UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

MULTISCALE SPATIOTEMPORAL ANALYSES OF SEISMICITY AND WASTEWATER INJECTION IN OKLAHOMA

A THESIS

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

MASTER OF SCIENCE

By

JACKSON A HAFFENER
Norman, Oklahoma
2017

MULTISCALE SPATIOTEMPORAL ANALYSES OF SEISMICITY AND WASTEWATER INJECTION IN OKLAHOMA

A THESIS APPROVED FOR THE
CONOCOPHILLIPS SCHOOL OF GEOLOGY AND GEOPHYSICS

BY

_____
Dr. Xiaowei Chen, Chair

_____
Dr. Kyle Murray

_____
Dr. Nori Nakata

Acknowledgements

First I want to thank the ConocoPhillips School of Geology and Geophysics for enabling me to receive my Master of Science degree. I would like to acknowledge all the excellent professors I have had here at the University of Oklahoma and specifically the School of Geology and Geophysics who have made learning about geology and geophysics so enjoyable. I would like to thank my advisor, Xiaowei Chen, for her tenacity, encouragement, and humor. I would like to thank my committee members Kyle Murray and Nori Nakata for reviewing my work. I would like to thank the staff of the School of Geology and Geophysics, and especially Rebecca Fay who is an absolute magician.

I also want to thank fellow researchers who aided this work. Thanks to Kyle Murray for inputs on pressure regions in the state. Thanks to Chen Chen for providing his relocated earthquake catalog. Thanks to Colin Pennington and Yan Qin for listening to my ramblings. Thanks to Xiaowei Chen for her excellent insight and ideas that have guided me throughout my project as well as the fundamental code to cluster seismicity, fits planes to earthquake clusters, and calculate diffusion fits. I would also like to acknowledge the Oklahoma Geological Survey and the ConocoPhillips School of Geology and Geophysics who both provided teaching and research assistant funding for my time at OU.

Finally, I would like to thank my best friends Sarah and Jefferson for always being there, my father and siblings for their inexhaustible support and encouragement, and my mother for watching over me.

Table of Contents

List of Tables

List of Figures

Abstract

The increase in seismicity in Oklahoma since 2009 has been primarily attributed to wastewater disposal, however, the extent and nature of the relationship is not clear. On the state level, there is no statistically significant relationship between injection volume and nearby seismic activity for individual wells. However, there is a clear trend between combined regional injection volumes and seismicity. This relationship requires a minimum association distance of 25 km to emerge and 40 km to stabilize. Analysis of the temporal relationship seen between injection and seismicity shows a clear 215-day delay at the state level and time delay between 0 and 180 days at a 20 km spatial scale. The time delay exhibits spatial variability, which may suggests spatial heterogeneity in the hydraulic fluid properties.

Furthermore, clusters of seismicity are used to understand the underlying triggering processes and fluid movement in Oklahoma. Around 52% of the seismic clusters in Oklahoma show statistically significant diffusive migration , which is an indicator of pore pressure triggering, with a mean diffusivity around 0.05 m^2/s. At the regional scale, there are also signs of diffusive migration away from areas with high injection rates with an average diffusivity of 1.8 m^2/s. The large-scale diffusion likely reflects the pressure front propagation within the Arbuckle Group, compared to the cluster based diffusion showing pressure movement in the crystalline basement within a single fault system. Finally, the individual clusters show the majority of earthquakes occur along steeply dipping seismogenic faults that have strike directions between 30° and 120° E. Many of these seismogenic faults occur on sub-optimal fault orientations

based on fault hazard maps, suggesting current fault hazard assessment may not properly represent the actual hazard.

Chapter 1: Introduction and Background of Induced Seismicity

Oklahoma has experienced a substantial increase in the number of M3+ earthquakes since 2009. In the first years of the seismic increase the cause was largely unknown, with natural causes being put forward as a likely reason for events such as the 2011 M5.7 Prague earthquake (Keller and Holland, 2013). Increasingly, research scientists have multiple lines of evidence that it is not natural seismicity but tied to the disposal of wastewater in the state. The first evidence came from the close proximity of the 2011 Prague sequence to the injection site of actively injecting wastewater disposal wells (Keranen et al., 2013). At a state scale, this is also clearly evident, where in general, the location of densest disposal activity is the location of the densest seismic activity (Figure 1). Further support came from similar cases of induced seismicity being documented elsewhere in the mid-continent (Horton, 2012; Kim, 2013). In both Youngstown, Ohio, and Guy-Greenbrier, Arkansas, there was a clear time dependency where injection activity started or increased followed shortly after by an increase in seismicity. This time dependency is also visible at the state level, the increase and decrease of disposal rates has been paralleled by seismic activity (Figure 1). More recent studies involving pressure modeling (Keranen et al., 2014) or spatiotemporal analysis (Walsh and Zoback, 2015; Weingarten et al., 2015; Yeck et al., 2016) indicate that wastewater disposal in Oklahoma is a major contributing factor for the increase in seismicity.

There is also often a distinction made between triggered and induced seismicity. Triggered seismicity is seismicity resulting from activity causing a stress change that "is only a small fraction of the ambient level" acting on the fault; while induced seismicity

is activity causing "a stress change that is comparable in magnitude to the ambient shear stress acting on a fault" (McGarr, 2002). Although the distinction is worthwhile, it is extremely difficult to show if seismicity is triggered or induced, so in this study the two terms are used interchangeably and the results could be representing either triggered or induced seismicity.

Historically, seismicity was first considered human induced from mining activity in the early 20th century (McGarr, 2002). Since then, seismologists have found cases of induced seismicity tied to petroleum production, reservoir impoundment, high pressure fluid injection, and natural gas production (McGarr, 2002). The first documented case of fluid injection induced seismicity was in the 1960's at Rocky Mountain Arsenal (RMA), Colorado, where an isolated disposal well was drilled into the Precambrian basement (Healy et al., 1968). At RMA seismic activity began approximately seven weeks after injection started and occurred within kilometers of the well (Healy et al., 1968). The level of seismic activity closely followed the bottom hole pressure of the disposal well throughout the five-year period the well was active (Healy et al., 1968). In addition to establishing that fluid injection can induce earthquakes, RMA also showed that events can occur after fluid injection has ceased. The largest three events at RMA, ranging from M5 to M5.5, occurred after the well was shut-in (Healy et al., 1968).

Before 2008, Oklahoma had only low seismic activity, with only eleven M3 or larger earthquakes between 1995 and 2005. This can be compared to an average of thirteen M3+ earthquakes occurring per week in 2015. In the 100 years prior to 2011, the 1952 M5.5 El Reno earthquake (Keller and Holland, 2013) was the largest on

record. However, paleoseismicity has shown Oklahoma has had extremely large earthquakes in the recent past. In 1990 two trench studies were carried out on the Meers fault. Located in southwestern Oklahoma, the Meers fault scarp is approximately 26 km in length and has about 3 meters of vertical throw (Crone and Luza, 1990). The studies found the fault last ruptured between 800 and 1600 years ago with an estimated magnitude of slightly greater than M_w 7 (Crone and Luza, 1990; Kelson and Swan, 1990). There was also a wide range of possible recurrence intervals, with anywhere from 1,300 (Kelson and Swan, 1990) to 100,000 years (Crone and Luza, 1990). The potential of damaging natural earthquakes from the Meers fault has been the main reason Oklahoma has had moderate damage potential in natural earthquake hazard assessments (Petersen et al., 2014). Although it is extremely unlikely that (induced) Coloumb stress changes would be sufficient to trigger the Meers fault, the presence of the fault and the hazard associated with the fault's size makes it pertinent to the study of seismicity in Oklahoma.

The majority of recent earthquakes occurring in Oklahoma are along reactivated faults at shallow (~6 km) depths (McNamara et al., 2015a). Most of these earthquakes are in the crystalline basement but some events occur in the basal sedimentary formations (McNamara et al., 2015a). For most of the state, the Arbuckle Group makes up the majority of sediments directly overlying the crystalline basement. The Arbuckle is also the primary reservoir for wastewater in the state, with about 70% of the disposed fluid in the state going into the Arbuckle (Murray, 2015). Based on its proximity and the volumes of wastewater disposed into it, injection into the Arbuckle and

corresponding communication with the basement is considered a major contributing factor for induced seismicity in Oklahoma (Walsh and Zoback, 2015; OCC, 2015).

The Arbuckle Group is late Cambrian to early Ordovician in age and is primarily comprised of carbonates with a few laterally consistent sandstones (Fritz et al., 2012). The Arbuckle overlies the comparatively thin Timbered Hills Group, which contains the basal Reagan Sandstone and Honey Creek Limestone and sits on top of the Precambrian basement (Fritz et al., 2012). The Arbuckle in ascending order contains six formations: the Fort Sill, Signal Mountain, McKenzie Hill, Cool Creek, Kindblade, and West Spring Creek (Fritz et al., 2012). The carbonates that make up the bulk of the group are restricted, shallow water marine carbonates and contain multiple upward-shoaling sequences (Fritz et al., 2012). The Arbuckle thickness varies greatly, with wells in the northern part of the state near Pawnee County placing the Arbuckle at approximately 1000 feet thick (Brizendine, 2017). While in the southern Oklahoma Ardmore basin, the Arbuckle Group thickens to over 8000 feet (Fritz et al., 2012). The formations in the Arbuckle also show a wide range of porosity and permeability. Arbuckle permeabilities measured in Oklahoma range from 0.16 to over 3,000 millidarcys (mD) (Morgan and Murray, 2015). Well-based porosity estimates show the Arbuckle section having from 4% to 12% porosity in northcentral Oklahoma (Brizendine, 2017). At the state scale, the Arbuckle exhibits a broad range in vertical thickness, fluid capacity, and hydraulic properties, which makes it extremely challenging to accurately model fluid behavior in the Arbuckle.

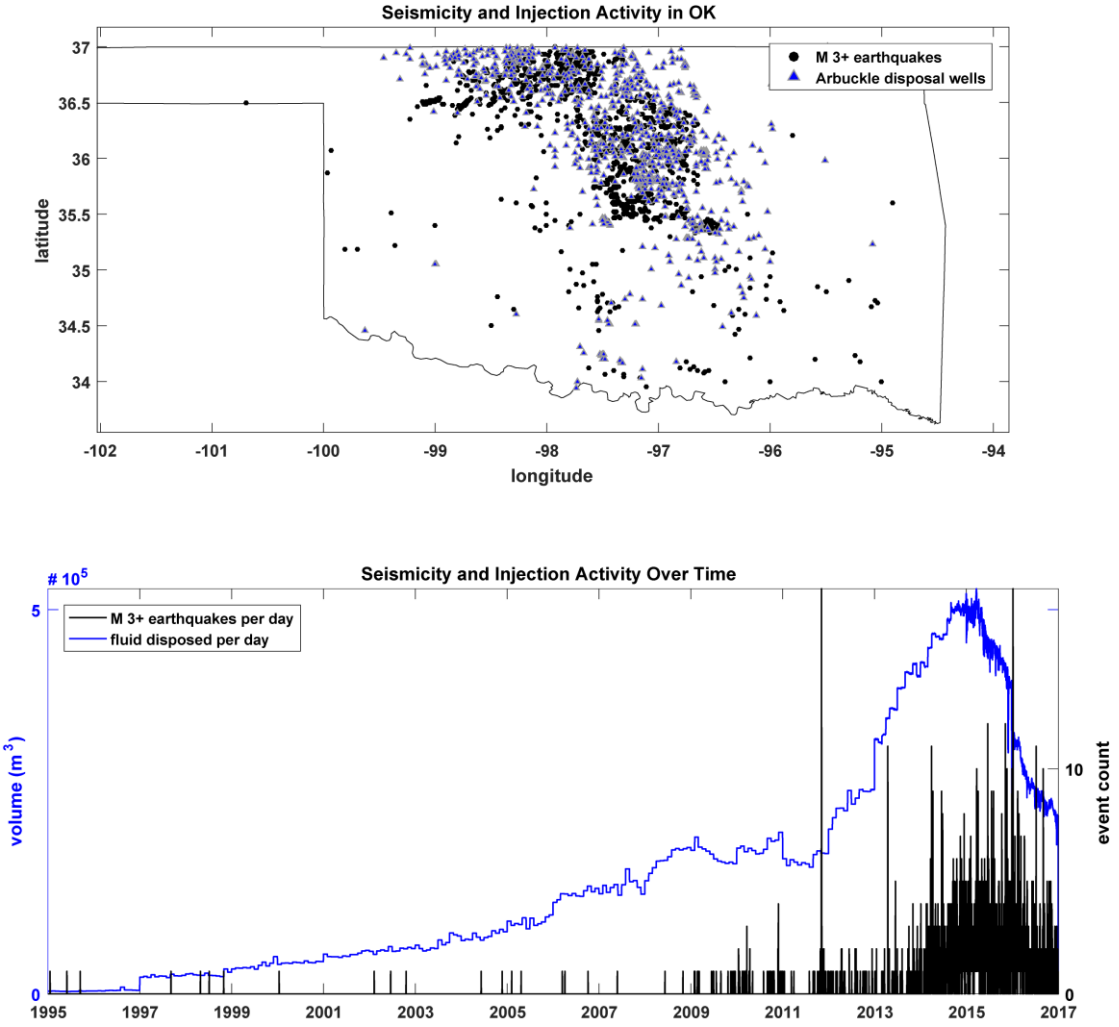

Figure 1. The location of M3+ earthquakes that have occurred in the last 22 years and Arbuckle disposal wells (top). The daily injection volume and number of events above M3 (bottom). The locations of disposal wells and seismicity roughly coincide. The daily volume of fluid disposed in the state has been steadily increasing since 1995, with a rapid increase beginning in 2012 that peaked in 2015. The number of daily events follows a similar pattern with a growth, peak, and decline between 2009 and 2017.

Chapter 2: Spatial Relationship Between Disposal Volume and Seismicity

Introduction

Although the increase in seismicity in Oklahoma is likely due to wastewater disposal, much of the details in the relationship are still uncertain and under debate (Keranen et al., 2014; Weingarten et al., 2015; Walsh and Zoback, 2015). Notably, whether limiting volumes for individual high-rate disposal wells would be more effective versus groups of clustered disposal wells is an open question (Weingarten et al., 2015; Walsh and Zoback, 2015). Although growing evidence shows that clusters of disposal wells have much greater spatial regions of influence (Yeck et al., 2016; Goebel et al., 2017,), the spatial influence region's size has not been systematically quantified.

The OCC originally took a reactionary approach to seismicity, with a traffic light system where disposal wells within 6 miles of an earthquake swarm or 3 miles of a known fault had to have their injection practices reviewed and possibly reduced or halted (OCC, 2015). After the continuation of seismicity, the OCC carried out more wide-spread volume reductions where wells within large 'Areas of Interest' had reductions based on their injection volume, with the goal of reducing the yearly volume to 40% below the 2014 total (OCC, 2016a; OCC, 2016b). These actions reflect the large uncertainty in the spatial scale of the well-earthquake relationship in the state. Accounts of induced seismicity outside of Oklahoma have shown earthquakes occurring up 16 km away from a single injection well (Yeck et al., 2015), while other documented cases either have a maximum distance of less than 15 km (Healy et al., 1968) or involve multiple wells and cannot assume a single distance (Davis and Frohlich, 1993; Horton,

2012; Kim, 2013). With over 850 Arbuckle disposal wells in the state, determining a definitive spatial influence of a single well or even a group of wells is challenging.

Multiple studies have shown that seismicity in Oklahoma is associated with nearby disposal well activities. The 2011 M5.7 Prague sequence may be associated with low-volume disposal wells less than 5 km away (Keranen et al., 2013), although the debate is ongoing (McGarr, 2014; Walsh and Zoback, 2016). A hydrogeologic model of central Oklahoma suggests that four high-rate wells triggered events up to 35 km away (Keranen et al., 2014). In a study of the mid-continent, high-rate disposal wells were found to be significantly more likely to be associated with seismicity within 15 km than low-rate wells (Weingarten et al. 2015). The micro-seismicity rate of the October 2014 Cushing OK earthquake sequence showed a high correlation to the injection rate of three Arbuckle disposal wells located within 10 km of the sequence (McNamara, 2015b). The 2016 M5.1 Fairview earthquake sequence, was likely triggered by multiple high-rate wells located 12–20 km away (Yeck et al. 2016). All of these findings depict a wide range of possible triggering distances.

Recent work has derived theoretical relationships between the volume of fluid disposal and the associated seismicity. McGarr (2014) demonstrated a theoretical maximum moment cap that scaled linearly with the volume disposed and agrees well with historic cases of induced seismicity. Dieterich et al. (2015) found the same scaling as the moment cap model from simulated earthquake catalogs. Using the same data as McGarr (2014), van der Elst et al. (2016) presented a different theory, where the number of events scales with volume, and the maximum expected magnitude is consistent with the extrapolation from Gutenberg-Richter distribution. Hereby these two

models will be referred to as the “moment-cap model” and “sample-size model”, respectively. Although similar, the two theories have important implications to hazard assessment. Putting a cap on the maximum moment possible, without taking into account the regions stress field and seismic potential, has the potential to underestimate the hazard of induced seismicity.

The goals of this chapter is to first determine a spatial scale for the relationship between injection and seismicity; then obtain the empirical relationships between injection operation parameters and earthquake parameters; finally, the empirical observations with model predictions will be compared. With an improved understanding of the spatial influence window and governing equations of the injection-seismicity relationship in Oklahoma, regulators can better shape their reduction guidelines and have a justification for their regulations, which will ultimately reduce the earthquake hazard in Oklahoma and beyond.

Data

The disposal well database used in this thesis is compiled from OCC monthly and daily volume reports. Only the 865 disposal wells that are injecting into the Arbuckle Group are used in this study. Unless otherwise noted, data before 2005 was trimmed due to poor data quality. Disposal data after 2016 is incomplete at the state scale, because disposal wells outside the OCC mandated “Areas of Interest” are not required to report daily injection volume. Based on the 2015 injection rates of wells without reports in 2016, not having the 2016 volumes for wells outside the areas of interest is underestimating 2016 volumes by approximately 7% of the true volume. Because of this, the 2016 data is excluded for the majority of the analysis but it is

included in the discussion to understand the statewide relationships with inclusion of the three M5+ events in 2016 (February M5.1 Fairview, November M5 Cushing, September M5.8 Pawnee). Inconsistencies that appear to be keystroke errors were removed from the volume database by replacing volume entries that were greater than 10 times the mean injection rate of the well with the injection rate just prior to the anomalous rate. The Oklahoma Geological Survey earthquake catalog was used in this study and events below the Mc = 2.5 removed. A synthetic earthquake catalog and well database was also generated to test the statistical significance of the results. The synthetic well database is created by randomly distributing the wells locations within the study area, while keeping the original injection history for each well. The synthetic earthquake catalog is made by randomly distributing the event locations and resampling the magnitude of events following a Gutenberg-Richter distribution. The synthetic datasets remove any spatial dependence of injection and earthquakes in the original datasets.

Oklahoma's seismicity demonstrates a two-step Gutenberg-Richter distribution that is challenging to fit with a single b-value for both small and large earthquakes (Figure 2). This may be due to the different magnitudes reported in the catalog, with moment magnitude (M_w) primarily used for larger earthquakes and local magnitude (M_L) for small earthquakes. For smaller events M_L typically gives larger magnitudes than M_w, with $M_L \propto 1.5\ M_w$ (Deichmann, 2017). The two-step Gutenberg-Richter distribution could also be due to the underlying physical processes governing small and large earthquake occurrences. The Mc is calculated using zmap software (Wiemer and Wyss, 2000), with a minimum Mc = 3.1 for a 95% goodness of fit, Mc = 2.2 for a 90% goodness of fit, and a Mc = 2.3 from maximum curvature. A Mc of 3.1 would make the

analysis difficult due to low sample size, so a Mc of 2.5 that is 0.2 magnitude above the maximum curvature is used (Figure 2). For a b-value estimate the maximum likelihood method was first tested, resulting in a b-value of 1.18, which fits the low magnitude distribution of events well (Aki, 1965). As an alternative estimate, a bootstrapped least-squares fitting method resulted in a b-value of 1.52 that has an improved fit of the higher magnitude events. Due to this wide range between methods, both b-values were used in the comparisons to theoretical equations.

In this analysis all units are converted to SI units: the injection volumes are converted to cubic-meter, and the seismic moment are in the unit of Nm, simply using catalog magnitudes as moment magnitude, as the moment magnitudes for small earthquakes are unavailable. There are often systematic difference in M_L and M_w, so further studies based on more careful moment scaling are required. However, for the purpose of this study, if the difference is systematic for all areas, it should not affect the first-order relationships for the entire state that are being estimated here.

Methods and Results

Relationship between total injected volume and seismicity – individual wells

McGarr (2014) compiled cases of induced seismicity and derived an upper limit on the seismic moment of induced seismicity. He showed that there is a reasonably linear relationship between the two variables where the disposed volume (V) controls the maximum moment (M_0max), with maximum moment falling under the line:

$$M_0max = G\Delta V \quad (1)$$

where G is the shear modulus (~30 GPa). Part of this derivation was the relationship between total moment ($\sum M_0$) and V:

$$\sum M_0 = 2GV \quad (2)$$

When (1) and (2) are viewed on a logarithmic plot, both M_0max and $\sum M_0$ have a slope of one, while the shear modulus dictates the y-intercept. An alternative theory put forward by van der Elst et al. (2016) showed that the volume controlled the total number of events, but the maximum magnitude is expected from the Gutenberg-Richter distribution. They show the most likely maximum magnitude follows:

$$M_{max} = \frac{1}{b}(\Sigma + log_{10}V) \quad (3)$$

where the maximum magnitude depends on the b-value b, the seismogenic index Σ, and the volume. Because of the scaling relationship between seismic moment and moment magnitude, the maximum moment would follow a slope of 1.5/b on logarithmic plot, with a b-value of 1.5 resulting in a linear scaling like equation (1). Both arguments, which will be referred to as the moment cap and sample size hypotheses, show that there is a clear trend between the injected volume and the maximum moment of corresponding events.

