTS

Note: In this section, since p-values are estimated by 10,000 random draws with replacement from the full permutation group, we determine statistical significance to be p-values whose lower 95% confidence interval bound is less than or equal to 0.05. For convenience, when p-values are mentioned in this section, the lower 95% confidence interval bound is implied.

4.1 California

In California, we apply our test statistic and permutation test and find that out of 87 blocks with both non-zero water injection and non-zero seismicity, 10 (or 11.5%) had p-values less than or equal to 0.05. Unlike (Goebel, et al 2016) [15] who found evidence of induced seismicity near the Tejon Oil Field, we did not detect any. Additionally, there are two blocks on the border of Kern County, a region with high wastewater injection and the area of analysis in (Goebel, et al. 2015) [11], which show induced seismicity in our analysis. There are two locations in common with (McClure 2017) [18] which have significant p-values. Please see Figure 2 for a map of the results. For full results including the coordinates of the significant blocks, please see the Appendix.

4.2 Oklahoma

In Oklahoma, we find that out of 84 blocks, there are 17 (or, 20.2%) with significant p-values at the 0.05 level. Our results are roughly consistent with those found in other research [14]. In comparison to (McClure 2017) [18], we again find 2 locations in common with significant p-values. Please see Figure 2 for a map of the results. See the Appendix for full results.

4.3 Shifted Grid

To examine the sensitivity of our test to the grid layout, we shift the grid in California by 0.1° three times – in the longitudinal direction, in the latitudinal direction, and both – and re-run our test. The percentage of blocks with p-values less than or equal to 0.05 is 14.4%, 15.1%, and 18.1% for each shift, respectively. The large range of blocks with significant p-values shows that our results may be sensitive to the layout of the grid.

Table 1. Simulation Results

	Mean p-values	Mean lower	Mean upper	% sig
Low background, low water, 3 months	0.411	0.403	0.419	25.0%
High background, low water, 3 months	0.497	0.489	0.505	18.2%
Low background, high water, 3 months	0.059	0.055	0.063	78.2%
High background, high water, 3 months	0.184	0.177	0.19	50.4%
Low background, low water, 9 months	0.436	0.428	0.443	24.6%
High background, low water, 9 months	0.531	0.523	0.539	13.8%
Low background, high water, 9 months	0.078	0.074	0.082	74.0%
High background, high water, 9 months	0.175	0.169	0.181	53.4%
No background, low water, 3 months	0.32	0.312	0.327	34.2%
No background, high water, 3 months	0.049	0.046	0.053	81.0%
Low background, no water, 3 months	0.613	0.607	0.62	13.8%
High background, no water, 3 months	0.646	0.638	0.653	9.2%