First to check if the relationships of equations (1) – (3) can be seen on an individual well scale, the volume V for each individual well was compared to the M_0max of all events within 10 km (Figure 3). Next the V versus M_0max data wee cross-plotted and a linear regression is fit to the $\log_{10}$ of both variables. The same steps were also taken to compare V versus $\sum M_0$. The goodness of fit to the linear regression was quantified using the adjusted coefficient of determination (referred as R^2):

$$R^2_{adj} = 1 - \left(\frac{n-1}{n-p}\right)\frac{SS_E}{SS_T} \quad (4)$$

where the adjusted coefficient of determination R^2_{adj}, is similar to the normal coefficient of determination (one minus the sum of squared error SS_E divided by the sum of

squared total SS_T) except for the modifying factor of the number of observations (n) and the number of regression coefficients (p). The adjusted coefficient of determination takes into account the spurious increase in the coefficient when the number of regression coefficients (in this case two), is of similar order of magnitude to the number of samples (Figure 4).

As seen in Figure 5, there is a very poor relationship between volume and seismicity at the individual well scale, with both V versus M_0max and V versus M_0 showing a poor R^2 fit of the linear regression. There are similar low correlations between the average injection rate and seismicity rate. The low correlation is also persistent with further distances (e.g., up to 30 km). This result suggests that relationship between individual wells and seismicity is statistically insignificant at the state level.

Individual clusters relationship to nearby volume

As the next test of the effect of injection volume, the size of individual clusters was compared to regional volume. For this test, events are first clustered using nearest neighbor rescaled distance after the method of Zaliapin and Ben-Zion (2013). A more detailed explanation of this method is explained in Chapter 4. The clustering resulted in 82 clusters with over 20 events. Next, for each cluster, the total volume of wells within 5 km of the cluster was summed. The log10 of the $\sum M_0$ for each cluster and log10 of the V of nearby wells was fit with a linear regression and the R^2 and slope was calculated following equation (4). This process was repeated from 5 km to 100 km radial distances in 5 km increments. The same steps were also carried out to compare the number of earthquakes (N_{eq}) of a cluster against the V near each cluster. As seen in

Figure 6, there is no statistically significant relationships between the N_{eq} or $\sum M_0$ of an individual cluster and the nearby V at any distance. The lack of correlation for both individual cluster analysis and individual well analysis suggest that it is difficult to establish robust relationship at individual level for the entire state. However, this could also suggest that not all Arbuckle wells contribute to induced seismicity, and there could be cases where individual relationships can be established, but it may require detailed local-scale analysis to establish clear spatial-temporal correlation.

Gridded spatial binning of volume and seismicity

Walsh and Zoback (2016) suggested that it is likely that wells within wider areas contributed to the seismicity (e.g., at the spatial scale for individual counties). In order to quantitatively determine if a relationship becomes apparent at larger scale quantitatively, the relationship between total injection volume and seismicity was examined with increasing spatial windows. The study area (34 to 37 degrees latitude and -100 to -95 degrees longitude) was divided into different grid cells, with the grid sizes ranging from 0.1 degree (approximately 14 km) to 1.2 degree (approximately 172 km) in 0.01 degree increments. Below 0.1 degree the grid size was too small to encompass entire clusters of seismicity (Figure 7), while above 1.2 degree grid size there were too few samples for a statistically significant result. Then, for each fixed grid size, the injection volumes and earthquake data in each cell were summed (Figure 8). Next the relationship was examined for four sets of gridded data: a) V versus N_{eq}, b) V versus $\sum M_0$, c) V versus M_0max, and d) volume disposed before maximum moment (V_{M_0max}) versus M_0max. For each dataset, the following equation was examined: $D = AV^S$, where D is the observational data related to earthquakes, and V is the

corresponding injection volume. To determine the parameters A and S, a linear regression wass fit to the $\log_{10}$ of both variables ($\log_{10} D = \log_{10} A + S \log_{10} V$), and the results were saved for the adjusted R^2, the p-value (to determine if the linear regression model is statistically significant, typically, p-value <0.05 is required to declare a significant relationship), and the slope (=S) and y-intercept (=$\log_{10}$A) of the linear regression. The same steps were also carried out for the synthetic well database and earthquake catalog.

The results show a clear volume-seismicity relationship across all four sets of variables. In all datasets, there is a gradual increase in the R^2 with increasing grid size until around 0.6 degree, where the R^2 plateaus (Figure 9). A similar pattern is apparent in the variation of the slope of the best-fit line with increasing grid size, where the slope of all four datasets rapidly increases to around 0.3 degrees (~ 43 km) grid size (Figure 10). At increasing grid sizes, the slope slowly increases up to around 0.6 degrees (~ 86 km), where the slope no longer changes. The fact that a stable slope to the best fit line occurs above 0.3 degrees suggests 0.3 degrees is the minimum grid size possible to observe the volume – event relationship with this method. Above 0.6 degrees both the R^2 and slope have only minor fluctuations, representing a minimum grid size for a consistent volume-seismicity relationship. The variation in y-intercept with grid sizes is also similar to the variation in slope (Figure 11). The synthetic tests show low R^2 (Figure 9), near zero slope (Figure 10) and high p-values (Figure 11) for all data pairs, confirming the high correlation of the observed data is unlikely to randomly occur. The increase in 95th percentile of the R^2 for large grid sizes using synthetic datasets is almost

entirely due to the reduction in sample points, resulting in a wider distribution of results between iterations (Figure 4).

Unsupervised clustering of wells and spatially binning seismicity

Although the gridded method shows minimum spatial scale to the observed well-seismicity relationship, the square-shaped cells may artificially truncate clustered wells, and are not consistent with the typically assumed radial diffusion model. To circumvent this, a second statistical test was developed to quantify the relationship. Volume-weighted random partition k-means clustering was used to group the disposal wells and examine the relationship between injection volumes for each group with nearby earthquake data. K-means clustering is a way of naturally grouping data based on mean distances (Hamerly and Elkan, 2002). The method used employs the following steps:

1. Randomly place n_s initial seed points, ranging from 5 seeds to 100 seeds in 5 seed increments, throughout the study area.
2. Each well is associated with the seed point closest to it, and seeds that have no associated wells are removed.
3. Each seed point is relocated based on the injection rate weighted average location of its associated wells by:

$$S_x = \frac{\sum_i^N Wx_i * Wv_i}{\sum_i^N Wv_i} \qquad S_y = \frac{\sum_i^N Wy_i * Wv_i}{\sum_i^N Wv_i} \qquad 4)$$

where S is the seed point, N is the number of wells associated with a given seed point, Wx and Wy are the location of each well, and Wv is the average nonzero injection rate of each well.

4. Steps 2 and 3 are iterated until the seed locations no longer change, then the seed is assigned with the summation of volumes from the associated wells.
5. The number of events and total moment of events within a distance d_r of each final seed point is calculated for distances of 5 km to 100 km in 5 km increments.
6. A linear regression is fitted to the $\log_{10}(V)$ vs $\log_{10}(N_{eq})$ and the $\log_{10}(V)$ vs $\log_{10}(\sum M_0)$ of each seed n_s.
7. For each linear regression model, the adjusted R^2, slope, y-intercept, and p-value are calculated for each possible combination of n_s and d_r.
8. Steps 1–7 are repeated 100 times to test the sensitivity of the outcome to the initial seed placement, and obtain the 5, 50, and 95 percentiles for each combination of n_s and d_r.
9. Steps 1–8 are repeated for the synthetic datasets to test the statistical significance of the observations.

The final position of the seed points represents the volume weighted center location of similarly located wells. For a small n_s, the seeds represent very large regional injection zones, for a large n_s, the seeds represent local groups of tens of wells. An example of the clustering of wells and spatial summing of events can be seen in Figure 12. It is worth noting that in contrast to the gridded method, this method can count events multiple times if they fall within the radii of multiple final seed locations.

The 100 iterations of the observed dataset show a peak in both the median R^2 and median slope of the linear regression between 35 and 55 km radii for both V versus $\sum M_0$ and V versus N_{eq} (Figure 13). The max slope of both datasets is within the same

range as the 'stable slopes' reached by large grid sizes in the gridded method. Also observable is the low coefficient of determination ($R^2 < 0.4$) and rapidly decreasing slope for radii below 25 km, again showing a minimum area of influence is required to define the volume-earthquake relationship (Figure 14). The statistical results show low p-values for radii above 10 km for the observed data, suggesting the resulted parameters are significant, however, the synthetic tests have high p-values and a near zero R^2 for all seed number and radii, confirming the observed results are unlikely to occur at random chance (Figure 15).

Empirical relationships

Both the gridded and k-means methods demonstrate a clear volume-seismicity relationship: for the gridded method, the slope stabilizes at 0.3 degrees (or 43 km), while the R^2 and slope plateaus at 0.6 degrees (or 86 km); for the k-means method the R^2 and slope rapidly increases until approximately 25 km, then both stabilize until peak values at about 45 km radius. To compare these results to the "moment cap" and "sample size" hypotheses, the four datasets were again examined: a) V versus N_{eq}, b) V versus $\sum M_0$, c) V versus M_0max, and d) V_{M_0max} versus M_0max (Figure 16). For each dataset, a best-fitting relationship was derived based on the average of the parameters for all grids greater than 0.6 degrees from the gridded method, and a standard deviation was also estimated based on these grids. The range of standard deviations for both the slope and y-intercept (Table 1) were used to create lines of estimated uncertainty. The data of 0.5 and 1 degree grids were used as examples to demonstrate the relationship. For V versus $\sum M_0$, the theoretical relationship of equation (2) was added for comparison. For V versus M_0max, the compiled data of McGarr (2014) as well as the

theoretical relationships of equations (1) and (3) were added for comparison. For equation (3), which has two additional independent variables, the higher b-value from the statewide observation - 1.52, and seismogenic indexes (Σ) of -1.0 from least squares fit were used, to generate the theoretical relationship, which is closer to observation than the lower b-value of 1.18 from maximum-likelihood fit.

The four empirical equations relate volume and seismicity from the average of grid sizes of 0.6 degrees and larger are:

$$\text{A) } N_{eq} = 10^{-2.46} V^{0.62} \quad 6)$$

$$\text{B) } \sum M_0 = 10^{9.64} V^{0.83} \quad 7)$$

$$\text{C) } M_0 max = 10^{10.09} V^{0.68} \quad 8)$$

$$\text{D) } M_0 max = 10^{10.54} V_{M_0 max}^{0.64} \quad 9)$$

Most notable from these empirical equations is that, in Oklahoma, neither N_{eq}, $\sum M_0$, nor $M_0 max$ scales linearly with volume. It is also notable that the best-fit relationship of V versus $M_0 max$ has a shallower slope than the theoretical equations (1) and (2) (Figure 16). In Figure 16(c) and (d), the most notable outlier is the grid containing the M5.7 2011 Prague earthquake for both 0.5 and 1.0 degree grid sizes. With the addition of 2016 data (Figure 17), both the M5.8 Pawnee and M5.1 Fairview earthquakes also represent as outliers for the 0.5-degree grids, while the M5.1 Fairview earthquake moves closer to state average with 1-degree grids. Although these larger earthquakes appear as outliers compared to empirical observations in Oklahoma, they are shifted closer to the compilation of the McGarr (2014) study, and closer to theoretical predictions. However, none of these earthquakes appear as significant outliers in the V versus N_{eq} and V versus $\sum M_0$ datasets compared to state average. So what is

happening? If a uniform b-value for the entire state is assumed, then all grids should follow the same relationship as predicted. Therefore, it seems possible that b-value varies significantly in Oklahoma.

b-value map and location of large magnitude events

To assess the spatial variability of b-value, the nearest 150 events for each earthquake were selected, which were used to estimate a local Mc and b-value using a least-squares fitting method (Figure 18). The results using maximum likelihood method were also examined (Aki, 1965), which exhibit similar spatial patterns of low and high b-values (Figure 19). Next, the b-value distributions with spatial density of earthquakes of different magnitudes is compared. To do so, the earthquake data was spatially gridded for the state on a grid of 0.02-degree (approximately 3 km). For each cell, the number of events above a given magnitude were counted (i.e., M2.5, M3, M3.5, M4, M4.5). The data was smoothed with a Gaussian spatial filter using 3 adjacent cells (approximately 9 km). The data for cells with density greater than the mean of all cells was then contoured (Figure 18 and 19). For example, the M4+ contour lines represent where, after smoothing, the density of M4+ events is greater than the mean of all other cells. Because there are so few M4+ events, these contours center on the epicenters of all M4+ events. The resulting map shows that areas of large magnitude events have lower b-values than areas with only smaller magnitude events, and there exists wide range of b-values for Oklahoma (ranging from 0.8 to 1.8). It is noteworthy that the present analysis maybe affected by the inconsistent magnitude in the catalog as noted before, however, the systematic spatial variation suggests that there likely exists some physical control of the spatial distributions of earthquakes with different sizes. If b-

value is a representation of differential stress (Schorlemmer et al., 2005), then this suggests that large earthquakes tend to occur within areas of high differential stress.

Discussion

Observations at different scales

The individual well analysis shows no statistically significant relationship in the data, and high-volume wells do not tend to have higher levels of seismicity. This suggests that a single well's injection, whether high or low, is likely not solely responsible for inducing seismicity. This is not too surprising: the Arbuckle has high end-member permeability between 3–30 darcys and the average vertical height that disposal wells are injecting into the Arbuckle is 380-meters, implying wells have the potential to rapidly communicate disposed fluids (Carr et al., 1986; Franseen and Byrnes, 2012; Morgan and Murray, 2015). If rapid fluid transport is occurring between wells then the summed injection of a regional groups of wells, whether multiple smaller injectors or a group of large injectors, will be the controlling factor of triggering.

These results seemingly contradict the findings of Weingarten et al. (2015), who found high-rate injection wells are more likely to be spatiotemporally associated with nearby seismicity in the mid-continent. However, in Oklahoma high-rate wells are more often located in areas of dense injection activities. The Arbuckle disposal well catalog shows that wells with average injection rate below 10,000 bbl/month and above 300,000 bbl/month have an average distance to the five nearest wells of 11.8 km and 4.8 km respectively. Furthermore, the 10,000 bbl/month wells have had on average 2.1×10^7 m^3 of total fluid disposed within 15 km, compared to an average of 5.4×10^7 m^3 for wells with an average rate above 300,000 bbl/month. It seems likely the greater

association of high-rate wells to seismicity could be due to these wells are preferentially located in areas of denser disposal well spacing and higher overall regional injection.

Similarly, the cluster based method shows no statistically significant relationship between nearby total injection volume and the seismic activity within each individual cluster at distances from 5 km to 100 km. This would be consistent with the modeling observations from Dieterich et al., (2015). They inject into a single fault zone, but with the embedded stress heterogeneity and rate-state evolution, spatio-temporal clustering naturally occurs within the pressurized fault zone. When considering individual clusters from the same pressurized region, which is essentially the individual cluster analysis presented here, one would not expect statistically significant observations between the activity within each cluster and the total injection volume for the whole fault. On the other hand, the large-scale analysis combines both injection volumes and seismic activities over multiple clusters within the same spatial window, and the minimum required spatial window to observe clear relationship would be analogous to the dimension of the pressurized zone to some degree.

Both the gridded and k-means methods yield similar results: a minimum spatial window of about 25 km to start to see a R^2 greater than 0.5 and a slope to start to stabilize; and a 43 km radial window and 86 km square window for the methods to reach the maximum R^2. However, both methods have obvious drawbacks: the gridded method has the potential to split apart regions of potentially related disposal wells and seismicity; while the k-means method handles this association problem better, but events can be counted multiple times, so it can’t be directly compared to the theoretical equations.

The observation of large association distance between disposal volume and events is consistent with the pressure modeling of the Jones Swarm, where the majority of events where between 15 km and 30 km of the high-rate disposal wells that likely triggered the seismicity and the potential far-field triggering of the 2015 Fairview seismicity from 12 to 20 km away (Keranen et al., 2014; Yeck et al., 2016). As a conflicting example, Keranen et al. (2013) attributed the Prague sequence to a pair of wells within 5km of the sequences initiation point. However as pointed out by McGarr (2014), this would require very small volume wells to have triggered the M5 foreshock believed to initiate the Prague sequence (Sumy et al., 2014), and the data point would be substantially above the theoretical lines of both the moment cap and sample size hypotheses. Accordingly, it seems possible the Prague sequence was triggered by numerous wells further afield in the region.

That the M5.7 Prague event was likely triggered by Coulomb stress changes generated by the M5 foreshock has interesting implications to the 2016 M5.8 Pawnee event. The Pawnee event is similarly an outlier in the statewide relationship between V versus M_0max and V_{M_0max} versus M_0max (Figure 17). For both events the amount of fluid injected generated an unexpectedly large event relative to the rest of the state, although still under the theoretical line of the moment cap model. Furthermore, the Pawnee event had multiple M3+ foreshocks prior to the mainshock that likely played a role in the triggering of the mainshock (Pennington and Chen, 2017). It seems possible the Pawnee event was also triggered by Coulomb stress changes, which would explain its high moment relative to the regional volume.

Empirical relationship and theoretical relationship

To facilitate the discussion, theoretical relationships between the slopes for different datasets were derived (details in Appendix A). In Dieterich et al. (2015), similar to McGarr (2014), they suggest that the slope between V and M_0max (S_{M_0max}) should be 1 (i.e. would match equation (3) if b-value = 1.5), based on the logic that the earthquake rupture is limited to the pressurized volume when the initial stress is subcritical. However, they also acknowledge that if the pressure is dominated by fracture permeability along the faults, the same argument would predict S_{M_0max} =1.5, which would match equation (3) if b-value = 1.

Among all the studies focused on the maximum seismic moment, only McGarr (1976, 2014) derived relationship between volume and total seismic moment. If the earthquake magnitude-frequency distributions follow Gutenberg-Richter law, then there should be an identical slope between V and $\sum M_0$ (S_{M_0}) and S_{M_0max}, while the slope between V and N_{eq} (S_{Neq}) should be equal to $\frac{b}{1.5} * S_{M_0max}$. Next these theoretical relationships are compared with the observed data.

First, there is a notable discrepancy between the observed and theoretical lines of Figure 16. The observed data shows best-fitting S_{M_0max} of 0.83, which is shallower than both the 3D volume pressurization, and 2D fault pressurization would predict. To match the shallow slope, it would require a b-value of 1.8 for equation (3), which is on the higher end of the spatial varying b-value, so this is not unrealistic. However, if only data from 2005 to 2015 is used, the slopes are $S_{M_0} = 0.83 \pm 0.07$ and $S_{M_0max} = 0.68 \pm 0.07$, while including the 2016 data with all three M5 earthquakes in 2016, the slopes are $S_{M_0} = 0.9 \pm 0.08$, and $S_{M_{0max}} = 0.83 \pm 0.08$. The inclusion of 2016 data

results in both S_{M_0} and $S_{M_{0max}}$ shifting closer to the theoretical predictions, and closer to each other. A way to explain the discrepancy is that all the theoretical relationships, and observations in Dieterich et al. (2015), are based on earthquake sequences and injections that have been somewhat completed and the largest earthquakes have occurred. However, the Oklahoma activities are still ongoing. So the results in 2015 could be incomplete due to the delayed response to injection, which lead to a slightly better match with theoretical scaling by including the 2016 data.

Second, for the 2016 data, the best-fitting observed relationship is almost parallel to the prediction in McGarr (2014), but systematically shifted lower. The McGarr (2014) predicts scale factor between V and $\sum M_0$ of 2G (G is the shear modulus, typically assumed as 30 GPa), and McGarr (1976) predicts a scale factor of G. The difference comes from the assumptions whether the injection changes the pressure (McGarr 2014) or modifies the deviatoric stress field (McGarr, 1976). Either way, the scale factor is about an order of magnitude larger than the scale factor observed here: $10^{9.64}$ (~4.4 GPa). Given the uncertainties in the measurement of the y-intercept, the relationship in McGarr (1976) would be closer to the upper bound of the observation here (Figure 17b), which could suggest that some degree of elastic stress responses on the state level should be considered. There are several possible reasons for this: (1) The shear modulus where the earthquakes occur in Oklahoma is lower – this could be true given the overall shallow depth focus compared to other intraplate earthquakes. Given the uncertainties in the measurement, the scale factor can vary between 1.3 GPa and 15.3 GPa, and the higher-end of the scale factor is close to the typical shear modulus. This possibility could be validated with detailed 3D velocity model and density

structure. (2) Not all wells or not all the volume injected contributed to the induced seismicity. That not all wells contribute would be possible given that only a small percentage of hydraulic fracturing wells are associated with seismicity (Ghofrani and Atkinson, 2016), and that there is a lack of relationship at individual well scale. However, this would require detailed efforts to find out wells that probably do not contribute to seismicity, which is difficult to validate in Oklahoma since high-rate wells are often clustered. The fact that the upper end of permeabilities in the Arbuckle Group, with a 75^{th} percentile and maximum horizontal permeability measurements of 18.83 mD and 3088.61 mD for the upper Arbuckle Group and 55.86 and 171.62 mD for the lower Arbuckle Group (Morgan and Murray, 2015), range above what is considered seismogenic permeabilities, 0.5 mD to 50 mD (Talwani et al., 2007), suggest a portion of disposed fluid is propagating through more permeable formations without causing seismogenic pore pressure changes. (3) The Oklahoma earthquakes are not over yet, and the total seismic moment would continue to increase – this is very likely. There are volume reductions requested, and the injection volume to Arbuckle Group has been significantly reduced (Langenbruch and Zoback, 2016). However, as observed in Dieterich et al., (2015), continuing seismicity can occur after the shut-in (or reduction) of injection activities. In addition, there is often time-delay associated with induced seismicity due to both the diffusion process and delayed nucleation of earthquakes (Segall and Lu, 2015). For example, combining poroelastic modeling and rate-and-state friction, Barbour et al., (2017) demonstrated a good match between predicted seismicity rate and observed seismicity rate. This possibility can be validated, but requires continuing seismic monitoring for much longer time period.

The third observation, the M5 earthquakes (M5.7 Prague, M5.8 Pawnee, M5.1 Fairview) appear as outliers on the V_{M_0max} versus M_0max plot, but there are no clear outliers for the V versus $\sum M_0$ and V versus N_{eq} plots (Figure 16 and 17). The slope S_{M0} is nearly sub-parallel to the theoretical predicted slope of 1 (McGarr, 2014). However, the S_{Neq} is much lower than the seismogenic index model used in Van der Elst (2015), which predicts $S_{Neq} = 1$, based on the assumption of homogeneous medium with uniformly distributed cracks (Shapiro et al., 2010). Because of the inherit scaling between magnitude and seismic moment, $S_{Neq} = \frac{b}{1.5} S_{Mo}$, they cannot both equal to 1, unless b=1.5. The shallower slope for S_{Neq} could be due to that the magnitude completeness is actually much higher than 2.5, which could result in the systematic shift observed here. The magnitude of completeness is less an issue for the total seismic moment, as this will be dominated by larger earthquakes. However, this does not affect the key observation here, which is both N_{eq} and $\sum M_0$ are closely related to volume, including cells with M5 earthquakes.