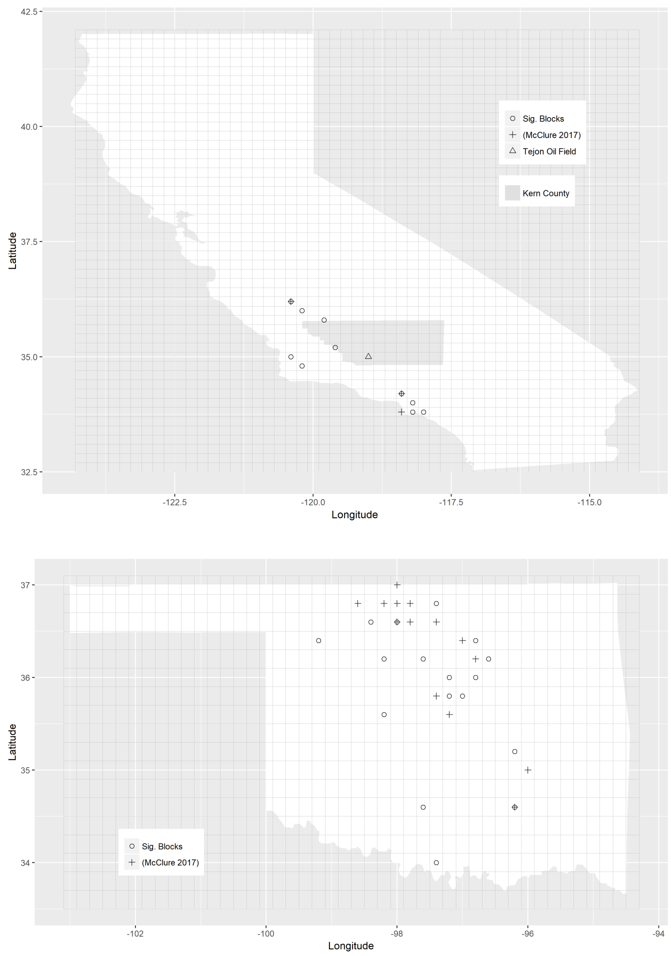


Figure 2. CA and OK Maps with Results

5 OPEN QUESTIONS AND FUTURE WORK

To assess the validity of our test statistic, we checked the performance of our statistic against synthetic earthquake data. As mentioned in section 3, we used the Poisson model proposed by (McClure 2017) [18]. The model provides a convenient tool to generate synthetic data that depends on water injections. At first glance, the model seems to roughly emulate earthquake behavior to a satisfactory degree. However, as noted before, there has been little research to prove the legitimacy of this model. Further, we acknowledge that the relationship between wastewater injections and earthquakes is probably more complicated than that proposed by the model where the mean rises linearly with injections. We recommend further research on the validity of the Poisson model or trying different models to generate synthetic data in order to validate or invalidate our test statistic.

Additionally, the decision to choose six months as the size of our time cells to permute is somewhat arbitrary. Six months was chosen as a somewhat conservative estimate given earthquake behavior. However, there may be a more appropriate method to force exchangeability in this case.

In this study, only counts of earthquakes are considered. There are obvious deficiencies in this decision as a 5.0 magnitude earthquake should not be considered equal to a 3.0 earthquake. In this paper, magnitudes were ignored since they were reported in different scales that were impossible to perfectly relate. However, future research may look into approximating relationships between different magnitude types in order to consider magnitude size in the study design.

Similarly, we did not consider depth (either of earthquakes or of injections) in this analysis. Recent research in Oklahoma suggests a strong relationship between injection depth and induced seismicity [19]. Further research in California concerning this relationship may nuance this study.

As (McClure 2017) [18] mentions in his paper, the grid system could be improved upon since it may not make sense for an earthquake at the border of one block to be associated with an injection well at the other side of the block and not an injection well immediately next it on the other side of the border. Some sort of method that softens the boundaries of the grid may improve upon this model.

Additionally, the association we look for may not be the physically appropriate association. We are essentially testing to see if relatively large amounts of water injection line up with large amounts of earthquakes for some lag. However, the absolute values or cumulative values of water injection may be more relevant. Large changes in water volume may also be more relevant as (Goebel, et al. 2015) used [11].

ACKNOWLEDGMENTS

My Dinh, Michael Jetsupphasuk, Lisa Jian, and Frank Mei each contributed to the design and statistical methods of the study. Although all four members of the team aided in the development in each aspect of the project, the main responsibilities were as follows:

Dinh and Jetsupphasuk conducted the literature review; Jetsupphasuk cleaned the data; Jetsupphasuk, Jian, and Mei coded the simulations; and Jetsupphasuk and Mei wrote this report. Special thanks to Professor Philip Stark for supervising, guiding, and contributing to the project.