The overall agreement between S_{M0} and the theoretical prediction suggesting consistency with the moment-cap theory. The maximum seismic moment relationship in McGarr (2014) was derived based on homogeneous Gutenberg-Richter distribution, which contradicts with the observation here (see the strong spatial variability of b-value in Figure 18 and 19). Taking this into account, the sample-size hypothesis proposed by van der Elst et al. (2016) can plausibly be incorporated. With the N_{eq} and $\sum M_0$ constrained, the re-distribution of seismic moment onto individual earthquakes would depend on the local magnitude distributions. The M5.7 Prague and M5.8 Pawnee earthquakes are the two most robust outliers in the maximum seismic moment plots (the

M5 Fairview earthquake merges with the other cells with 1.0 degree bins), and both have lower b-value. Following equation (3), the effect of spatial variability is tested by plotting the 'local slope' that should be present for an area based on the b-value of the maximum moment event in each cell (Figure 20). In Figure 20, there are apparently two populations of the slopes – steeper slopes with lower b-values for relatively larger earthquakes, and shallower slopes with higher b-values for relatively smaller earthquakes. Combining them, a shallow slope is obtained from the least squares fitting to all the data. The steeper slope is consistent with the scenario of pressurization of fault surface suggested by Dieterich et al., (2015), and also consistent with observations of low b-value during fault activation stage (e.g. Shelly et al., 2016).

Fourth, another notable result is the only minor differences between the R^2 and best-fit lines of V versus M_0max and V_{M_0max} versus M_0max (Figure 9, 16). The V_{M_0max} versus M_0max test is performed to see if size of the largest event is controlled by the amount of fluid injected before it occurs (moment cap hypothesis), while the V versus M_0max test does not consider the timing of the largest earthquake in each cell (sample-size hypothesis). If the sample size hypothesis is true, then there should be a weaker correlation between V_{M_0max} versus M_0max, which is not observed here (Figure 16 and 17). This could be affected by two reasons: First, the volume after maximum moment is being influenced by OCC regulations. Regional injection activity is often halted or reduced after a large magnitude event, and this is reflected in the data, at a 1 degree grid cell, on average only 31% of a region's volume is injected after peak seismicity (considering 2005–2015 data). As seen in Figure 21, the relative change in volume for each grid cell from V_{M_0max} to V is small, and V_{M_0max} and V are clearly

correlated. This suggests that in Oklahoma, it would be difficult to solely distinguish these two hypotheses based on the timing of largest earthquakes observed on individual clusters.

Finally, combining all the observations, it is proposed that the on average, the moment-cap theory seems to be consistent with the upper bound of the total seismic moment from injected volume, but the observations are systematically below the prediction. The sample size model would provide a good overall constraint for the number of earthquakes, but the relationship requires some modifications to match the observed slope. The maximum seismic moment would depend on local magnitude distribution, which would need to incorporate the sample-size theory, and fully consider the spatial heterogeneity in b-value, which likely reflects fault dimension (Shelly et al., 2015) or differential stress (Schorlemmer et al., 2005) distributions – further studies may be needed to distinguish the two. Some of the steeper slopes are consistent with the fault-facilitated permeability suggested by Dieterich et al., (2015).

As cautious notes, the dataset is subject to location uncertainties – there exist some relocated catalogs, but they either applies a lower magnitude cut-off to only retain best-recorded earthquakes, or discard earthquakes that are non-clustered (e.g, double-difference), so the original full catalog was used. However, this should not significantly affect the results, as the minimum spatial window of 25 km is much larger than the location uncertainty. The shallower slope of S_{Neq} could be due to systematically higher Mc than the 2.5 used here, or the delayed responses of seismicity, but could also due to a physical mechanism: if the seismic moment is more relevant to the physical processes

and stress tensors, and is more reliable, then the slope S_{Neq} would depends on the b-value, and is not necessarily equal to 1.

Conclusions

Oklahoma shows a clear relationship between injection volume and seismicity. The relationship is statistically significant on the scale of individual wells or individual clusters. This does not preclude cases where individual relationships can be established with local-scale analysis, but suggest that at the state level, groups of wells are often associated with groups of clusters. This contradicts with previous observations that individual high-rate disposal wells are more likely to be associated with seismicity, which is likely due to the fact that high-rate wells are preferentially grouped within high-volume zones.

On the state level, two different approaches both suggest that a minimum association distance over 20 km is required to define clear relationships between injection and seismicity, and a distance of 40 km to stabilize the relationship. These distances are larger than most documented cases of non-Oklahoma induced seismicity, which typically report induced triggering to a maximum distance of 10–15 km. These support the notion of far-field triggering from fluid injection, likely due to high permeability within the Arbuckle Group.

Empirical equations relating volume to earthquake numbers and seismic moment in Oklahoma are also derived. The maximum seismic moment shows overall consistency with moment cap hypothesis, but the M5 earthquakes show as outliers compared to the rest of the state. However, the total seismic moment show stronger correlation with total injection volume, including the M5 earthquakes, consistent with

the moment cap hypothesis. The scale factor is systematically lower than a shear modulus of 30 GPa, this could be due to lower shear modulus at hypocenters for Oklahoma earthquakes, which could be inferred from high-resolution 3D velocity models, or because the delayed response of seismicity, which requires long-term continuous monitoring. The delayed response is also evident in the higher slope when the 2016 is included (compared to only 2005–2015 data). When considering heterogeneity in b-value distributions, the sample-size hypothesis provides two groups of scaling between maximum seismic moment and total injection volume, suggesting the necessity in incorporating spatial variability. The clear demonstration of a statewide correlation between seismicity and fluid injection into the Arbuckle, and a nearly 1:1 relation for volume versus moment (with scaling factor of a few GPa), is a compelling argument for regulators reducing not just rate, but the total volume into the Arbuckle Group. Reducing injection rates to pre-2014 levels may reduce the seismicity rate, but may not prevent continued seismicity due to the delayed pressure diffusion and nucleation of earthquakes.

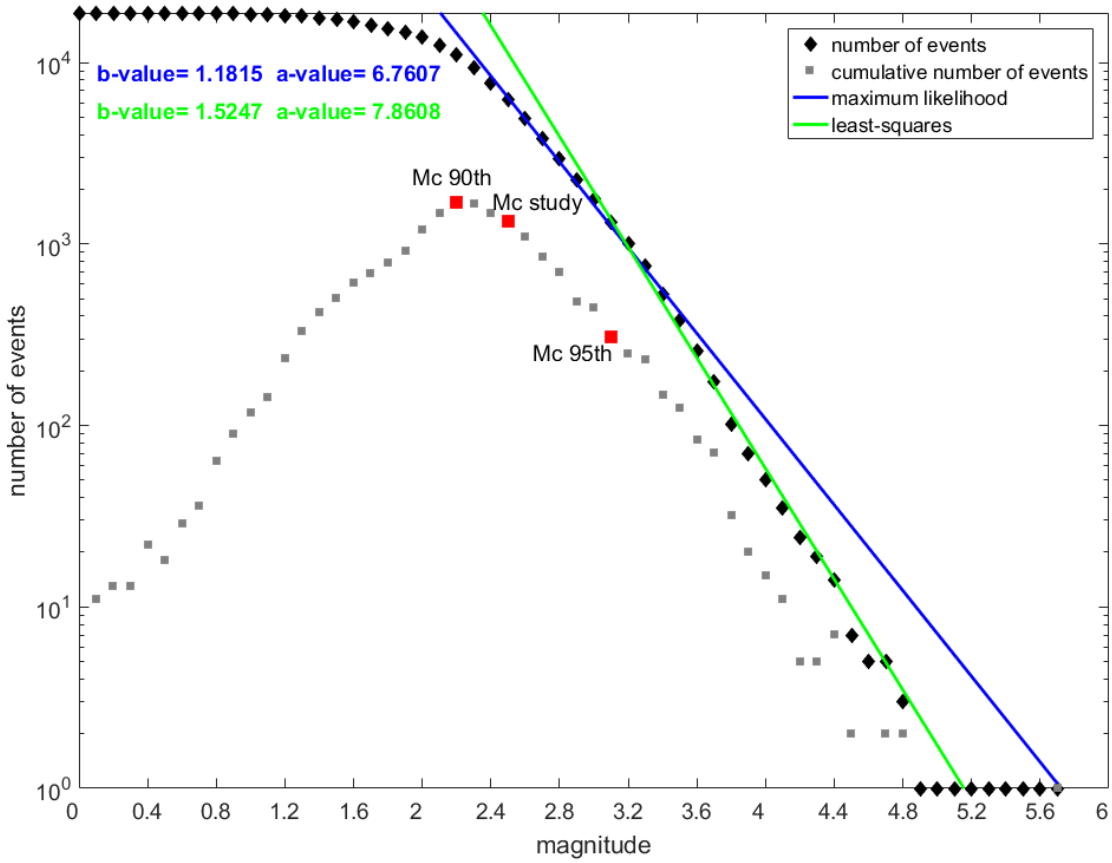

Figure 2. Cumulative frequency versus magnitude plot of Oklahoma seismicity between 2005-2016. The magnitude of completeness markers of Mc 90th and Mc 95th represent 90% and 95% of the data can be modeled by a power law fit (Wiemer and Wyss, 2000). The Mc used in this study, Mc = 2.5, represents a choice between the two estimates and two magnitude bins above the max curvature of 2.3. Also depicted is two b-value fits to the data, the maximum likelihood method which fits the small magnitude data well, and a bootstrapped least-squares fitting method which fits the high magnitude distribution wells except for the M5.7 Prague earthquake.

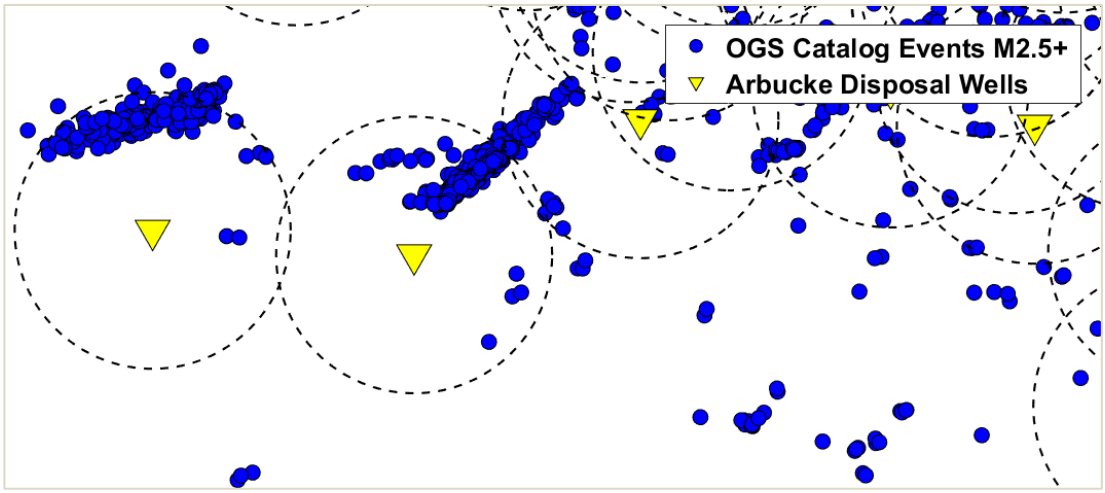

Figure 3. Example of the circular grouping of events within 10 km of an Arbuckle disposal well.

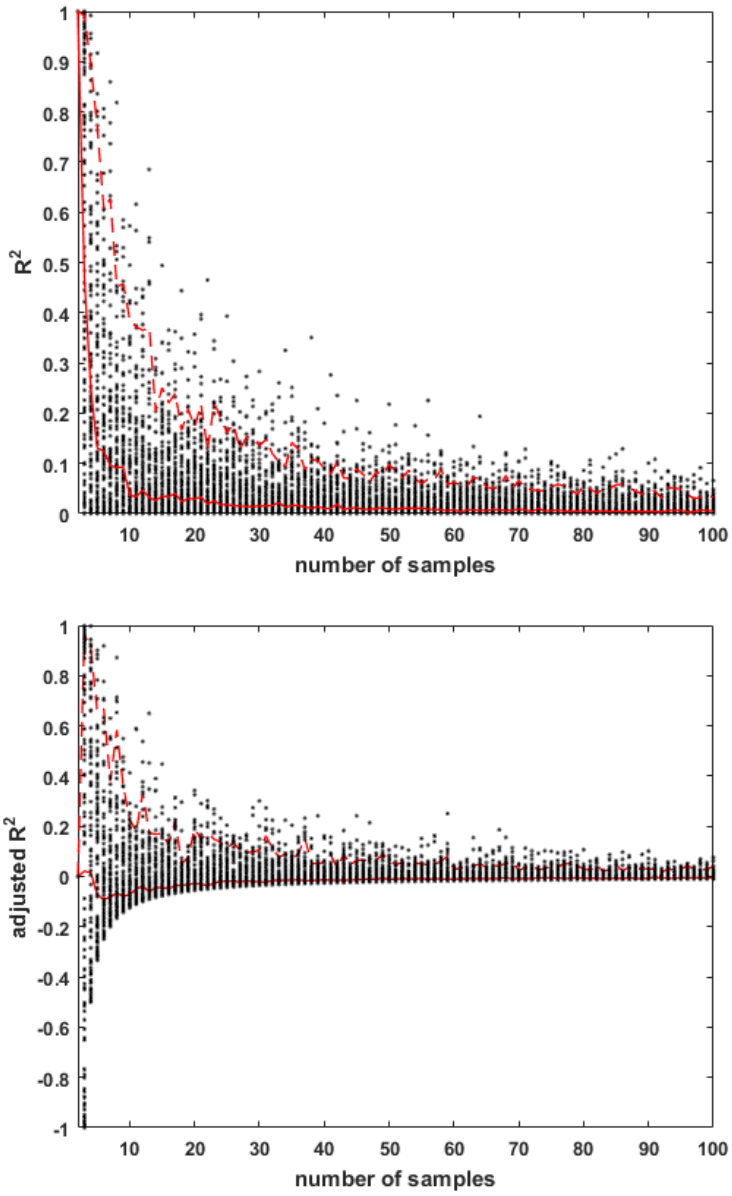

Figure 4. Visualization of R^2 (top) and adjusted R^2 (bottom) for small sample counts. Both plots depict the R^2 fit of a linear regression for 100 iterations of randomly distributing n = (1:100) number of samples between x = (0,1) and y = (0,1). The solid red and dashed red lines depict the median and 95th percentile of the 100 iterations for each sample count n. Both plots show a small spread and low R^2 for large sample counts. As the number of samples gets closer to the number of variables (in this case 2 variables, slope and y-intercept of a linear regression) the median R2 increases despite no trend in the data. The adjusted R^2 still has an increase in the 95th percentile with small sample counts but maintains the median R^2 at around zero.

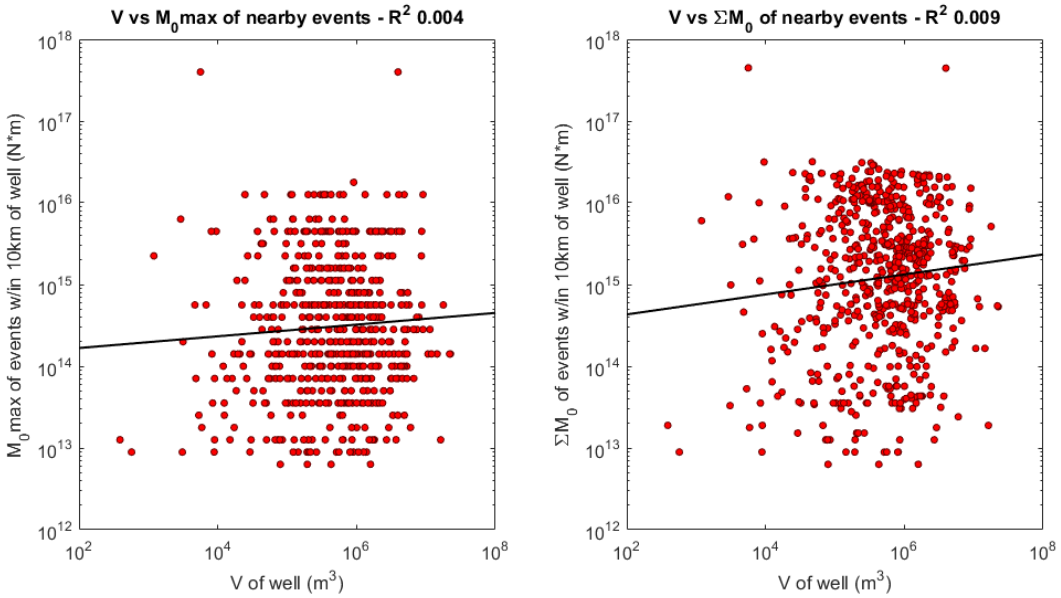

Figure 5. The total volume disposed of a given well against the maximum moment of events within 10 km of each well (left) and against the total moment of events within 10 km of each well (right). The black line on each plot depicts a best-fit linear regression to the log10 of both variables. Both cross-plots demonstrate no trend in the data. Also noticeable is the high moment release near the two wells within 10 km of the M5.7 Prague event.

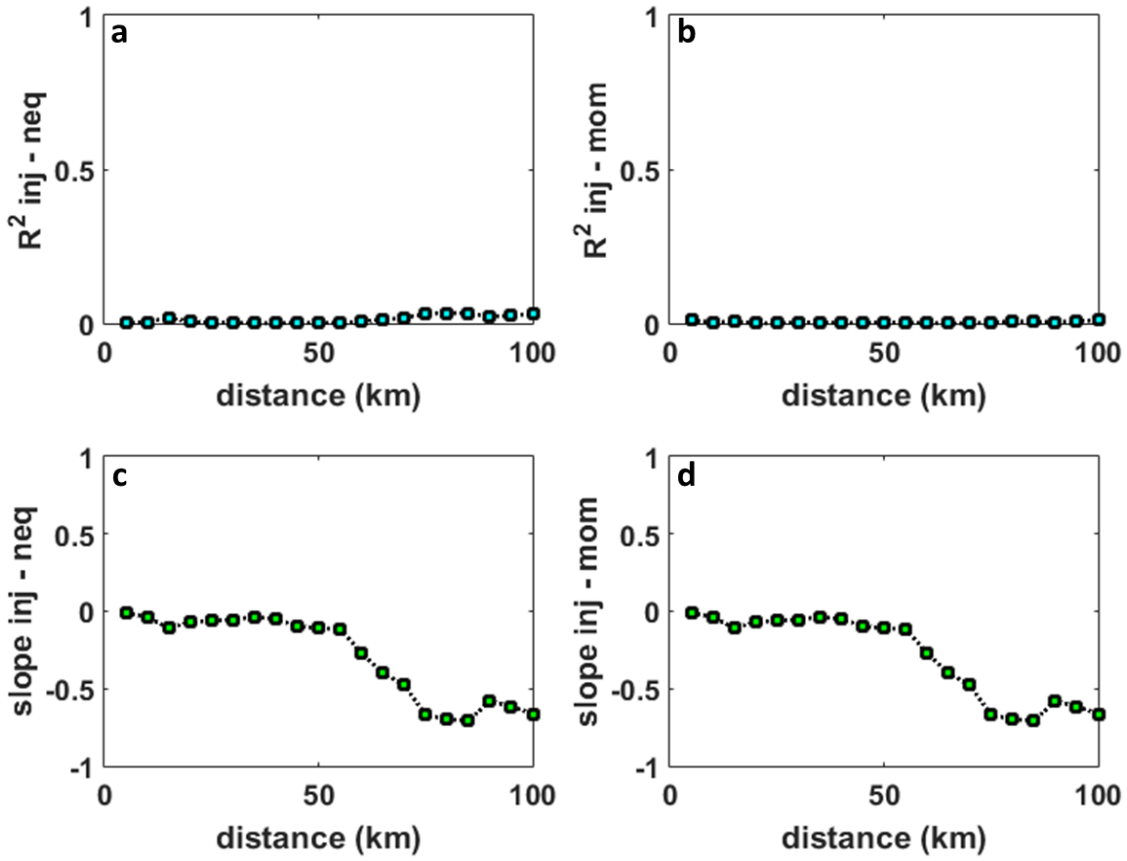

Figure 6. Cluster based spatial relationship testing between volume and seismicity. The variation in R^2 for the linear regression fit to $\boldsymbol{V}$ versus $\boldsymbol{N_{eq}}$ (**a**), $\boldsymbol{V}$ versus $\sum \boldsymbol{M_0}$ (**b**) of clusters at various distances (top). The variation in slope for $\boldsymbol{V}$ versus $\boldsymbol{N_{eq}}$ (**c**), $\boldsymbol{V}$ versus $\sum \boldsymbol{M_0}$ (**d**) of clusters at various distances (bottom). The results show there is little trend in the data, with extremely small R^2 and no positive relationship in the slope.