APPENDIX A: SPEARMAN RANK CROSS-CORRELATION

A1 Dependence

Spearman's rank correlation, r_s, of two processes, X and Y, is defined as follows:

$$r_s = 1 - \frac{6D}{N^3 - N}, \tag{A.1}$$

where r_s is the rank correlation, N is the total number of observations in the sample and D is defined as follows:

$$D = \sum [\text{rank}(X_i) - \text{rank}(Y_i)]^2. \tag{A.2}$$

Equivalently, we can write r_s in terms of the moments of the ranked data $rg_X = \text{rank}(X)$, $rg_Y = \text{rank}(Y)$, which is Pearson's correlation coefficient computed on ranked data:

$$r_s = \frac{\gamma_{rg_X, rg_Y}}{\sigma_{rg_X}\sigma_{rg_Y}}, \tag{A.3}$$

where

$$\gamma_{rg_X, rg_Y} = \text{cov}(rg_X, rg_Y). \tag{A.4}$$

More specifically, in our context, to compute the cross-correlation for two series at a particular time lag is to compute the correlation between the two series shifted by the appropriate lag. For instance, if we consider the two processes $X = \{x_1, x_2, ..., x_n\}$ and $Y = \{y_1, y_2, ..., y_n\}$ and we want to compute the cross-correlation at lag t, this is equivalent to computing Spearman's correlation for $X_t = \{x_1, x_2, ..., x_{n-t}\}$ and $Y_{-t} = \{y_t, y_{t+1}, ..., y_n\}$, which we define as:

$$r_t(X, Y) = \frac{\gamma_{rg_{X_t}, rk(Y_{-t})}}{\sigma_{rk(X_t)}\sigma_{rk(Y_{-t})}}. \tag{A.5}$$

A statistical test on the correlation itself could be defined where the null hypothesis H_0 is $r_t \neq 0$ at a fixed t. If at least one of the series is exchangeable given the other then under the null hypothesis, all permutations of one variable against the other are equally likely to occur and each r_t generated is equally likely. Therefore a p-value could be generated by computing 1 minus the percentile of the original r_t in relation to the distribution of r_t generated by the permutations.

A2 Time Lag Effect of Induced Seismicity

Initially, we look at the Spearman's cross-correlations of the data at time lags ranging from one month to a year (12 months) and choose the maximum correlation among them. Then, our test statistics is $T = \max_{t \in [1,12]} r_t(W, S)$, where W is wastewater injection and S is seismicity.

However, that statistic ignores the information provided by other lags, which may also have some sensitivity to the overall relation between water injection and earthquakes. For example, if we have one sequence of correlations $\{0, 0, 1, 0.9, 0\}$, and another $\{0, 1, 0, 0, 0\}$, both have a max correlation of 1. But the first sequence also has a correlation of 0.9 at another lag, which is large. And thus we want to give the first one a higher score by incorporating the correlations at other lags in some way.

We can combine different lags with a p-norm. The p-norm of a vector $x = (x_1, x_2, ...x_n)$ is

$$\|x\|_p = (\sum_{i=1}^{n} |x_i|^p)^{1/p}. \tag{A.6}$$

This is a generalization of the method of taking the maximum correlation among lags because as p approaches ∞, the p-norm becomes the maximum, $\|x\|_\infty = \max(|x_1|, |x_2|, ..., |x_n|)$.

Furthermore, the p-norm takes each entry at its absolute value, while our objective is to differentiate a positive and negative correlation given that a positive correlation between data is what is desired. Specifically, there may be cases that the high p-norm value is majorly contributed by a negative correlation taken to its absolute value. Therefore, we choose to discard negative correlations at lags when calculating the statistics, and let $T_t = \max(0, r_t(W, S))$ for each lag t.

Thus, in the end, the test statistics is set to be $T = \|(T_t, t \in [1, 12])\|_2$.

APPENDIX B: PERMUTATION TEST

The null hypothesis H_0 implies that the distribution of T is invariant under permutations of time cells. Let $\mathcal{G}$ be the finite group of transformations that permute among time cells of size 6 months, and let its size $|\mathcal{G}| = G$. For permutation $g \in \mathcal{G}$, $T(X, Y) \sim T(X, gY)$, where X is wastewater injection and Y is seismicity.

We take 10,000 random permutations from the whole permutation group $\mathcal{G}$ and look at the distribution of our test statistic T under these permutations. We find a p-value by measuring how extreme is our statistic for the original data. Let G' be the number of permutations $g \in \mathcal{G}$ such that $T(X, Y) \geq T(X, gY)$. We reject H_0 at level α if $P = \frac{G'}{10000} \leq \alpha$.

Given that we are estimating a permutation test, there is uncertainty in the observed p-value. Since any given draw of the whole permutation group yields a test statistic either below or above our original test statistic, we can model the p-value as a $\text{Binom}(n, \hat{p})$ where n is our sample size, and $\hat{p}$ is the observed p-value. We then compute a 95% confidence interval by the Clopper-Pearson method.

APPENDIX C: BETAS

$$\beta = [4.74090419e-08, 1.74408118e-08, 1.01501462e-08]; \tag{C.1}$$

$$[2.37045210e-07, 8.72040592e-08, 5.07507312e-08]; \tag{C.2}$$

$$\begin{aligned}&[2.12601517e-08, 1.52335644e-08, 1.09153258e-08,\\ &7.82117273e-09, 5.60411515e-09, 4.01552397e-09,\\ &2.87724865e-09, 2.06163875e-09, 5.21125884e-09];\end{aligned} \tag{C.3}$$

$$\begin{aligned}&[1.06300759e-07, 7.61678218e-08, 5.45766292e-08,\\ &3.91058636e-08, 2.80205757e-08, 2.00776199e-08,\\ &1.43862433e-08, 1.03081937e-08, 2.60562942e-08]\end{aligned} \tag{C.4}$$

APPENDIX D: CALIFORNIA TEST RESULTS

Table A1. Final California Results

Grid	P-value	P-value lower bound	P-value upper bound	Longitude	Latitude
358	0.282	0.273	0.291	-124.2	40.6
572	0.765	0.757	0.773	-122.2	39.8
623	0.237	0.229	0.246	-122.2	39.6
624	1.000	1.000	1.000	-122	39.6
676	1.000	1.000	1.000	-121.8	39.4
726	0.137	0.130	0.143	-122	39.2
777	1.000	1.000	1.000	-122	39
778	1.000	1.000	1.000	-121.8	39
880	1.000	1.000	1.000	-121.8	38.6
931	1.000	1.000	1.000	-121.8	38.4
982	0.559	0.549	0.569	-121.8	38.2
1033	0.104	0.098	0.110	-121.8	38
1086	0.694	0.685	0.703	-121.4	37.8
1087	1.000	1.000	1.000	-121.2	37.8
1136	0.700	0.691	0.709	-121.6	37.6
1235	1.000	1.000	1.000	-122.2	37.2
1289	1.000	1.000	1.000	-121.6	37
1347	0.200	0.192	0.208	-120.2	36.8
1398	0.685	0.676	0.694	-120.2	36.6
1446	0.153	0.146	0.160	-120.8	36.4
1448	0.468	0.458	0.478	-120.4	36.4
1449	0.186	0.178	0.193	-120.2	36.4
1450	0.101	0.095	0.107	-120	36.4
1451	0.584	0.574	0.593	-119.8	36.4
1495	1.000	1.000	1.000	-121.2	36.2
1499	0.000	0.000	0.000	-120.4	36.2
1500	1.000	1.000	1.000	-120.2	36.2
1501	0.696	0.687	0.705	-120	36.2
1547	1.000	1.000	1.000	-121	36
1548	1.000	1.000	1.000	-120.8	36
1550	0.742	0.733	0.751	-120.4	36
1551	0.001	0.000	0.002	-120.2	36
1552	0.809	0.801	0.817	-120	36
1553	1.000	1.000	1.000	-119.8	36
1603	0.491	0.481	0.501	-120	35.8
1604	0.048	0.044	0.053	-119.8	35.8
1605	0.521	0.511	0.530	-119.6	35.8
1608	0.874	0.867	0.880	-119	35.8
1654	0.828	0.820	0.835	-120	35.6
1655	0.912	0.906	0.917	-119.8	35.6
1656	1.000	1.000	1.000	-119.6	35.6
1657	0.179	0.171	0.186	-119.4	35.6
1658	1.000	1.000	1.000	-119.2	35.6
1659	0.864	0.857	0.870	-119	35.6
1660	0.477	0.468	0.487	-118.8	35.6
1706	0.714	0.705	0.723	-119.8	35.4
1707	0.878	0.871	0.884	-119.6	35.4
1708	0.329	0.319	0.338	-119.4	35.4
1709	0.142	0.135	0.148	-119.2	35.4
1710	0.240	0.232	0.249	-119	35.4
1711	1.000	1.000	1.000	-118.8	35.4
1753	0.130	0.123	0.137	-120.6	35.2
1758	0.004	0.003	0.005	-119.6	35.2
1759	1.000	1.000	1.000	-119.4	35.2
1760	0.668	0.659	0.677	-119.2	35.2
1761	1.000	1.000	1.000	-119	35.2
1762	0.271	0.263	0.280	-118.8	35.2