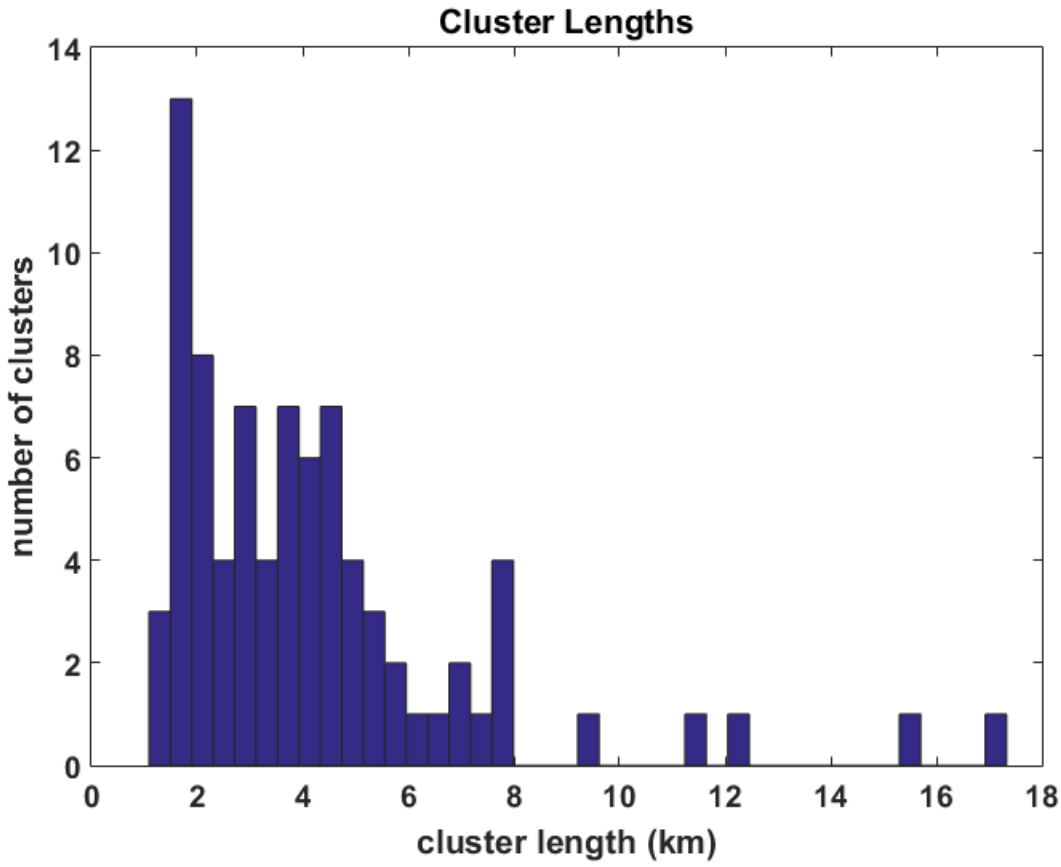

Figure 7. Approximate length of the 82 seismic clusters. The 0.1 degree grid size (approximately 14 km) is the smallest size tested to still capture the entire length of a fault.

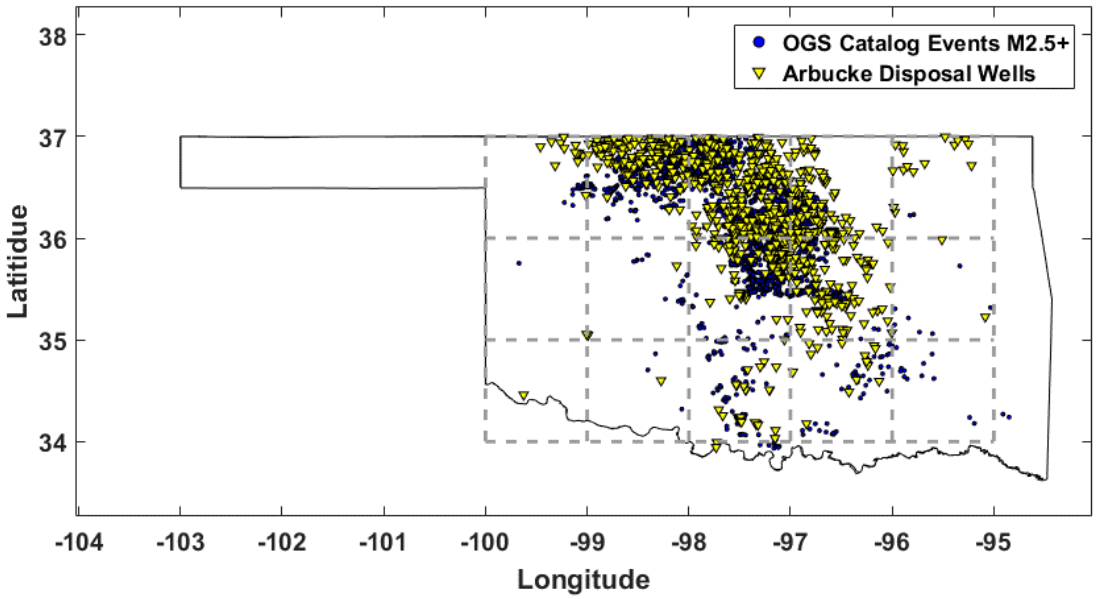

Figure 8. Example of the spatial summing of seismic and well data into a 1.0 degree grid.

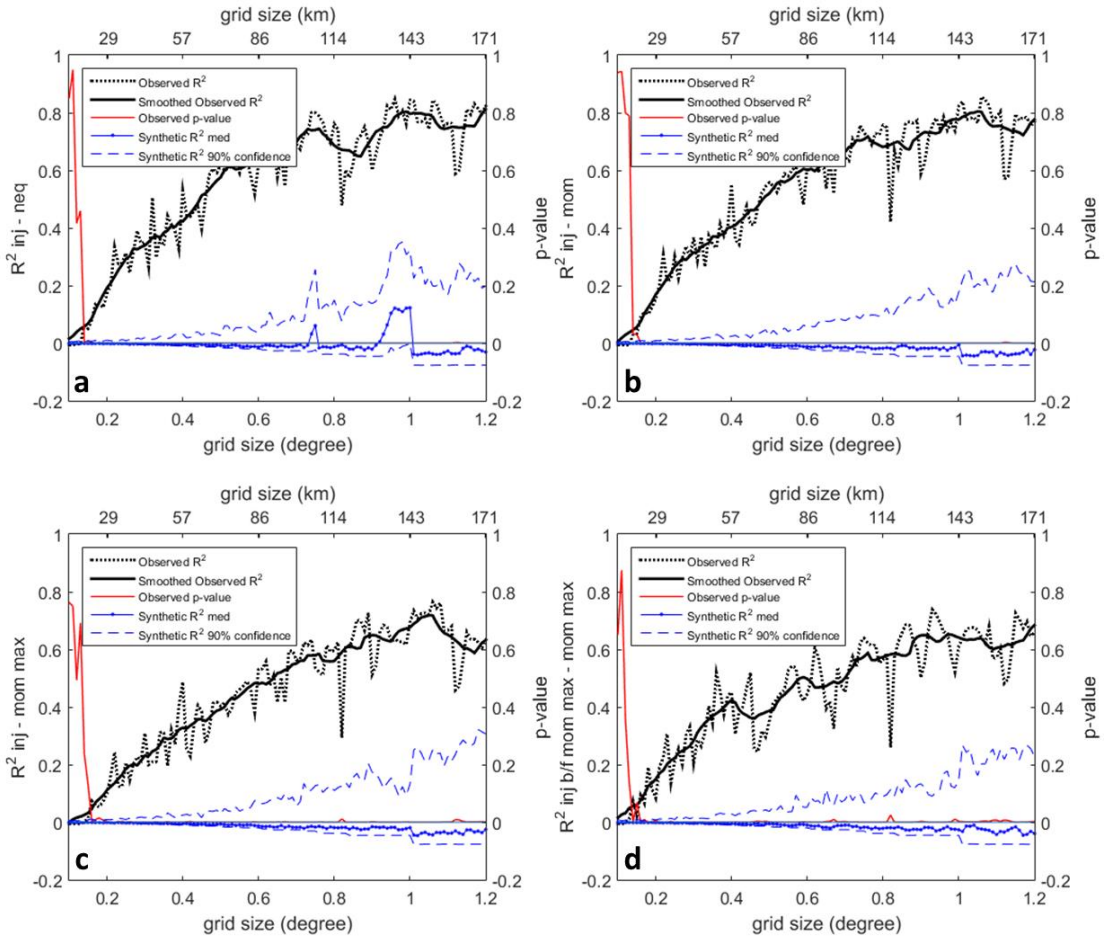

Figure 9. Coefficient of determination (R^2) for observed and synthetic data, as well as the p-value for the observed data, for grid sizes from 0.1 to 1.2 degrees for the datasets $\boldsymbol{V}$ versus $\boldsymbol{N_{eq}}$ (**a**), $\boldsymbol{V}$ versus $\boldsymbol{\sum M_0}$ (**b**), $\boldsymbol{V}$ versus $\boldsymbol{M_0 max}$ (**c**), and $\boldsymbol{V_{M_0 max}}$ versus $\boldsymbol{M_0 max}$ (**d**). All datasets show increasing coefficient of determination with increasing bin size and low p-values above 0.2 degrees. The synthetic dataset shows a low R^2 regardless of grid size.

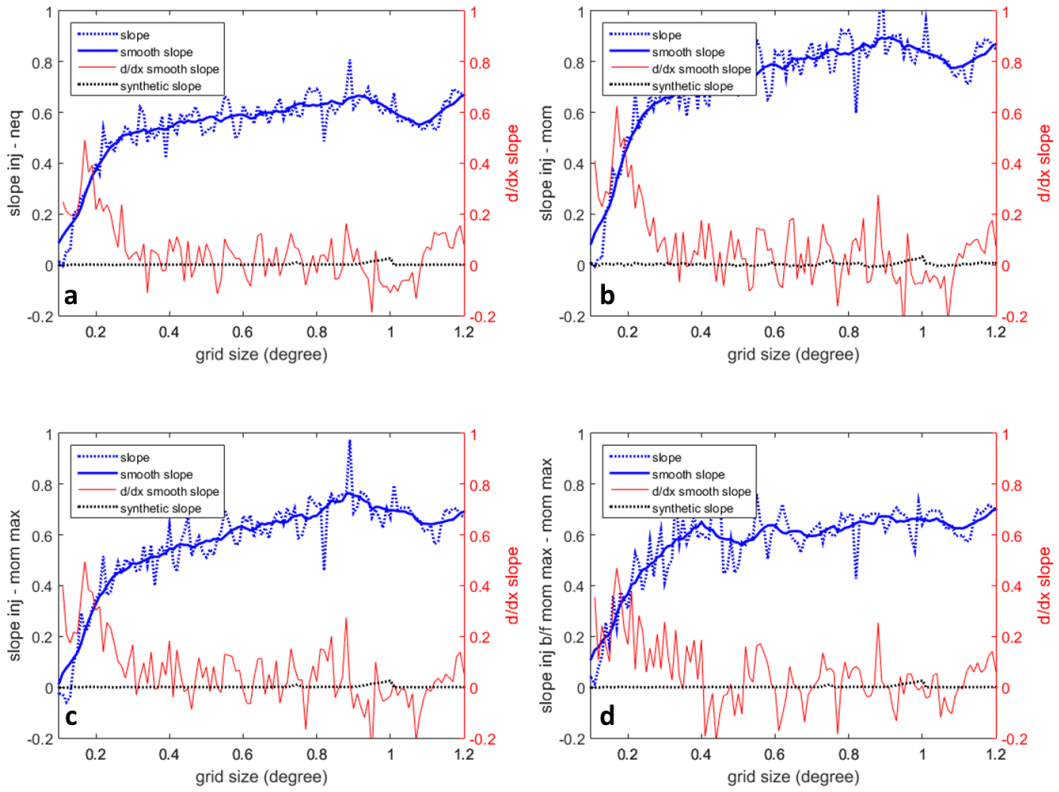

Figure 10. The slope of a linear regression fit to $\boldsymbol{V}$ versus $\boldsymbol{N_{eq}}$ (**a**), $\boldsymbol{V}$ versus $\boldsymbol{\sum M_0}$ (**b**), $\boldsymbol{V}$ versus $\boldsymbol{M_0 max}$ (**c**), and $\boldsymbol{V_{M_0 max}}$ versus $\boldsymbol{M_0 max}$ (**d**) at various grid sizes. Also represented is the slope of the synthetic test and the rate of change of the smoothed slope of the observed data. All datasets show the slopes increasing from small bin sizes until around 0.3 degrees, when the rate of change becomes close to zero. From 0.3 degrees the slope gradually increases until around 0.6 degrees, after which the slope has only minor fluctuations.

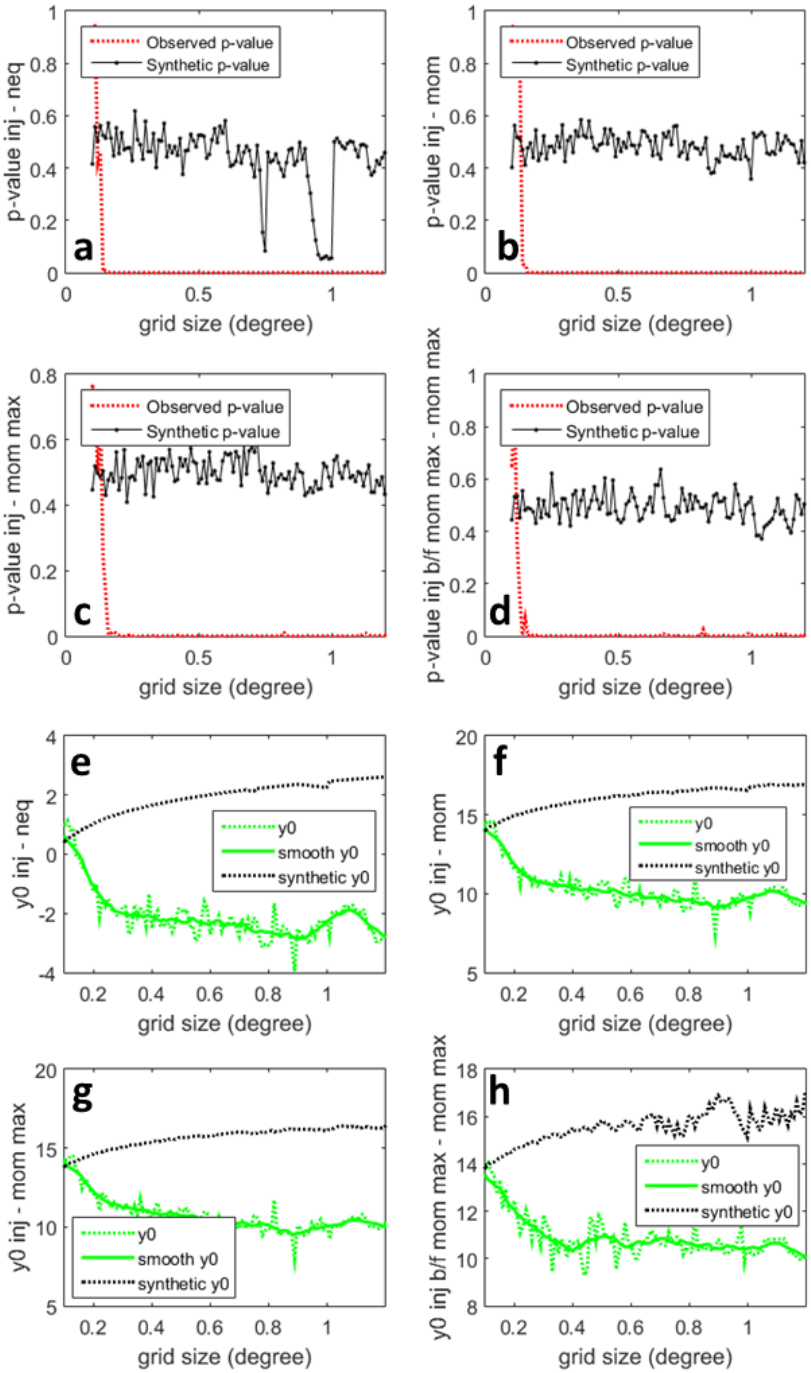

Figure 11. Variation in p-value (left) and y-intercept (right) for the linear regression fit to $\boldsymbol{V}$ versus $\boldsymbol{N_{eq}}$ (**a**,**e**), $\boldsymbol{V}$ versus $\sum \boldsymbol{M_0}$ (**b**,**f**), $\boldsymbol{V}$ versus $\boldsymbol{M_0 max}$ (**c**,**g**), $\boldsymbol{V_{M_0 max}}$ versus $\boldsymbol{M_0 max}$ (**d**,**h**) at various bin sizes. All observed datasets show a low p-value (below 0.05) at around 0.19 grid size, while synthetic datasets show a high p-value for all grid sizes. Plots (**e**,**f**,**g**) show a sharp decline in y0 until around 0.3 grid size, after which a gradual decline until a stable value is found at around 0.6 degrees.

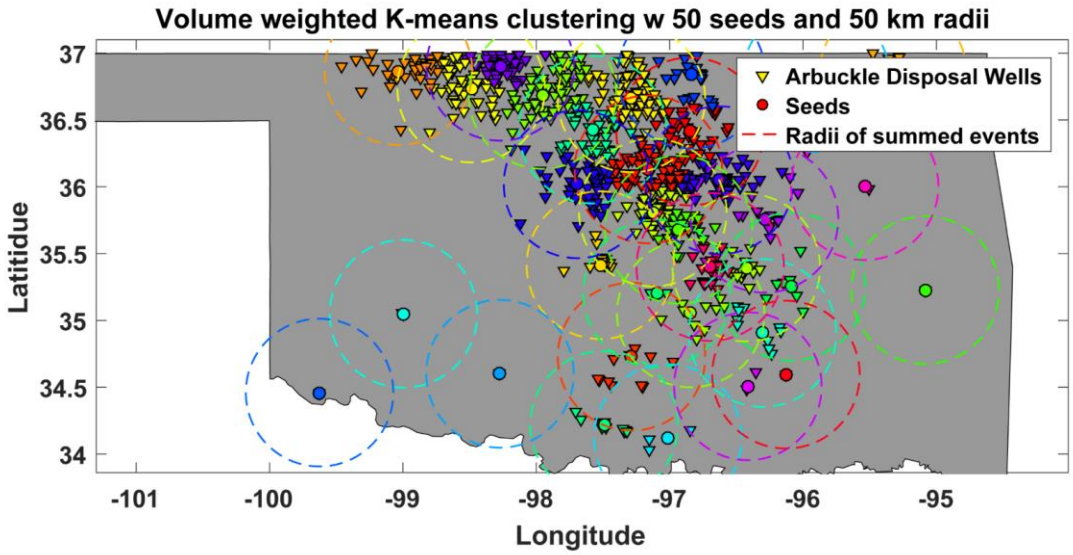

Figure 12. Example of the k-means method showing Arbuckle disposal wells in the same color as the seed they are associated with after convergence has been reached in the relocation of the seeds. Also depicted is the circular bin of 50 km that events will be tallied and cross-plotted against the seed's associated wells' volumes.

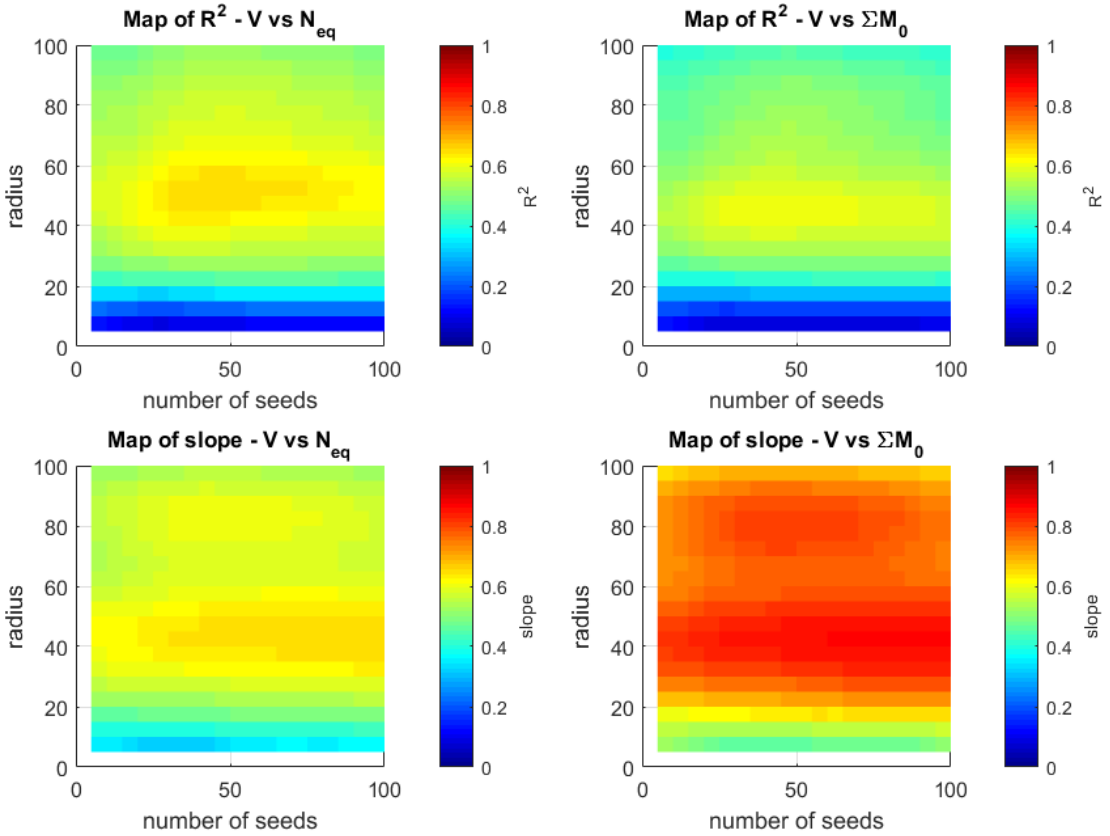

Figure 13. 2D map results of linear regression fit of k-means method. (Top) Median R^2 for 100 runs for each combination of initial seed count and bin radius for $\boldsymbol{V}$ versus $\boldsymbol{N_{eq}}$ (left) and $\boldsymbol{V}$ versus $\boldsymbol{\sum M_0}$ (right). Median slope of best-fit line for 100 runs for each seed and radius combination for $\boldsymbol{V}$ versus $\boldsymbol{N_{eq}}$ (left) and $\boldsymbol{V}$ versus $\boldsymbol{\sum M_0}$ (right). All four cross-plots show a local max between 35 km and 55 km and require a minimum distance of about 25 km to show consistent results.

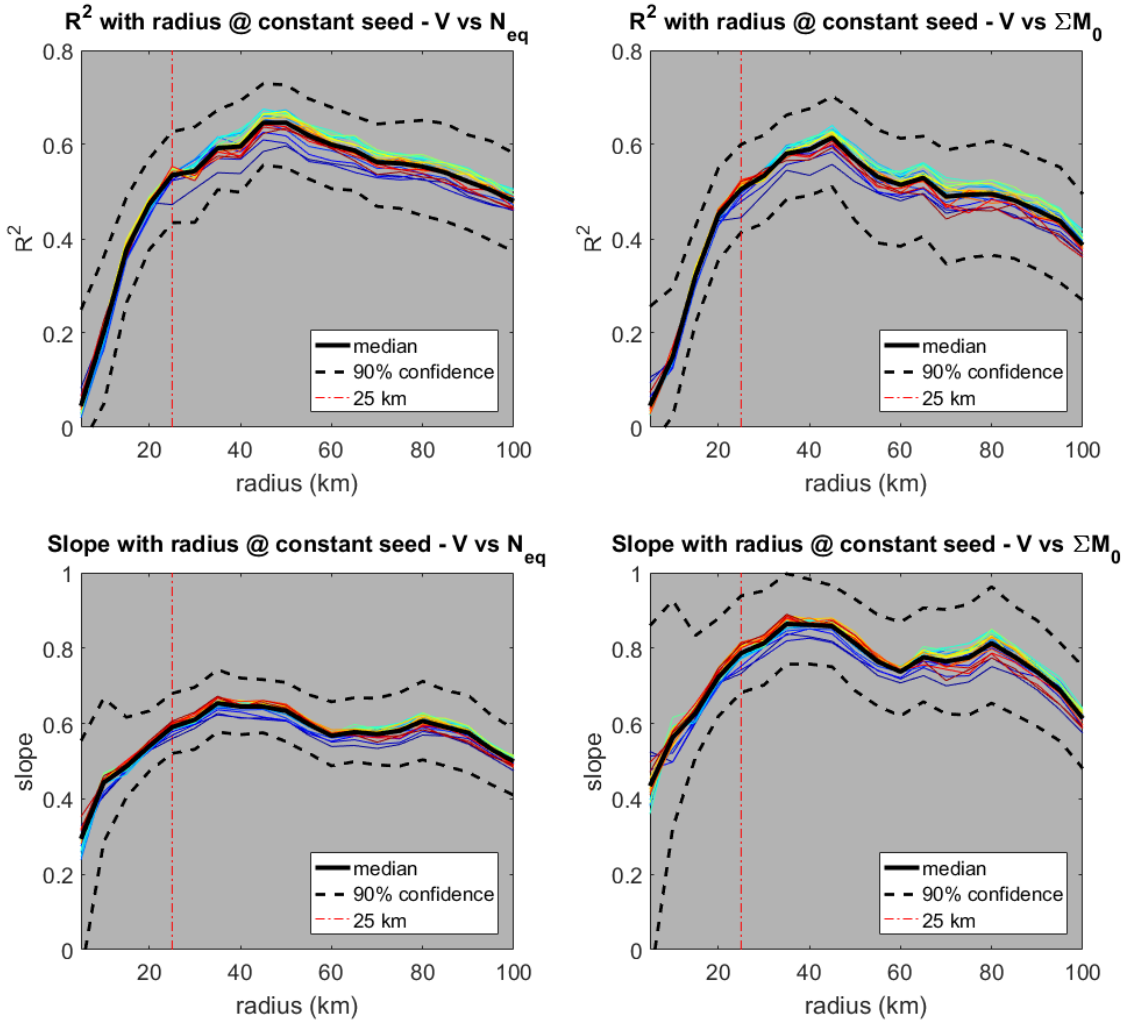

Figure 14. For the 100 iterations of all initial seed numbers, colored lines represent the median R^2 (top) and slope (bottom) of a linear regression fit to the data for each initial seed count, as well as the average of the median and average of the 90% confidence intervals between initial seed counts for $\boldsymbol{V}$ versus $\boldsymbol{N_{eq}}$ (**a**,**c**) and $\boldsymbol{V}$ versus $\boldsymbol{\sum M_0}$ (**b**,**d**). All plots show a sharp increase until about 25 km where the R^2 and slope gradually increase to a local max between 40 – 45 km.

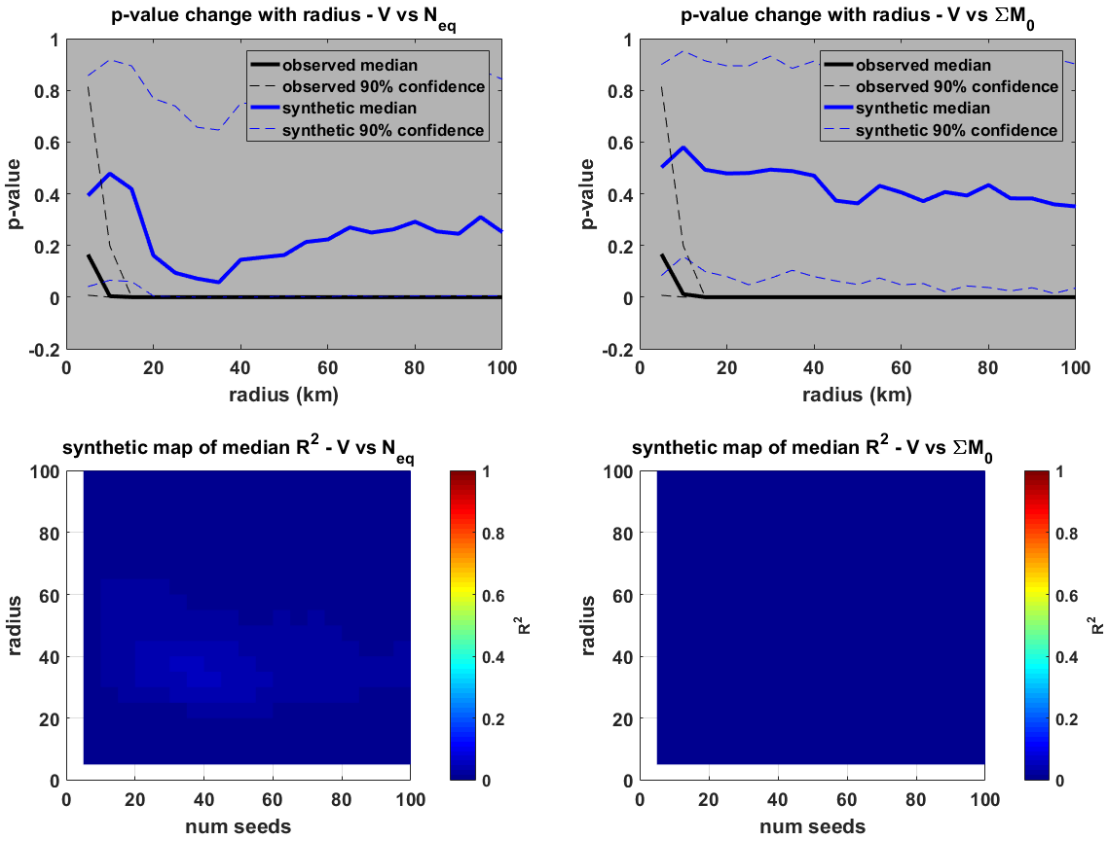

Figure 15. (Top) P-value variation with bin radius for observed and synthetic data for $\boldsymbol{V}$ versus $\boldsymbol{N_{eq}}$(left) and $\boldsymbol{V}$ versus $\sum \boldsymbol{M_0}$ (right). (Bottom) 2D map of median coefficient of determination for 100 synthetic well and earthquake datasets for $\boldsymbol{V}$ versus $\boldsymbol{N_{eq}}$(left) and $\boldsymbol{V}$ versus $\sum \boldsymbol{M_0}$ (right).

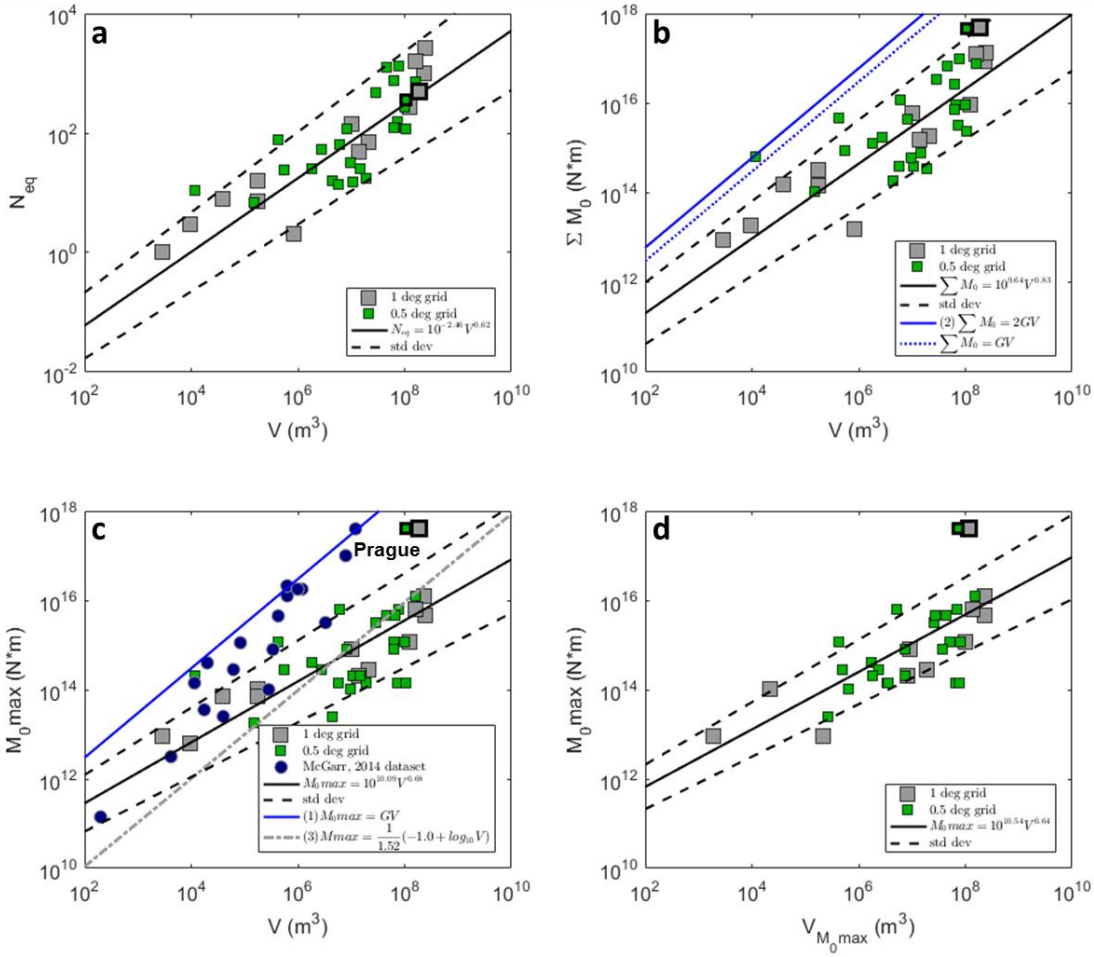

Figure 16. Plots of $\boldsymbol{V}$ versus $\boldsymbol{N_{eq}}$ (**a**), $\boldsymbol{V}$ versus $\boldsymbol{\sum M_0}$ (**b**), $\boldsymbol{V}$ versus $\boldsymbol{M_0 max}$ (**c**), and $\boldsymbol{V_{M_0 max}}$ versus $\boldsymbol{M_0 max}$ (**d**). All cross-plots have the data points of 1 and 0.5 degree bins and the average best-fit lines. $\boldsymbol{V}$ versus $\boldsymbol{\sum M_0}$ (**b**) has the theoretical line of equation (2). $\boldsymbol{V}$ versus $\boldsymbol{M_0 max}$ (**c**) has the data points of McGarr (2014) as well as the theoretical lines of equation (1) and (3). Grid cells with events M5+ are outline in black. The pronounced outlier in (**b**), (**c**) and, (**d**) is the 2011 M5.7 Prague event.

Table 1. The mean and standard deviation for the slope and y-intercept of bins above 0.6 degrees both excluding and including 2016 data.

	2005–2016	**2005–2017**
S_{Neq}	0.617 ± 0.055	0.634 ± 0.054
S_{M0}	0.834 ± 0.072	0.909 ± 0.078
S_{M0max}	0.682 ± 0.071	0.827 ± 0.076
S_{M0max_volume}	0.643 ± 0.056	0.81 ± 0.068
$y0_{Neq}$	-2.456 ± 0.440	-2.487 ± 0.438
$y0_{M0}$	9.644 ± 0.542	9.306 ± 0.593
$y0_{M0max}$	10.091 ± 0.493	9.346 ± 0.549
$y0_{M0max\_volume}$	10.537 ± 0.387	9.664 ± 0.490

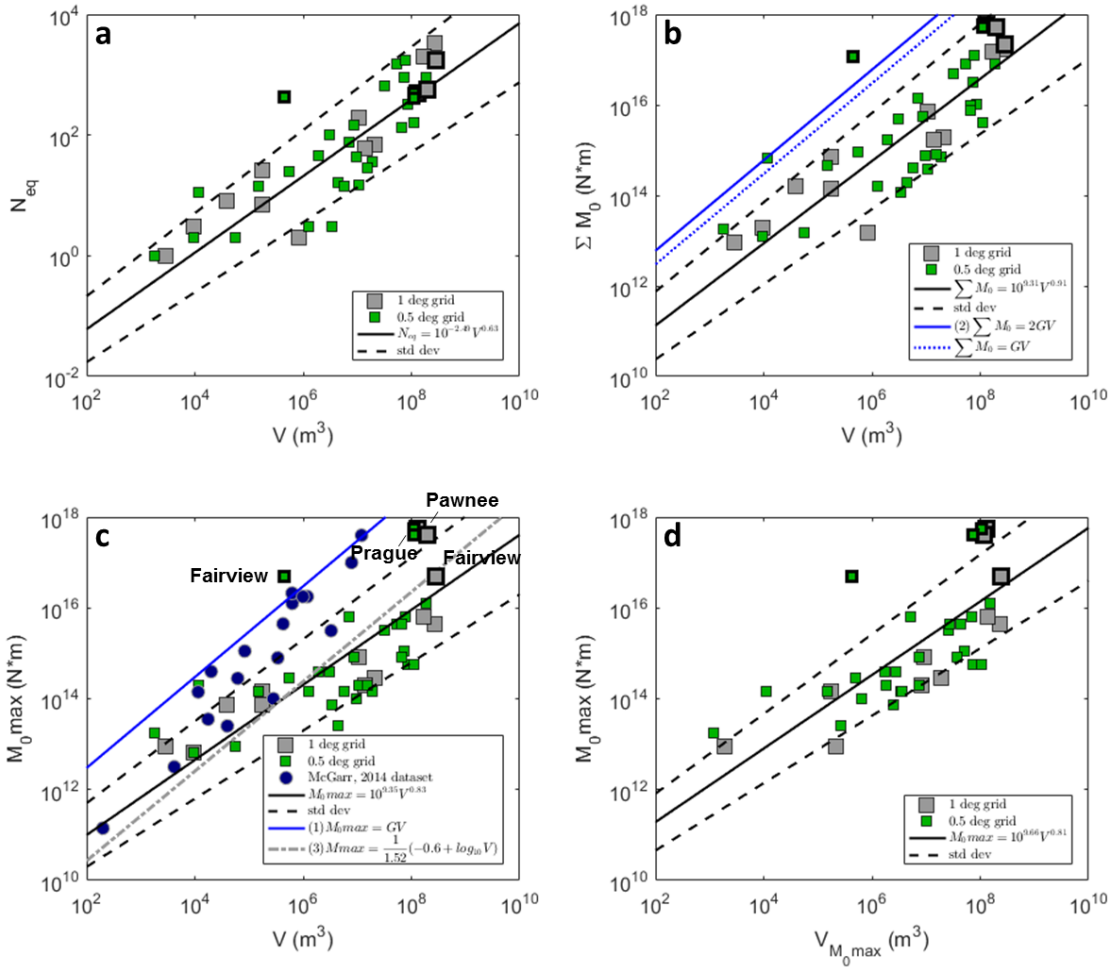

Figure 17. . Plots of $\boldsymbol{V}$ versus $\boldsymbol{N_{eq}}$ (**a**), $\boldsymbol{V}$ versus $\boldsymbol{\sum M_0}$ (**b**), $\boldsymbol{V}$ versus $\boldsymbol{M_0 max}$ (**c**), and $\boldsymbol{V_{M_0 max}}$ versus $\boldsymbol{M_0 max}$ (**d**) including 2016 data. All cross-plots have the data points of 1 and 0.5 degree bins and average best-fit lines. $\boldsymbol{V}$ versus $\boldsymbol{\sum M_0}$ (**b**) has the theoretical line of equation (2). $\boldsymbol{V}$ versus $\boldsymbol{M_0 max}$ (**c**) has the data points of McGarr (2014) as well as the theoretical lines of equation (1) and (3). Grid cells with events M5+ are outline in black. The 2011 M5.7 Prague event is joined by the M5.8 Pawnee and M5.1 Fairview events as outliers in (**b**), (**c**) and, (**d**). The violation of equation (1) by the Fairview 0.5 grid size point is due to the small grid size not including the high-volume wells to the northeast of the Fairview cluster.

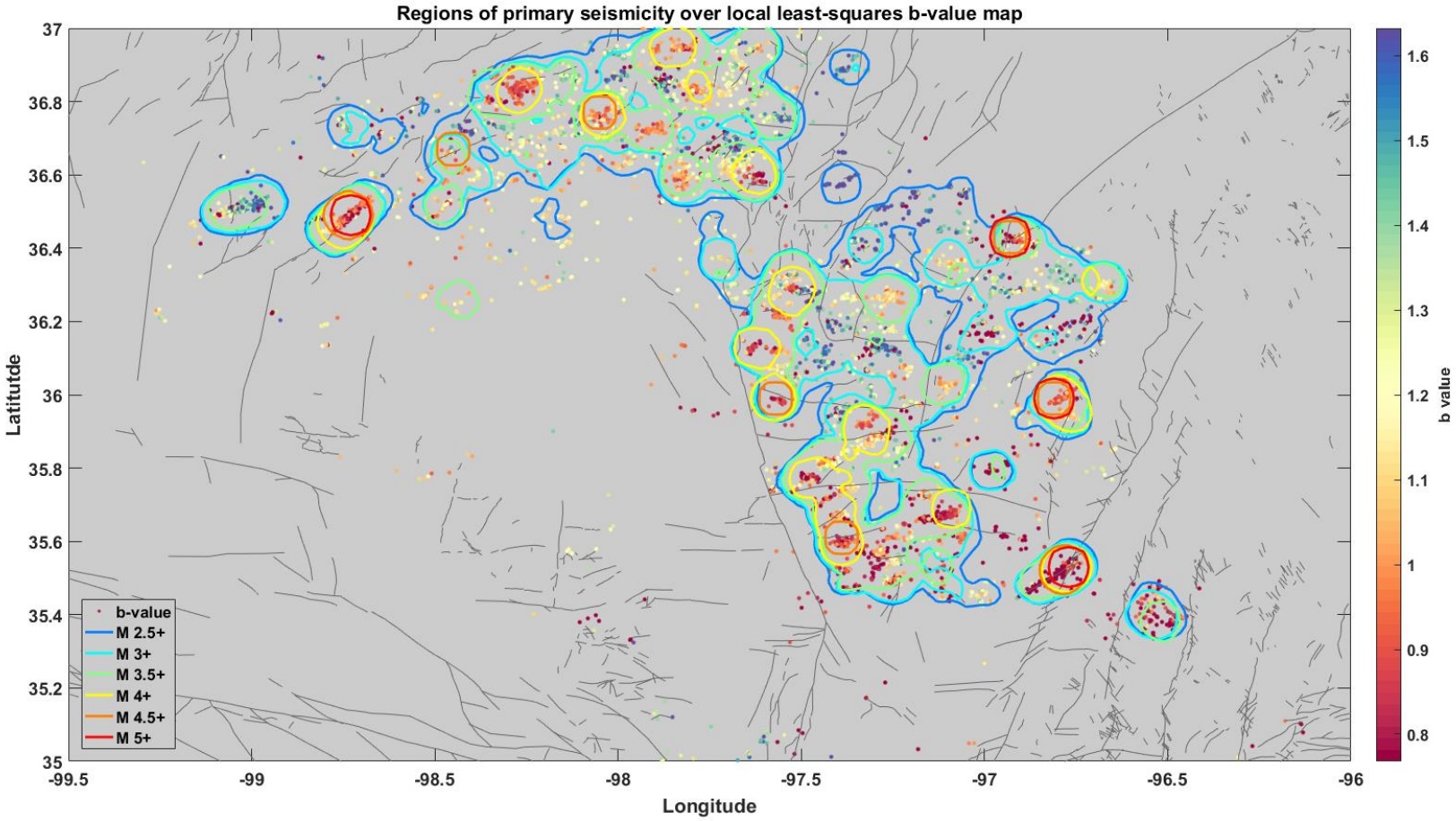

Figure 18. Map of Oklahoma local variations in least-squares b-value and earthquake magnitudes. The dots are earthquakes color coded by their local b-value estimated using a least-squares fitting method The contour lines represent where the density of events above a given magnitude are greater than the state average. There is a clear trend in lower b-values in areas with larger (M4+) events