Table A2. Final California Results (continued)

Grid	P-value	P-value lower bound	P-value upper bound	Longitude	Latitude
1804	0.572	0.562	0.581	-120.6	35
1805	0.000	0.000	0.000	-120.4	35
1810	0.680	0.671	0.689	-119.4	35
1811	0.481	0.472	0.491	-119.2	35
1812	1.000	1.000	1.000	-119	35
1813	0.717	0.708	0.726	-118.8	35
1855	0.390	0.381	0.400	-120.6	34.8
1856	0.109	0.103	0.115	-120.4	34.8
1857	0.002	0.001	0.003	-120.2	34.8
1908	1.000	1.000	1.000	-120.2	34.6
1915	0.162	0.155	0.170	-118.8	34.6
1962	1.000	1.000	1.000	-119.6	34.4
1963	0.119	0.113	0.125	-119.4	34.4
1964	0.231	0.223	0.239	-119.2	34.4
1965	0.575	0.566	0.585	-119	34.4
1966	1.000	1.000	1.000	-118.8	34.4
1967	1.000	1.000	1.000	-118.6	34.4
1968	0.187	0.179	0.195	-118.4	34.4
2015	0.851	0.844	0.858	-119.2	34.2
2017	0.715	0.706	0.724	-118.8	34.2
2019	0.036	0.032	0.039	-118.4	34.2
2070	0.828	0.821	0.835	-118.4	34
2071	0.000	0.000	0.000	-118.2	34
2072	0.719	0.710	0.728	-118	34
2073	0.455	0.446	0.465	-117.8	34
2121	1.000	1.000	1.000	-118.4	33.8
2122	0.037	0.034	0.041	-118.2	33.8
2123	0.024	0.021	0.027	-118	33.8
2125	0.124	0.117	0.130	-117.6	33.8
2174	1.000	1.000	1.000	-118	33.6