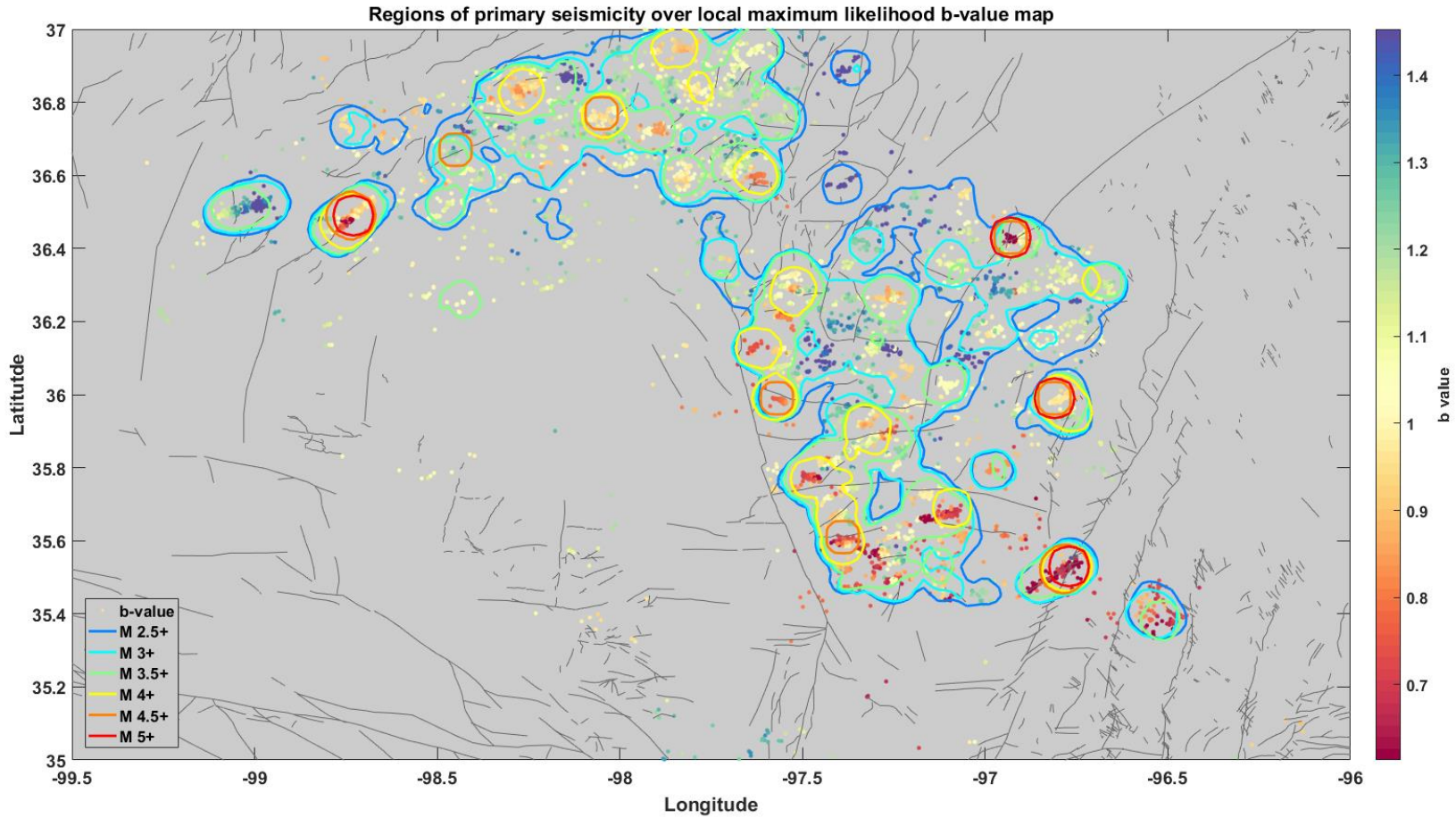

Figure 19. Map of Oklahoma local variations in maximum likelihood b-value and earthquake magnitudes. The dots are earthquakes color coded by their local b-value estimated using the maximum likelihood. method The contour lines represent where the density of events above a given magnitude are greater than the state average. There is a clear trend in lower b-values in areas with larger (M4+) events

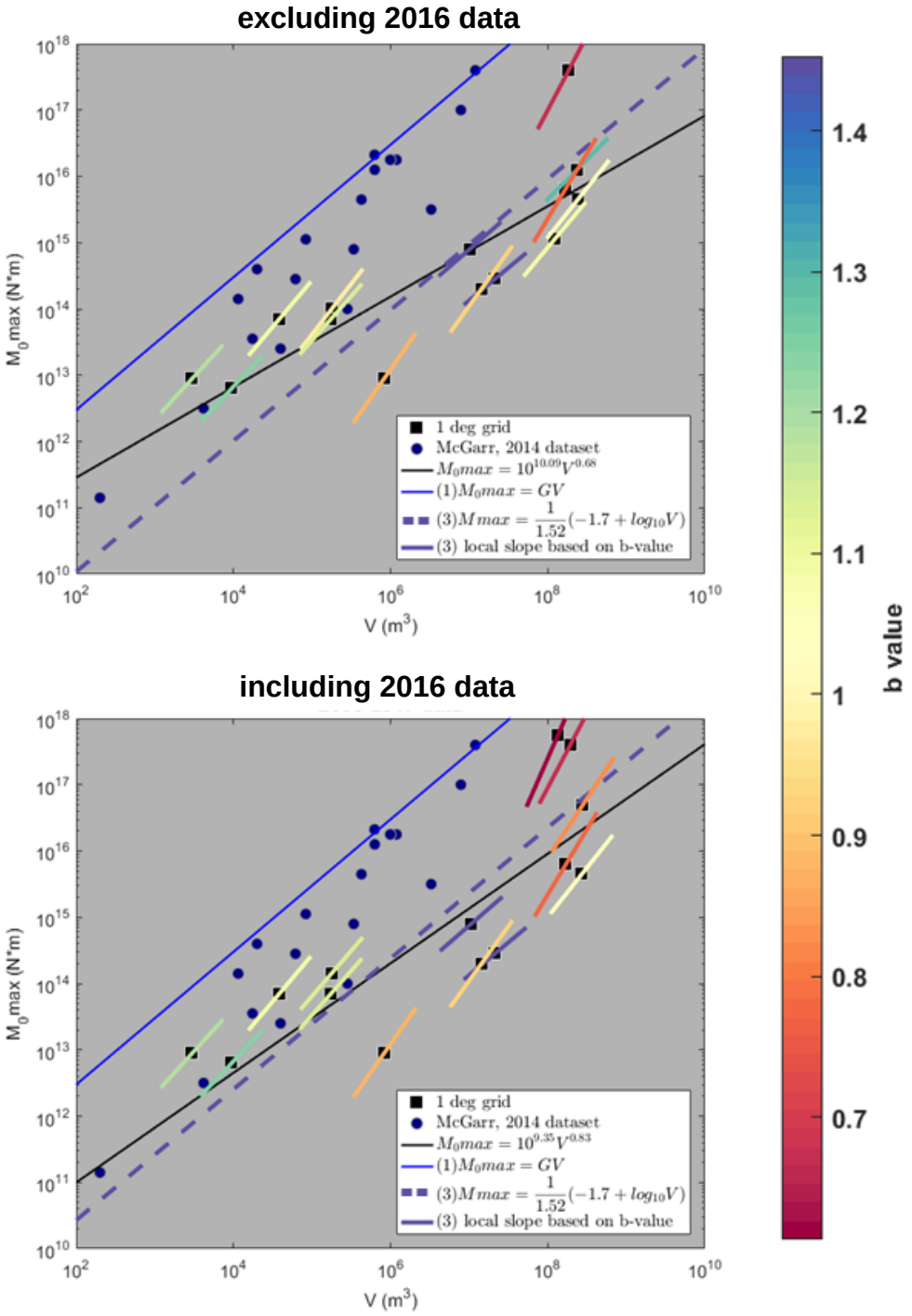

Figure 20. $\boldsymbol{V}$ versus $\boldsymbol{M_0 max}$ for a 1 degree grid with the addition of local slope lines both excluding (top) and including (bottom) 2016 data. The color-coded lines represent the expected slope of the $\boldsymbol{V}$ versus $\boldsymbol{M_0 max}$ line based on the sample size hypothesis of equation (3) using each maximum moment events local maximum likelihood b-value (colors) from Figure 19.

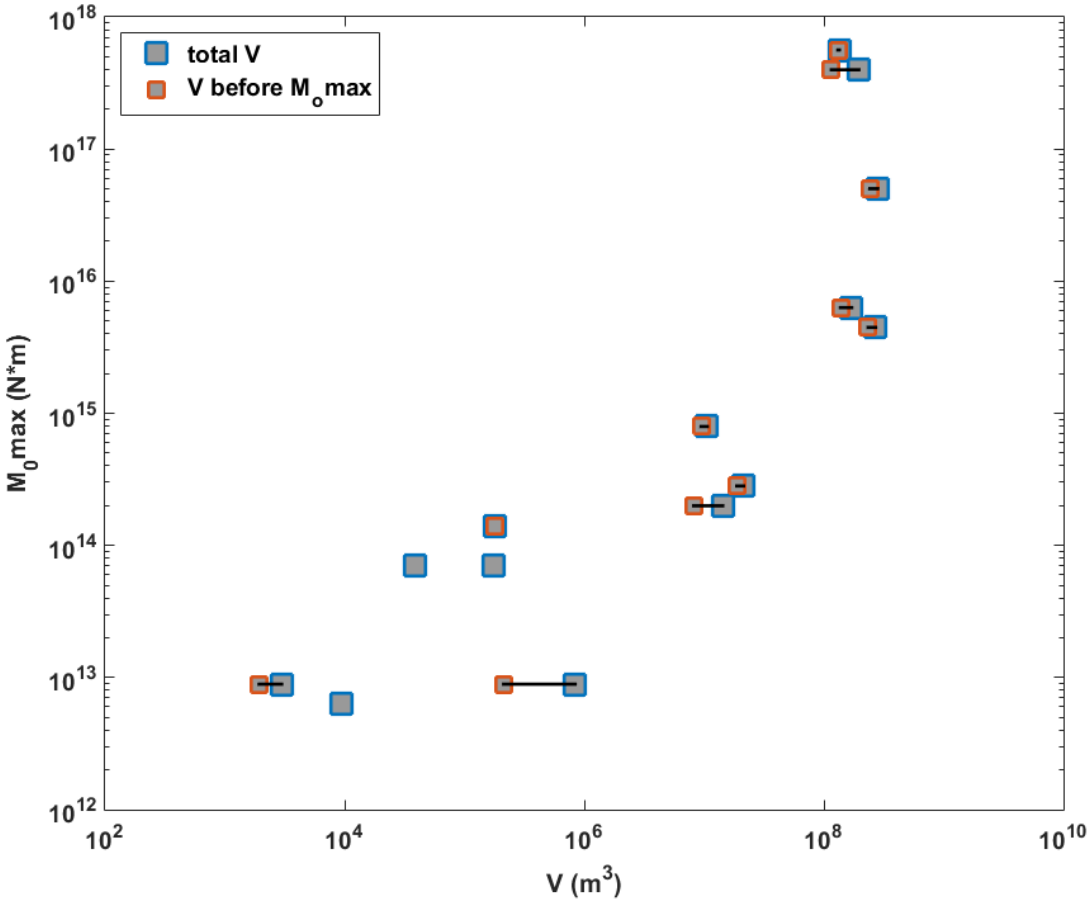

Figure 21. For a 1 degree grid cell, the volume before the maximum magnitude event ($\boldsymbol{V_{M_0max}}$) versus $\boldsymbol{M_0max}$ compared to the total volume in each grid cell ($\boldsymbol{V}$) versus $\boldsymbol{M_0max}$. Although large volume injection may occur in high volume regions after the maximum magnitude event, it is not a significant change when compared to the previous volume injected.

Chapter 3: Temporal Relationship Between Disposal Volume and Seismicity

Introduction

The expected time between when injection occurs and when seismicity is likely to occur is largely unknown in Oklahoma. On the state scale, injection began rapidly increasing during 2012, followed by rapid increase in seismicity rate about one year later (Figure 22). However, the decline of the injection rate and seismicity that began in 2015 has been near simultaneous. The volume restrictions of the OCC and reduction in oil and gas prices have both led to the decrease in disposal volume, which is the most likely reason that led to the decrease in seismicity. A statewide model of the volume-seismicity relationship based on modified Gutenberg-Richter equation suggests near historic seismicity levels will be reached in a few years (Langenbruch and Zoback, 2016). On a smaller scale however, there is substantial uncertainty. As Langenbruch and Zoback (2016) point out, the 2011 M5.7 Prague event occurred at rates below the currently mandated volume reductions. Furthermore, Oklahoma contains multiple structural terranes and varied basement geology (Shah and Keller, 2017), which can be easily broken into different pressure regions that likely have different fluid properties.

Understanding the time delay is important for multiple reasons. First, the time delay provides information about the fluid properties of the subsurface, e.g.. average permeability differences in the Arbuckle and basement, and spatial heterogeneity. Second, the time delay could provide some information about the triggering mechanism, i.e., if the triggering occurs ahead of the expected time scale of pressure

diffusion processes and if poroelastic stress transfer is required (Ellsworth, 2013; Chang and Segall, 2016). Furthermore, having a better understanding of the time delay can help regulators better understand the expected seismicity rate changes following fluid injection.

There are multiple cases of induced seismicity with a clear time delay between injection activity and seismicity, however historic examples show a wide range of delays. In the classic example of Rocky Mountain Arsenal, Healy et al., (1968) noted a similar pattern in the bottom hole pressure of the injection well and the seismicity rate with a 10-day delay. The Dallas-Fort Worth earthquake occurred at a time delay of approximately 90 days following the onset of injection from a disposal well within 0.5 km (Frohlich et al., 2011). Similarly, induced seismicity in central Arkansas had a peak correlation around 105 days between injection activity and seismicity (Horton, 2012). Another example comes from Oklahoma, where McNamara et al. (2015b) found a peak correlation at 17 days between injection rates from nearby wells and the earthquake rate of the November 2014 Cushing sequence. However, sometimes there is no clear time delay between injection and seismicity increases, such as the 2011 Prague sequence (Keranen et al., 2013).

The goal of this chapter is to test the hypothesis that there is a clear statewide time delay between injection increase and seismicity in Oklahoma. I will also quantify the variability of the time delay and determine if there is spatial coherency within different pressure regions that are defined by geological boundaries.

Method

Separating seismically active parts of the state into pressure regions

Three pressure regions, the western, northern, and eastern region, were defined based on the basement geology and possible fluid barriers (e.g., large regional faults) (Figure 23). There are five possible fluid barriers resulting in the spatial clustering of seismicity in the central and northwestern Oklahoma. The Cleveland anomaly, an area of regional magnetic and gravimetric anomaly, has a lower rate of seismicity, which is attributed to a decrease in the fracture intensity of the intrusive body (Shah and Keller, 2017). The Nemaha uplift is a similar area of low seismicity, likely because the faults surrounding the uplift are acting as fluid barriers (Shah and Keller, 2017). In central Oklahoma. the two large regional fault systems, the Wilzetta and Nemaha fault systems, are also possible fluid barriers based on the patterns of seismicity relative to these faults (Shah and Keller, 2017). Finally, the Anadarko basin, which deepens to the southwest of central Oklahoma, has almost no seismicity below 3 km basement depth (Shah and Keller, 2017). These possible fluid boundaries were used to define the three pressure regions (Figure 23).

Determining time lag between injection and seismicity

Assuming seismicity is triggered by the diffusion of pore pressure to critically stressed faults, a time delay between injection and seismicity is usually expected due to the time scale required for the pressure diffusion process. Sometimes shorter time delays are observed at relatively far distance, which is attributed to poroelastic stress transfer through rock matrix (e.g. Deng et al. 2016). To quantify a time delay between the two time-series across the state, the data was again spatially summed in grid cells.

Based on the results of the previous spatial analysis, 0.3 and 0.6 degree grid sizes were used for demonstration. The 0.3-degree grid is the minimum grid size required to demonstrate a relationship between injection volume and seismicity, while the 0.6-degree is required for the relationship to stabilize. To test the sensitivity of the results, grid sizes between 0.3 and 1.2 degrees were also tested.

For each fixed grid size, the data was processed as follows:

(1) For each grid cell, the daily number of events and volume disposed were summed and then smoothed with a 1-year running time window. The smoothing acts to reduce the effect of short-term spikes in seismic rate changes.

(2) For a time lag from 0 to 1000 days, the volume series was shifted forward and the correlation coefficient, or Pearson's r, of the two series was calculated. The Pearson's r correlation coefficient is a measure of the linear correlation between two time series:

$$r(A,B) = \frac{cov(A,B)}{\sigma_A \sigma_B} \quad 10)$$

where the correlation of times series A and B is based on the covariance of the two time series and their standard deviations σ_A and σ_B. The time lag for the maximum correlation was saved for each grid.

(3) The grid increment was set to half of the grid size. This was done to prevent representing an area only with potentially divided associated events and wells (Figure 24). As additional conditioning of the data, bins that did not contain at least 15 events and 3 wells were rejected due to the lack of data points for the correlation calculation.

(4) Two time periods were tested: 2005 to 2016, and 2012 to 2016. The first time period often produces misaligned peaks for the maximum correlation due to the

tail of low volume and low seismicity from 2005 to 2012 for most regions. For the second time period, the peaks of the two time series are often aligned more consistently.

(5) To determine spatial variability, individual results for the eastern and western pressure regions were also calculated. The northern pressure region was too small to apply this analysis. Because of the longer injection history in the eastern pressure region both 2005 and 2012 start dates were tested, however for the western pressure region only the 2012 start date was tested.

(6) Finally, a synthetic dataset was tested, where the event time was randomized, but the event and well location and the injection rate are held unchanged. This removes the temporal dependence between the two time series and tests the robustness of the observed datasets results.

Results

With grid size of 0.3 degrees, high correlation is observed in the western pressure region and the northern part of the eastern pressure region (Figure 25), with most time lags within 400 days. Poor correlation is often observed in areas along the edges of these regions, or areas with low injection and seismicity. Grid size of 0.6 degrees shows the same general trends, suggesting that the spatial variability is relative robust (Figure 26). Based on visual inspection of individual cells (Figure 27 and 28), a coefficient threshold of 0.75 is adopted to declare significant correlations between the two time series. As a general observation, cells have longer time lags tend to have lower maximum correlation. For all cells with significant correlation, the time lags are mostly under 500 days (Figure 29). The synthetic catalog shows significantly lower correlations than the observed result, confirming the robustness of the observed results,

however, the synthetic still data still exhibits the trend of decreasing correlation with increasing lag time at max correlation (Figure 30). A portion of the trend of decreasing correlation with increasing time lag is an artifact of the correlation calculation that is not extremely significant and is addressed in depth in Appendix B.

When comparing results for all grid sizes from 0.3 to 1.2 degrees, larger grid size tends to have narrower range of time lags from cells with significant correlation. The median time lags for all grid sizes range between 75 and 200 days (Figure 31). When looking at only the western pressure region, the same trend is observed, and the median time lags range between 150 and 200 day lags (Figure 31). A similar trend is apparent in the eastern pressure region for both 2005–2016 and 2012–2016 time periods (Figure 31). However, the longer time period (2005–2016) exhibits a wider range for lag times compared to the shorter time period, which is likely due to the lack of temporal correlation between 2005 and 2012.

Discussion

Both empirically and theoretically, the level of seismic activity scales with injected volume (McGarr, 2014; van der Elst et al., 2016); therefore, a temporal correlation between injection rate and seismicity rate is often expected. The calculation of the time delay works under the assumption of a time-dependent growth of the pressurized zone due to pressure front propagation, which is related to the diffusivity of the subsurface. With a larger grid size, the maximum distance tested in the volume-seismicity relationship increases, but it may likely average the time delay between smaller groups of associated wells and earthquakes.

Statewide observation clearly demonstrates that an increase in injection rate corresponds to an increase in seismicity rate, while a decrease in injection rate corresponds to a decrease in seismicity rate (Figure 22). What is being observed is not the effect of injection from a combined well triggering all combined seismicity with a fixed day lag, but the averaged effect of tens to hundreds of associated groups with a range of time delays. This trend is also visible in the spatial mapping of time lags for cells with high correlation for different grid sizes (Figure 31). At smaller grid sizes, the spatial distribution is more heterogeneous than larger grid sizes.

The 0.3-degree grid represents the smallest scale at which the volume seismicity relationship is observable (see chapter 2), therefore, it also represents the smallest scale to view the small-scale heterogeneity of time-lags. If events in a given grid cell are being triggered by well activity further than 0.3 degree, then there should be a low correlation between the two time series in that cell. As such, only the high correlation cells are used to infer the inter quartiles and median of time lags. For the whole state, western region, and eastern region the 25^{th}-50^{th}-75^{th} percentiles from the 0.3-degree grid are 0-90-178, 19-160-343, and 2-128-174 days, respectively. For the entire state, entire western region, and entire eastern region the max correlation occurs at a time lag of 215, 242, and 127 days respectively. All of these values have substantial uncertainty that could affect the time lag. A large source of error is the distance between wells and the associated seismicity likely varies among cells. Another reason is that the time lag represents combination of lateral diffusion through the Arbuckle with relatively higher diffusivity, and lower diffusivity within basement faults through downward diffusion

(Shapiro et al., 1997). However, the range of time delays provides an empirical range for regulators to consider.

There is clear spatial variability both between the two major pressure regions and within individual pressure regions. Overall, the western region has longer time delay than the eastern region, which could be an indicator of lower Arbuckle or fault diffusivity in the western region. Along the southern edge of the eastern pressure region, there are low correlations at both 0.3 and 0.6-degree grid sizes (Figure 25 and 26). This area includes the earliest likely induced seismicity in the state that began in 2009 (Keranen et al., 2014). The seismicity in the region, labeled as the "Jones swarm", was likely triggered by a group of high volume wells along the southern edge of the Nemaha fault (Keranen et al., 2014). The low correlation is partially due to the lack of wells in the middle of the swarm area (Figure 32), but also the narrower time series (2012–2016 used for Figure 25 and 26). When the high-volume wells likely tied to the Jones swarm are compared to nearby seismicity with a longer time period of 2000–2016 data, there is still only a 50% correlation but the time dependency between the two series with a 180 day time lag is evident (Figure 33). The low correlation is likely due to the injection between 2005 and 2009, which isn't mirrored by seismicity, acted to build pressure near the well and it wasn't until 2009 onward that the pressure perturbation could migrate northeast up the structural dip (Keranen et al., 2014) and into the Jones swarm. This concept agrees well with the pressure modeling results of the region (Keranen et al., 2014).

Another interesting result are the handful of cells that show very small or zero-day time lags yet still have extremely similar time series (Figure 29). If the seismicity

was happening directly next to the disposal well in the subsurface, such short time delays could be due to pore pressure triggering, however, the majority of Oklahoma earthquakes have occurred in the crystalline basement (McNamara et al., 2015a), which indicates this seismicity could be due to rapid elastic stress triggering (e.g., Deng et al., 2016).

Conclusions

The time lag calculation gives a useful estimate of the time delay between injection and seismicity, as well as a general indicator of spatial variations of time delays across the state. At the state level there is a clear 215-day time lag between the peak in injection and seismicity. When the small scale is looked at an interquartile range between 0 and 179 days represents the likely range of possible time delays in a region around 40 km in diameter. The spatial variations show both large and small spatial scale's time delay in the western region are longer than in the eastern region. This could be showing a higher fault or Arbuckle permeability in the eastern region or could be caused by some other factor. Finally, there are a few areas where there are extremely small delays between injection and seismicity, which could be indicating poroelastic triggering.