APPENDIX E: OKLAHOMA TEST RESULTS

Table A3. Oklahoma Final Results

Grid	P-value	P-value lower bound	P-value upper bound	Longitude	Latitude
25	0.112	0.106	0.118	-98.2	37
26	1.000	1.000	1.000	-98	37
27	1.000	1.000	1.000	-97.8	37
28	1.000	1.000	1.000	-97.6	37
29	1.000	1.000	1.000	-97.4	37
66	0.730	0.721	0.739	-98.8	36.8
67	0.112	0.106	0.118	-98.6	36.8
68	0.102	0.096	0.108	-98.4	36.8
69	0.603	0.593	0.613	-98.2	36.8
70	0.387	0.378	0.397	-98	36.8
71	0.219	0.211	0.227	-97.8	36.8
72	0.467	0.457	0.477	-97.6	36.8
73	0.028	0.025	0.031	-97.4	36.8
110	0.615	0.605	0.624	-98.8	36.6
111	0.985	0.982	0.987	-98.6	36.6
112	0.013	0.011	0.016	-98.4	36.6
113	0.285	0.277	0.294	-98.2	36.6
114	0.001	0.001	0.002	-98	36.6
115	0.795	0.787	0.802	-97.8	36.6
116	0.885	0.878	0.891	-97.6	36.6
117	0.323	0.314	0.332	-97.4	36.6
118	0.191	0.184	0.199	-97.2	36.6
119	0.964	0.960	0.967	-97	36.6
152	0.013	0.011	0.015	-99.2	36.4
153	1.000	1.000	1.000	-99	36.4
154	0.988	0.986	0.990	-98.8	36.4
155	0.485	0.475	0.494	-98.6	36.4
156	0.433	0.423	0.443	-98.4	36.4
157	0.614	0.604	0.623	-98.2	36.4
158	0.889	0.882	0.895	-98	36.4
159	0.967	0.964	0.971	-97.8	36.4
160	0.895	0.889	0.901	-97.6	36.4
161	0.832	0.824	0.839	-97.4	36.4
162	0.732	0.723	0.740	-97.2	36.4
163	0.538	0.528	0.548	-97	36.4
164	0.003	0.002	0.004	-96.8	36.4
165	0.125	0.119	0.132	-96.6	36.4
198	0.751	0.742	0.759	-98.8	36.2
199	1.000	1.000	1.000	-98.6	36.2
200	0.422	0.413	0.432	-98.4	36.2
201	0.032	0.028	0.035	-98.2	36.2
203	0.804	0.796	0.811	-97.8	36.2
204	0.039	0.036	0.043	-97.6	36.2
205	0.139	0.132	0.146	-97.4	36.2
206	0.756	0.748	0.765	-97.2	36.2
207	0.849	0.841	0.855	-97	36.2
208	0.456	0.446	0.466	-96.8	36.2
209	0.002	0.001	0.003	-96.6	36.2

Table A4. Oklahoma Final Results (continued)

Grid	P-value	P-value lower bound	P-value upper bound	Longitude	Latitude
248	0.371	0.362	0.381	-97.6	36
249	0.393	0.383	0.403	-97.4	36
250	0.018	0.015	0.021	-97.2	36
251	1.000	1.000	1.000	-97	36
252	0.002	0.001	0.003	-96.8	36
292	0.364	0.355	0.374	-97.6	35.8
293	0.266	0.257	0.274	-97.4	35.8
294	0.009	0.007	0.011	-97.2	35.8
295	0.002	0.001	0.003	-97	35.8
296	0.230	0.221	0.238	-96.8	35.8
297	1.000	1.000	1.000	-96.6	35.8
333	0.012	0.010	0.014	-98.2	35.6
337	0.938	0.933	0.943	-97.4	35.6
338	0.196	0.188	0.204	-97.2	35.6
339	0.122	0.116	0.129	-97	35.6
340	0.899	0.893	0.905	-96.8	35.6
381	0.912	0.906	0.917	-97.4	35.4
382	0.430	0.420	0.439	-97.2	35.4
383	0.149	0.142	0.156	-97	35.4
384	0.774	0.765	0.782	-96.8	35.4
385	0.945	0.940	0.949	-96.6	35.4
386	0.163	0.155	0.170	-96.4	35.4
431	0.014	0.012	0.017	-96.2	35.2
467	0.875	0.868	0.882	-97.8	35
468	0.523	0.513	0.533	-97.6	35
474	0.892	0.886	0.898	-96.4	35
476	0.870	0.864	0.877	-96	35
511	0.650	0.641	0.660	-97.8	34.8
556	0.010	0.009	0.013	-97.6	34.6
563	0.017	0.015	0.020	-96.2	34.6
606	0.265	0.256	0.274	-96.4	34.4
648	1.000	1.000	1.000	-96.8	34.2
649	0.414	0.404	0.423	-96.6	34.2
689	0.017	0.015	0.020	-97.4	34
690	0.162	0.155	0.170	-97.2	34
693	0.250	0.241	0.258	-96.6	34

APPENDIX F: CODE AVAILABILITY

Please see our Github repository for code used in this project. R was used primarily for cleaning the data and Python was used for our simulations and test.