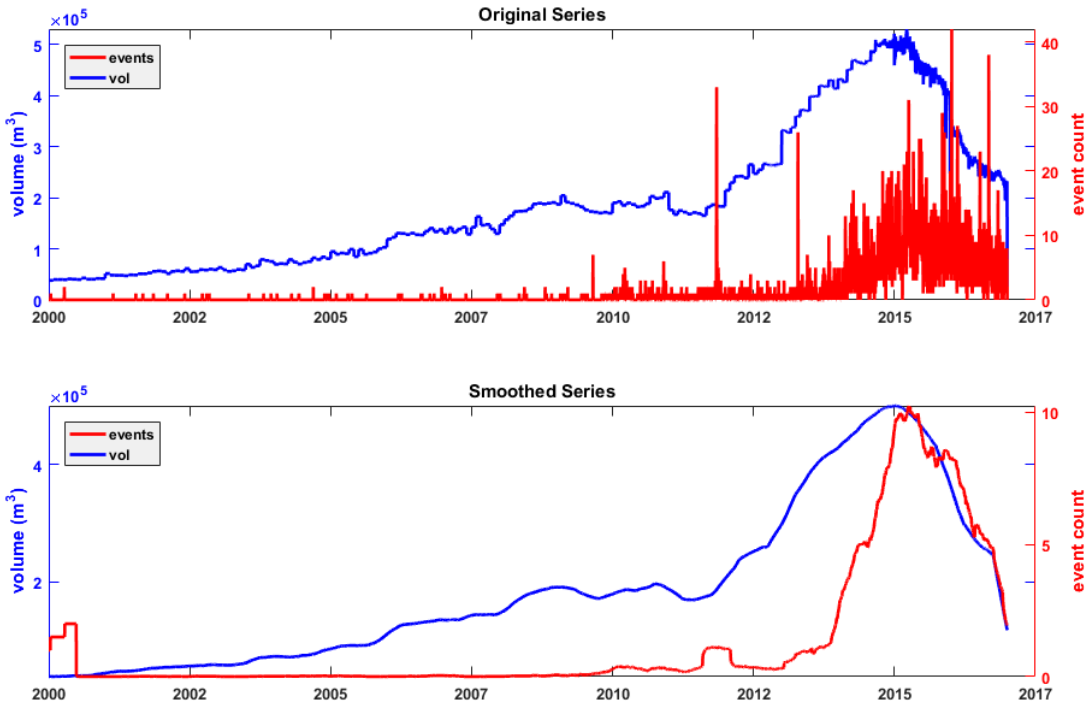

Figure 22. Statewide variation in Arbuckle disposal volume and earthquakes greater than M2.5 over time. (Top) Unsmoothed original data, (bottom) time series smoothed with a 180-day mean smoother. A time delay of about a year and a half is present between the increase in volume and the increase in seismicity. The two times series decline near simultaneously after the peak in volume in early 2015.

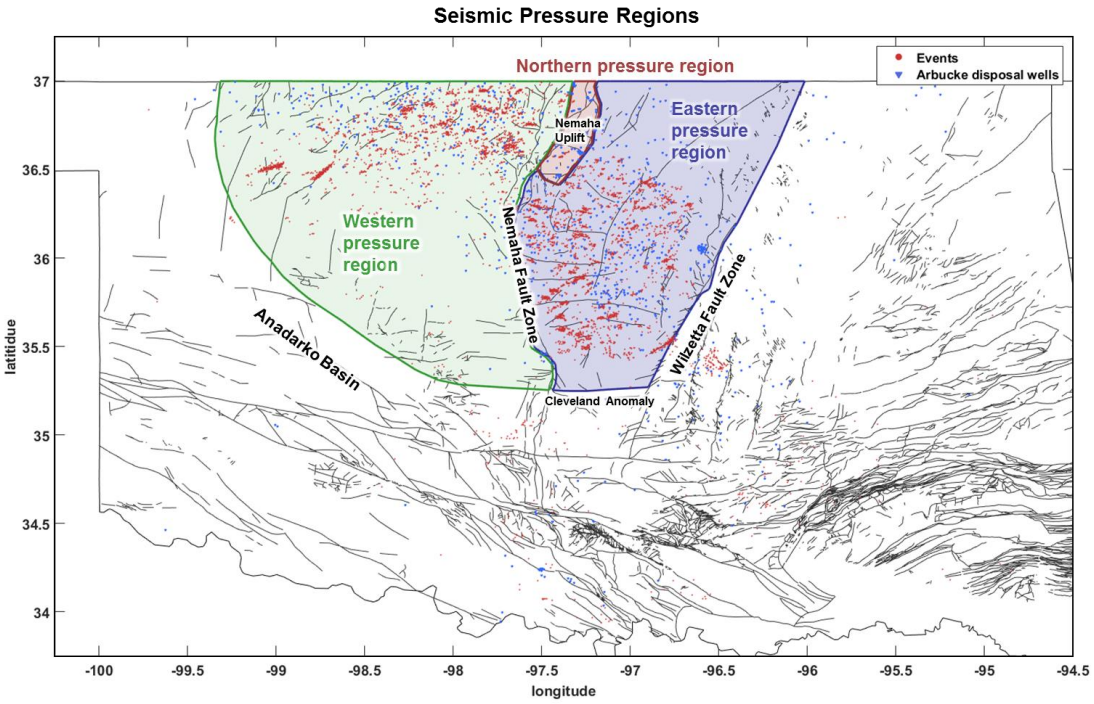

Figure 23. Map of Oklahoma depicting known faults (grey lines), earthquakes, Arbuckle disposal wells, and three pressure regions and their bounding structures. Faults from Holland (2015).

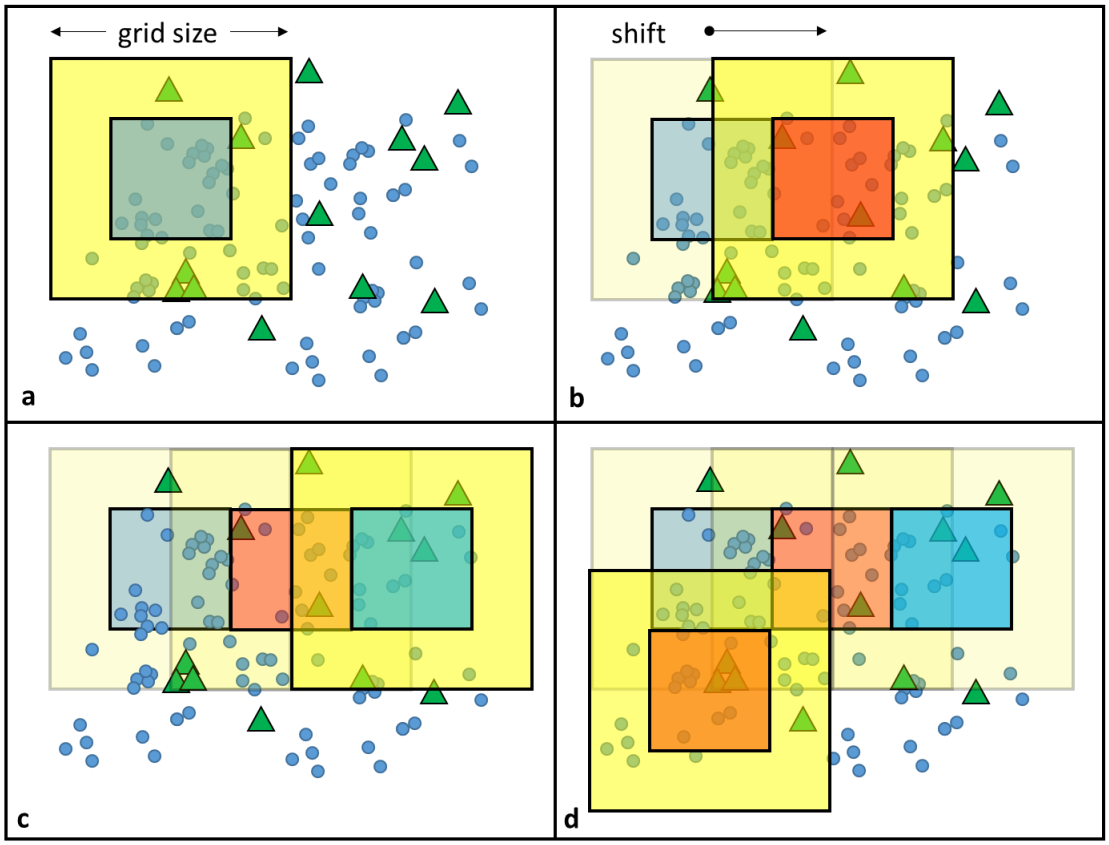

Figure 24. Example of the grid cell shifting and representation on the map of Figure 25 and Figure 26. At time (**a**) that volume and seismicity time series within the grid cell (yellow) is shifted for various lag times until the max correlation is found. A visualization cell half the dimension of the actual grid cell is placed to represent the magnitude of the correlation. At time (**b**) the grid cell is shifted half the length of the cell and the time series are again compared and the magnitude of the correlation is represented by the inner cell. This is repeated (**c**) until the edge of the study area is reached, in which case a new row is scanned. This also prevents potentially high correlations from being misrepresented by a well and associated seismicity being split between cells, for example the high (red) correlation at step (**b**) would be missed if overlapping cells were not used.

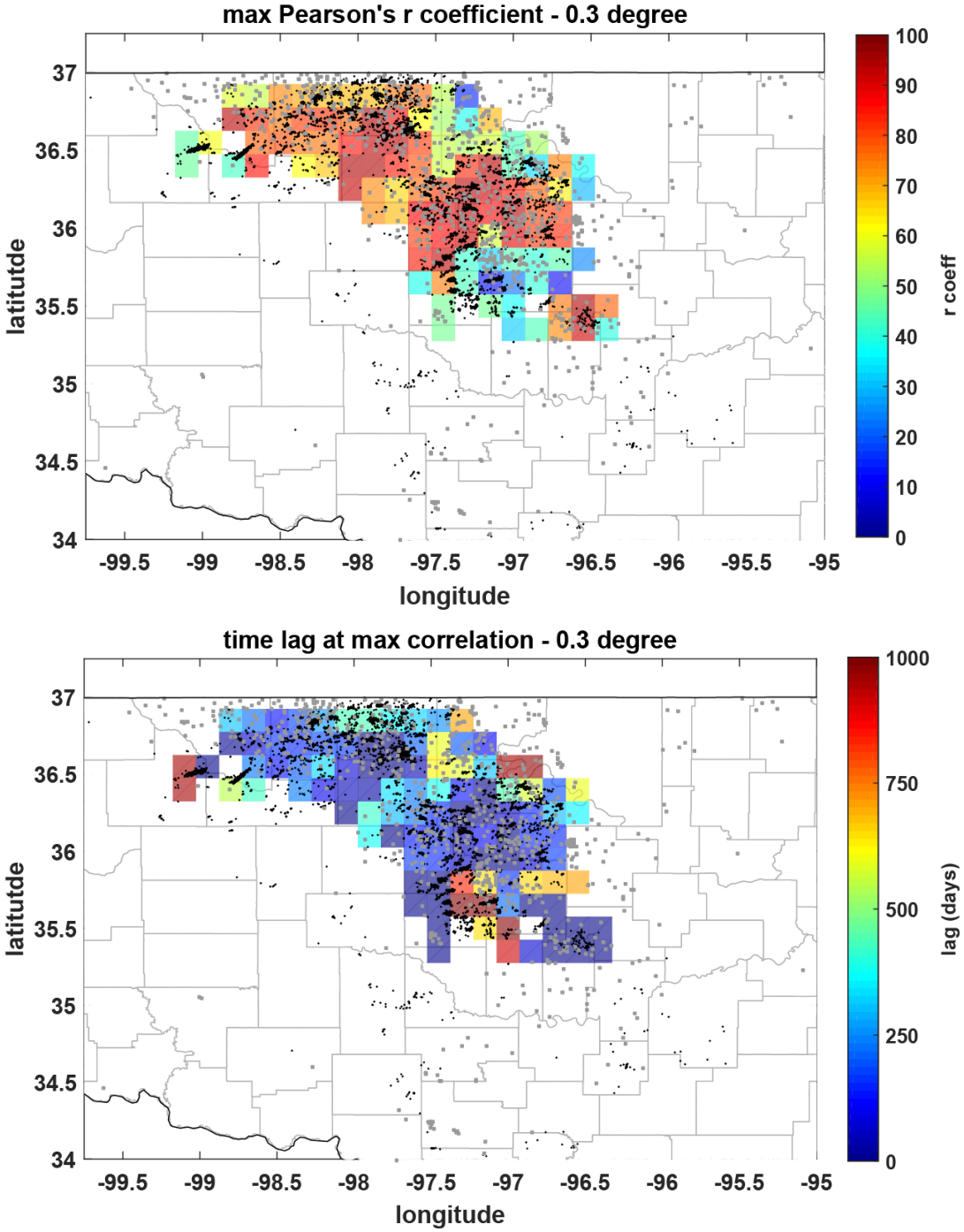

Figure 25. Map view of events (black) and wells (grey) on top of the max Pearson's r correlation coefficient (top) and the time lag at max correlation (bottom) for the daily volume and number of events time series with a 0.3 degree grid size. Most areas with high correlation are located where there is both dense wells and seismicity. As seen in Figure 24, the colored grid cells are half the actual grid dimension, so in this case 0.15 degrees.

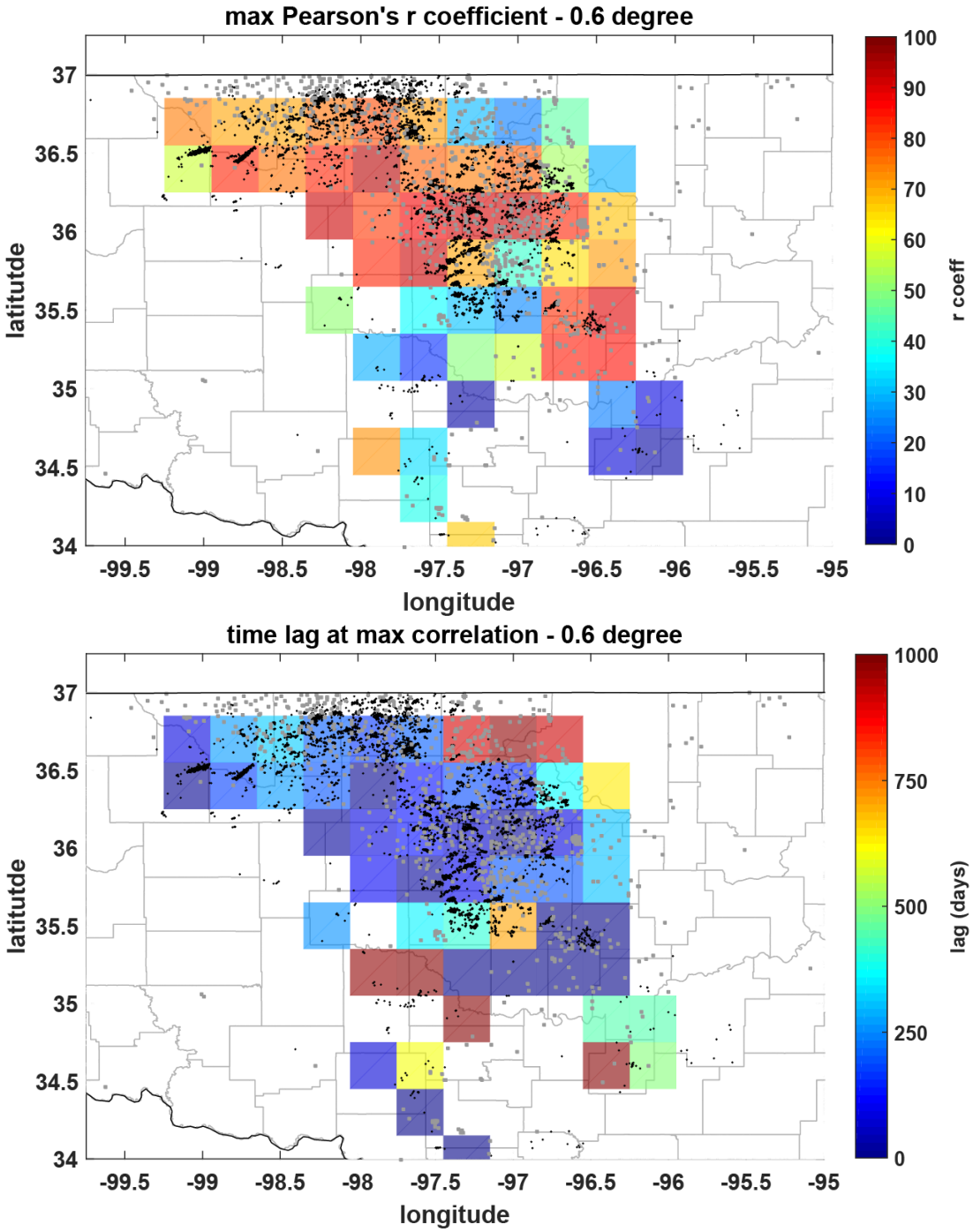

Figure 26. Map view of events (black) and wells (grey) on top of the Max Pearson's r correlation coefficient (Top) and the time lag at max correlation (Bottom) for the daily volume and number of events time series with a 0.6 degree grid size. Like the 0.3 degree grid, low correlation occurs along the edges of the main seismic regions where there is fewer wells and seismicity.

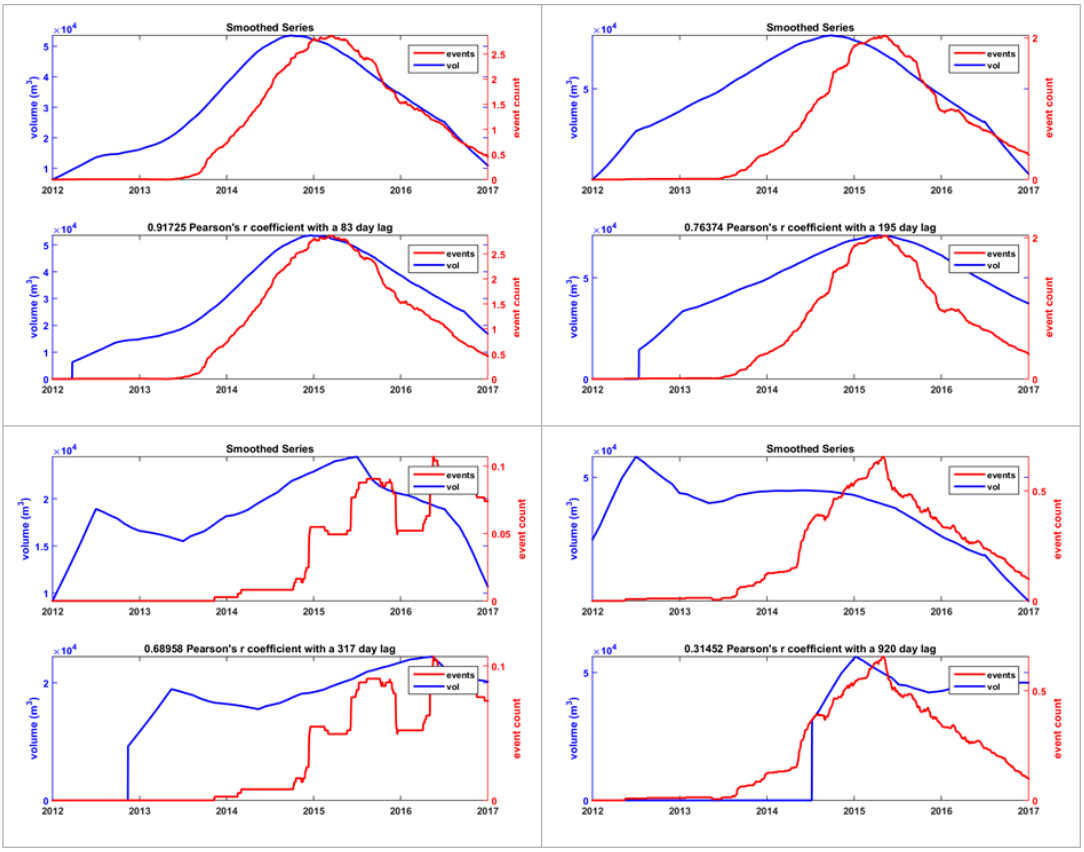

Figure 27. Example of four different grid cells that show a range of max correlations at long time lags. Each quadrant shows the original smoothed time series (top) and the volume shifted time series with the maximum correlation. The number of daily events is very low due to the one year smoothing of the time series.