REFERENCES

[1] R.A. Fisher. "Statistical methods for research workers". In: (1925).

[2] J. K. Gardner and L. Knopoff. "Is the sequence of earthquakes in Southern California, with aftershocks removed, Poissonian?" In: *Bulletin of the Seismological Society of America* 64.5 (1974), p. 1363. eprint: /gsw/content_public/journal/bssa/64/5/0037110664050001/3/bssa0640051363.pdf. URL: +%20http://dx.doi.org/.

[3] Karen R. Felzer and Emily E. Brodsky. "Testing the stress shadow hypothesis". In: *Journal of Geophysical Research: Solid Earth* 110.B5 (2005). B05S09, n/a–n/a. ISSN: 2156-2202. DOI: 10.1029/2004JB003277. URL: http://dx.doi.org/10.1029/2004JB003277.

[4] Nicholas J. van der Elst and Emily E. Brodsky. "Connecting near-field and far-field earthquake triggering to dynamic strain". In: *Journal of Geophysical Research: Solid Earth* 115.B7 (2010). B07311, n/a–n/a. ISSN: 2156-2202. DOI: 10.1029/2009JB006681. URL: http://dx.doi.org/10.1029/2009JB006681.

[5] Kate Hutton, Jochen Woessner, and Egill Hauksson. "Earthquake Monitoring in Southern California for Seventy-Seven Years (1932â"2008)". In: *Bulletin of the Seismological Society of America* 100.2 (2010), p. 423. DOI: 10.1785/0120090130. eprint: /gsw/content_public/journal/bssa/100/2/10.1785_0120090130/4/423.pdf. URL: +%20http://dx.doi.org/10.1785/0120090130.

[6] Bradley Luen. "Earthquake prediction: Simple methods for complex phenomena". PhD thesis. UC Berkeley, 2010. URL: https://escholarship.org/uc/item/22p7f44k.

[7] Cliff Frohlich. "Two-year survey comparing earthquake activity and injection-well locations in the Barnett Shale, Texas". In: *Proceedings of the National Academy of Sciences* 109.35 (2012), pp. 13934–13938. ISSN: 0027-8424. DOI: 10.1073/pnas.1207728109. eprint: http://www.pnas.org/content/109/35/13934.full.pdf. URL: http://www.pnas.org/content/109/35/13934.

[8] William L. Ellsworth. "Injection-Induced Earthquakes". In: *Science* 341.6142 (2013). ISSN: 0036-8075. DOI: 10.1126/science.1225942. eprint: http://science.sciencemag.org/content/341/6142/1225942.full.pdf. URL: http://science.sciencemag.org/content/341/6142/1225942.

[9] David Kahle and Hadley Wickham. "ggmap: Spatial Visualization with ggplot2". In: *The R Journal* 5.1 (2013), pp. 144–161. URL: http://journal.r-project.org/archive/2013-1/kahle-wickham.pdf.

[10] Richard D. Andrews and Austin Holland. *Statement on Oklahoma Seismicity*. Oklahoma Geological Survey, 2015. URL: http://wichita.ogs.ou.edu/documents/OGS_Statement-Earthquakes-4-21-15.pdf.

[11] T. H. W. Goebel et al. "An objective method for the assessment of fluid injection-induced seismicity and application to tectonically active regions in central California". In: *Journal of Geophysical Research: Solid Earth* 120.10 (2015). 2015JB011895, pp. 7013–7032. ISSN: 2169-9356. DOI: 10.1002/2015JB011895. URL: http://dx.doi.org/10.1002/2015JB011895.

[12] Thomas Göebel. "A comparison of seismicity rates and fluid-injection operations in Oklahoma and California: Implications for crustal stresses". In: *The Leading Edge* 34.6 (2015), pp. 640–648. DOI: 10.1190/tle34060640.1. eprint: https://doi.org/10.1190/tle34060640.1. URL: https://doi.org/10.1190/tle34060640.1.

[13] Susan E. Hough and Morgan Page. "A Century of Induced Earthquakes in Oklahoma?A Century of Induced Earthquakes in Oklahoma?" In: *Bulletin of the Seismological Society of America* 105.6 (2015), p. 2863. DOI: 10.1785/0120150109. eprint: /gsw/content_public/journal/bssa/105/6/10.1785_0120150109/3/2863.pdf. URL: +%20http://dx.doi.org/10.1785/0120150109.

[14] F. Rall Walsh and Mark D. Zoback. "Oklahoma's recent earthquakes and saltwater disposal". In: *Science Advances* 1.5 (2015). DOI: 10.1126/sciadv.1500195. eprint: http://advances.sciencemag.org/content/1/5/e1500195.full.pdf. URL: http://advances.sciencemag.org/content/1/5/e1500195.

[15] T. H. W. Göebel et al. "Wastewater disposal and earthquake swarm activity at the southern end of the Central Valley, California". In: *Geophysical Research Letters* 43.3 (2016), pp. 1092–1099. ISSN: 1944-8007. DOI: 10.1002/2015GL066948. URL: http://dx.doi.org/10.1002/2015GL066948.

[16] Cornelius Langenbruch and Mark D. Zoback. "How will induced seismicity in Oklahoma respond to decreased saltwater injection rates?" In: *Science Advances* 2.11 (2016). DOI: 10.1126/sciadv.1601542. eprint: http://advances.sciencemag.org/content/2/11/e1601542.full.pdf. URL: http://advances.sciencemag.org/content/2/11/e1601542.

[17] Manoochehr Shirzaei et al. "Surface uplift and time-dependent seismic hazard due to fluid injection in eastern Texas". In: *Science* 353.6306 (2016), pp. 1416–1419. ISSN: 0036-8075. DOI: 10.1126/science.aag0262. eprint: http://science.sciencemag.org/content/353/6306/1416.full.pdf. URL: http://science.sciencemag.org/content/353/6306/1416.

[18] Mark McClure et al. "Identifying potentially induced seismicity and assessing statistical significance in Oklahoma and California". In: *Journal of Geophysical Research: Solid Earth* 122.3 (2017). 2016JB013711, pp. 2153–2172. ISSN: 2169-9356. DOI: 10.1002/2016JB013711. URL: http://dx.doi.org/10.1002/2016JB013711.

[19] "Oklahoma's induced seismicity strongly linked to wastewater injection depth". In: *Science* (2018). ISSN: 0036-8075. DOI: 10.1126/science.aap7911. eprint: http://science.sciencemag.org/content/early/2018/01/31/science.aap7911.full.pdf. URL: http://science.sciencemag.org/content/early/2018/01/31/science.aap7911.

[20] URL: https://earthquake.usgs.gov/earthquakes/search/.

[21] URL: ftp://ftp.consrv.ca.gov/pub/oil/GIS/.

[22] URL: http://www.occeweb.com/og/ogdatafiles2.htm.

[23] "A comparison of long-term changes in seismicity at The Geysers, Salton Sea, and Coso geothermal fields". In: *Journal of Geophysical Research: Solid Earth* 121.1 (), pp. 225–247. DOI: 10.1002/2015JB012510. eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1002/2015JB012510. URL: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2015JB012510.