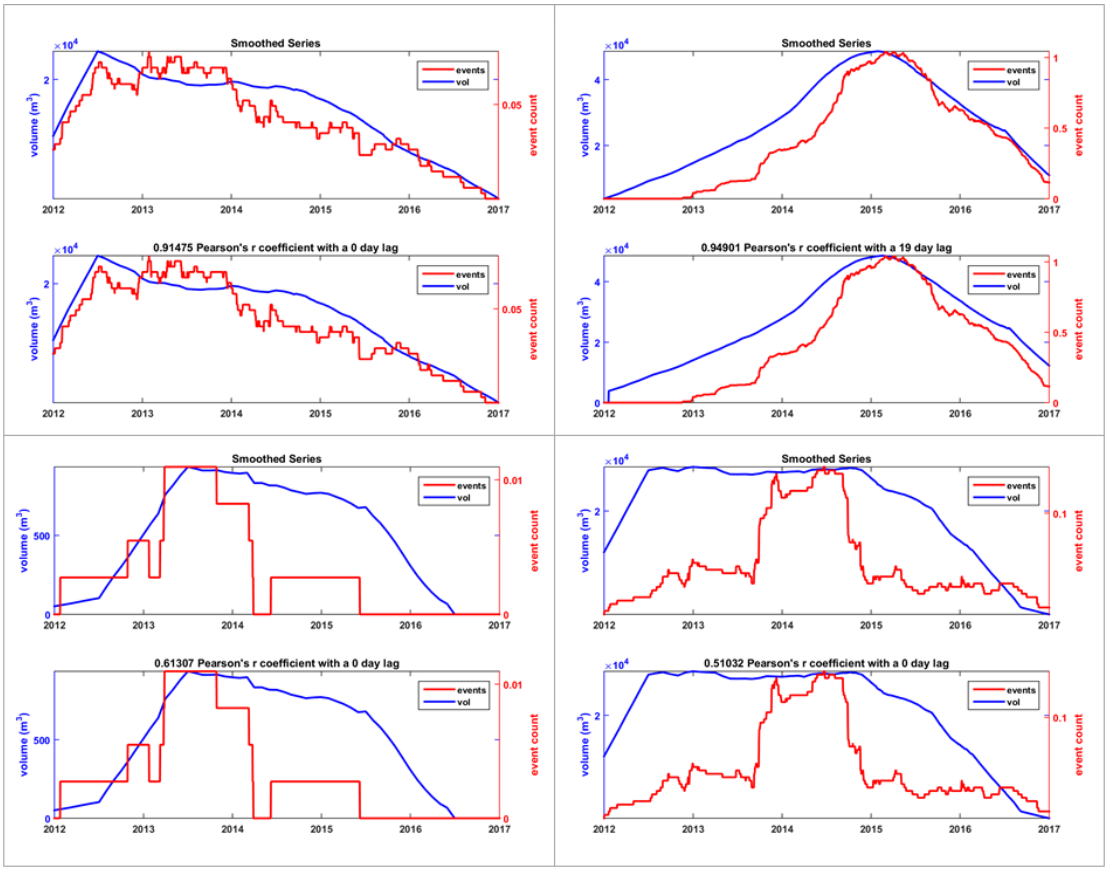

Figure 28. Example of four different grid cells that show a range of max correlations at short time lags. Each quadrant shows the original smoothed time series (top) and the volume shifted time series with the maximum correlation.

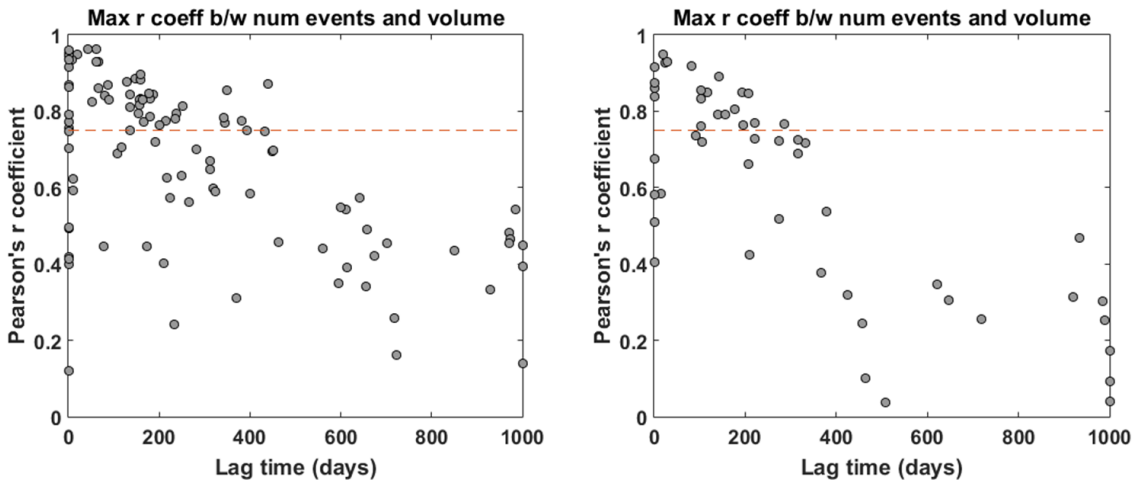

Figure 29. Max Pearson's r correlation coefficient against the corresponding time lag for the daily volume and daily number of events time-series with a 0.3 (left) and 0.6 (right) degree grid size. The dashed line represents the 75% r value, used as a cutoff to group high correlation grid cells. Both the 0.3 and 0.6 degree grids show a decrease in correlation coefficient with increasing lag time and several cells with max correlation at a zero day time lag.

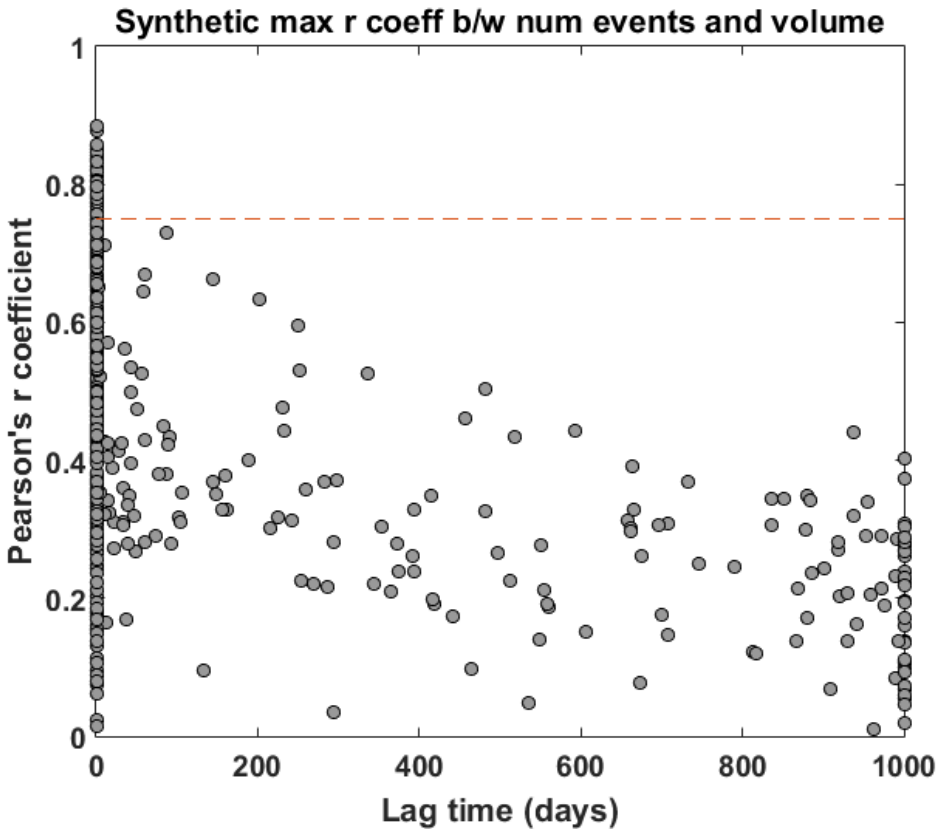

Figure 30. Result of 10 bootstrapped iterations of randomly distributing the event times between 2012 and 2017 while keeping locations of events and wells and injection rate of wells constant. There is still a trend of decreasing maximum correlation with increasing lag in that data.

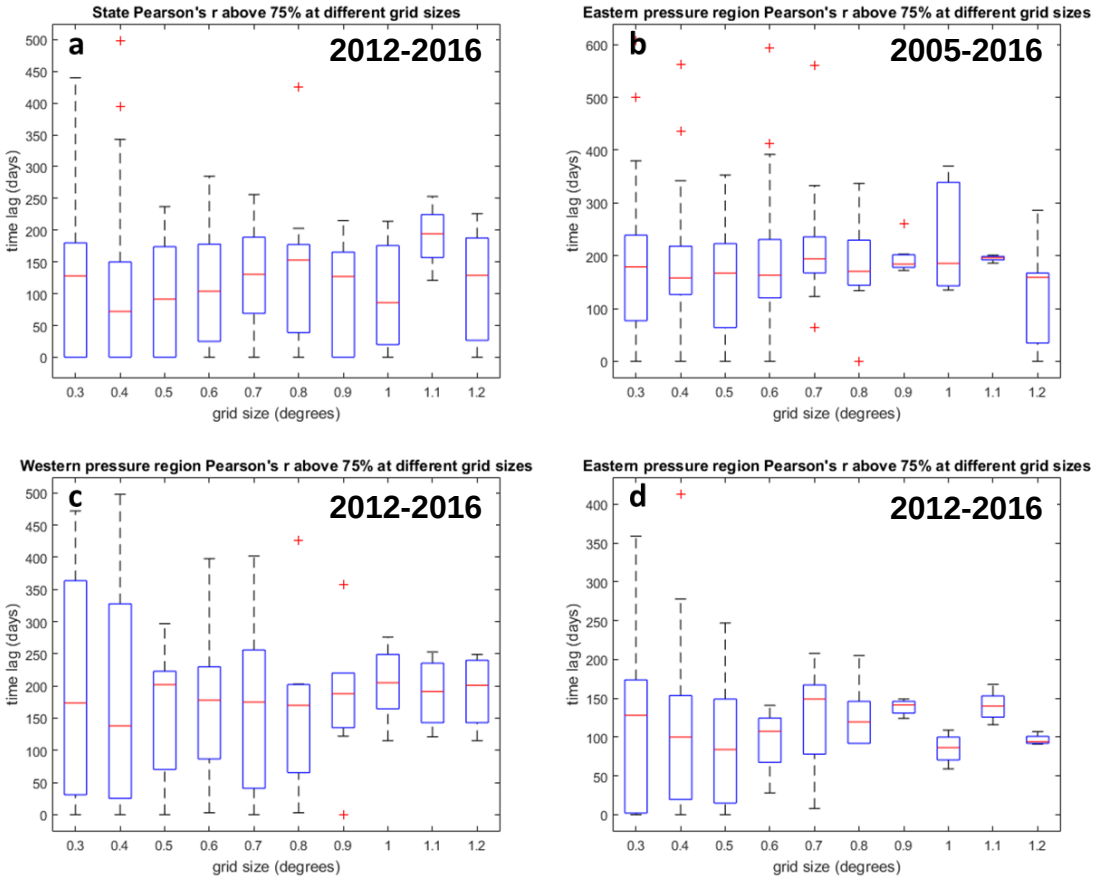

Figure 31. Box and whisker plot of the distribution of time lags for high correlations at a range of grid sizes. The red plus sign represents outliers, the whiskers represent the 99.3% coverage of the data assuming a normal distribution. The outer box edges are the 25th and 75th percentiles of the data and the red line is the median. Each increase in grid size represents fewer cells that are representing the data. At the state level, the distribution stays consistently under a 200 day time lag (a). The eastern pressure region shows longer time lags when the 2005-2011 data is included (b), but both time series show a consistent median, for the 2005-2016 time series it is around 200 days (b), for the 2012-2016 time series it is around 125 days (d). The western pressure region also has a consistent median around 200 days and a narrower range of lags with increasing grid size.

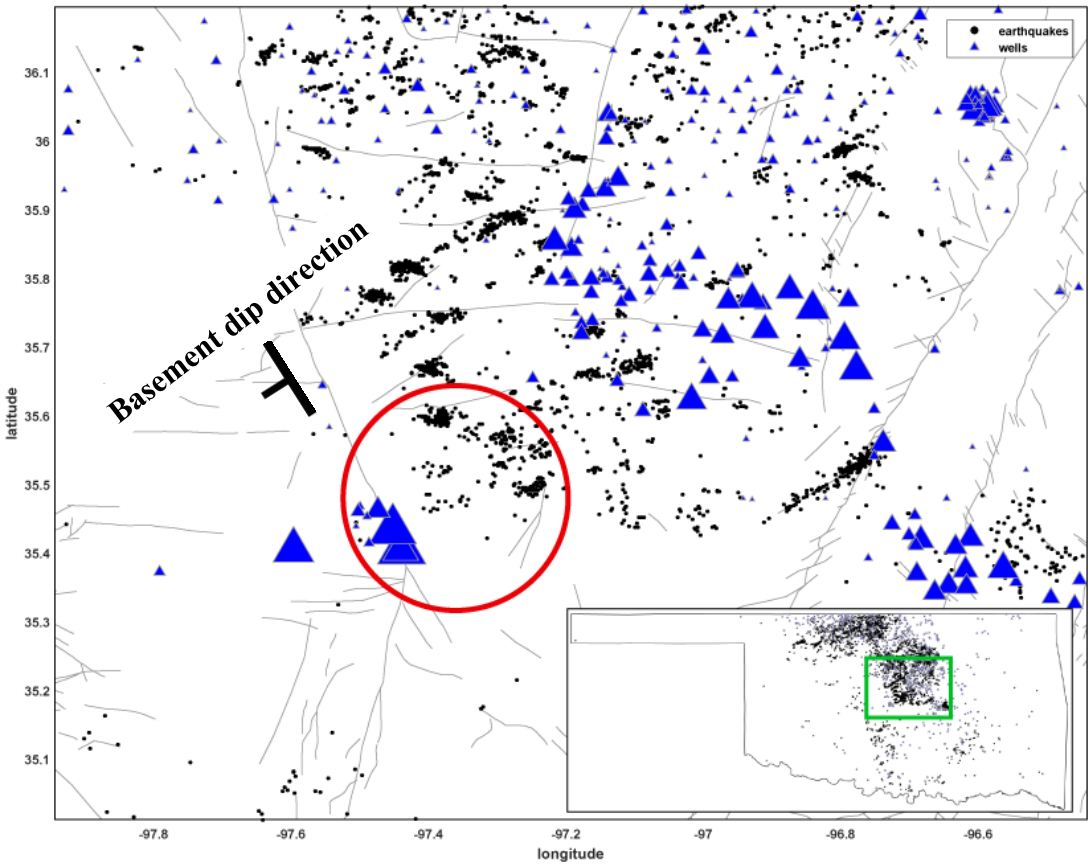

Figure 32. Map of the central Oklahoma disposal wells, scaled by volume, seismicity, and faults (grey lines). The high volume wells located at the southeastern edge of the eastern pressure region and nearby seismicity encircled in red are compared in Figure 27. The structural dip of the basement is to the southwest.

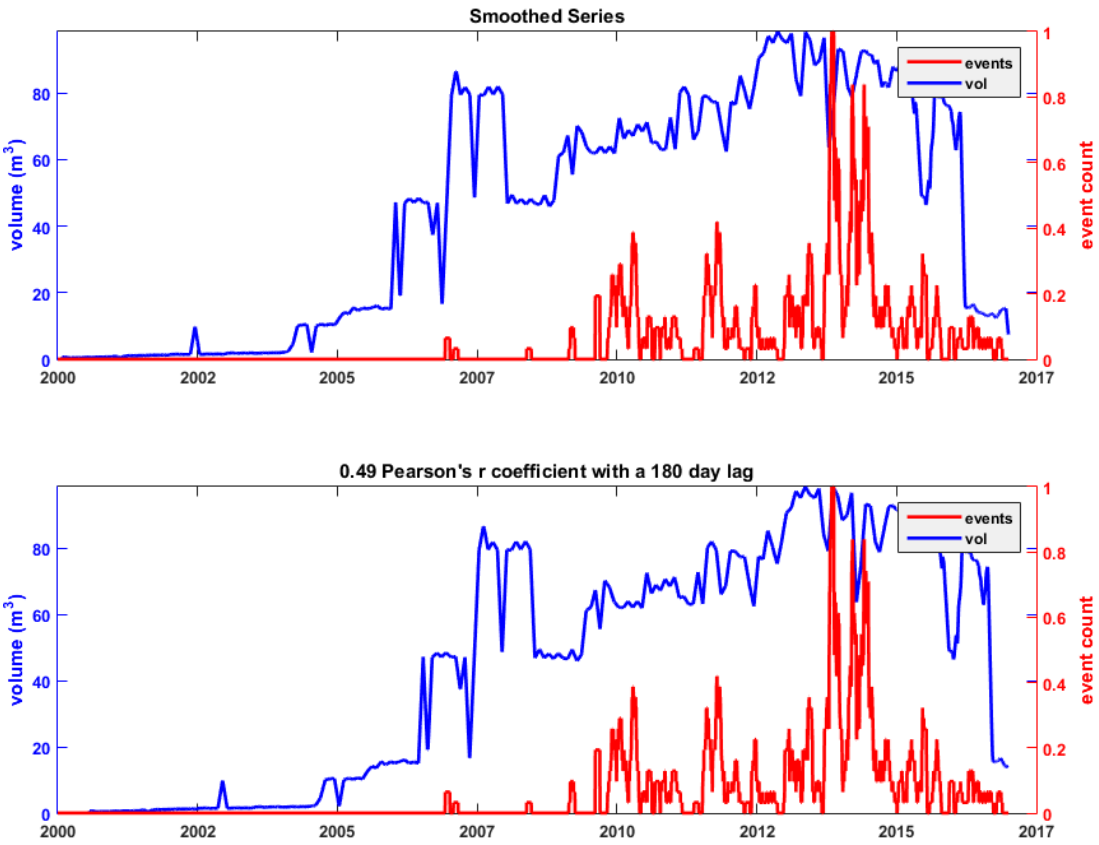

Figure 33. Daily injection and number of events over time for the southwestern edge of the eastern pressure region (top). The volume time series had a max correlation with a 180 day shift (bottom). Both time series are smoothed with a 90 day mean operator. While the correlation is low, there is still a clear delay and buildup in both injection and seismicity.

Chapter 4: Pore-Pressure Diffusion Triggering of Seismicity

Introduction

An important aspect of Oklahoma induced seismicity is the triggering mechanism. If the triggering mechanism is well understood, then a plan can be taken to reduce seismicity if it is human triggered. Since Oklahoma first began experiencing an increase in seismicity, different seismic episodes have been attributed to natural causes (Keller and Holland, 2013), pore pressure triggering from nearby well activity (Keranen et al., 2013), hydraulic fracture triggering (Holland, 2013a), and remote earthquake triggering (van der Elst et al., 2013). Although the majority of Oklahoma case studies hypothesize that pore pressure diffusion triggering is the most likely mechanism, a statewide analysis supporting pore pressure diffusion triggering has not been carried out.

If the seismicity is being directly caused by disposal fluid, there are two main triggering mechanisms. The first is the diffusion of pore pressure into a critically stressed fault. As the fluid pressure increases in a fault, it lowers the normal stress and allows shear failure by:

$$\tau_{crit} = \tau_0 + \mu(\sigma_n - P) \quad (11)$$

where the critical shear stress τ_{crit} is equal to the cohesion τ_0 plus the product of the frictional coefficient μ, and the difference between the normal stress $\sigma_n$and the pore pressure P (McGarr, 2002). When fluid enters a fault zone there is also a secondary chemical effect that reduces the frictional coefficient μ (Shapiro et al., 1997). The second mechanism is through elastic stress transfer. As pore pressure increases or decreases, deformation occurs in the host rock that results in an elastic stress change

(McGarr, 2002; Chang and Segall, 2016). Furthermore, there is a change in mass after injection, resulting in an increase in the vertical load and corresponding stress (McGarr, 2002; Ellsworth, 2013). The elastic and pore pressure mechanisms can be simultaneous, i.e. poroelasticity. When fluid pressure is increasing in a fault there will be both a decreasing normal stress as well as changes in the normal and shear stresses caused by elastic stress changes in the host rock. However, if the fault is not in communication with the fluid source, there will only be stress changes due to elastic stress transfer.

Typical examples of elastic stress transfer induced seismicity are cases of reservoir-induced seismicity, such as the M5.8 Kariba dam earthquake that occurred shortly after the reservoir reach maximum depth (Simpson et al., 1988). However, poroelastic models show it is a possibility for elastic stress due to a fluid injection generated pore pressure increase to trigger earthquakes on faults not hydraulically connected to the fluid source (Chang and Segall, 2016).

Most cases of fluid-injection induced seismicity are believed to be related to pore-pressure perturbations (McGarr, 2002; Ellsworth, 2013). Seismicity and pore-pressure diffusion generally are related by three lines of evidence. First, the events are spatially located close to the injection sources, such as the classic example of Rocky Mountain Arsenal, where all seismicity occurred within 10 km of the disposal well in a previously nascent area (Healy et al., 1968). Second, unlike instantaneous stress transfers, diffusion triggered events occur after some time delay, as the fluid pressure must first build up and propagate to the critically stressed fault (Simpson et al., 1988). Third, triggering by pore-pressure perturbations can result in the spatial migration of earthquakes. This spatial migration usually takes the form of a diffusion curve, where

events with increasing distance from the fluid source will only occur behind a spherically diffusing pressure front (Shapiro et al., 1997). This spatial migration has been used to approximate the crustal permeability through estimating the hydraulic diffusivity (Talwani and Acree, 1984; Shapiro et al., 1997; Chen and Shearer, 2011; Llenos and Michael, 2013) and is not observed in typical mainshock-aftershock sequences (Vidale and Shearer, 2006). In this chapter, I will be testing the hypothesis that Oklahoma seismicity shows signs of pore pressure diffusion in the spatial migration of seismicity. I will attempt to determine the scale of the migration, the rate of diffusivity, and how often the signature of pore-pressure diffusion is present in Oklahoma.

Method

Pre-processing earthquake catalog – clustering and aftershock removal

A hypoDD relocated OGS catalog containing 2010–2016 seismicity that employed an improved 3D velocity model was used for the diffusion analysis (Chen, 2016). A relocated catalog is important in diffusion analysis because the relative location of events dictates the strength of the diffusion pattern. Using the zmap software to calculate the Mc of the relocated catalog resulted in a minimum magnitude of Mc = 2.5 for a 90% goodness of fit (Wiemer and Wyss, 2000). The b-value was calculated as 1.11 using the maximum likelihood method (Aki, 1965), and 1.34 using a bootstrapped least-square fitting method (Figure 34). These values are different than the b-values estimated in chapter 2, due to the different catalog time range and magnitude distribution. Despite a low-end Mc of 2.5, only events below M2 were removed from the catalog. This was done to increase the number of events grouped into each cluster,

without which the diffusion fitting would be extremely limited. Next, the relocated catalog was clustered using nearest neighbor rescaled distance after the method of Zaliapin and Ben-Zion (2013). This method finds each event's distance and time to its closest neighbor scaled by the magnitude of the parent event. The rescaled space and time distances are:

$$T_{ij} = t_{ij}10^{-qbm_i}\ ;\ R_{ij} = (r_{ij})^{d_f}10^{-(1-q)bm_i} \qquad 12)$$

where T_{ij} and R_{ij} is the rescaled time and distance between event j, and its parent i, t_{ij} and $r_{ij}$is the original interevent time and distance, q is 0.5, $m_i$is the magnitude of the parent event, b is an average b-value of 1.2, d_f represents the dimension of the earthquake hypocenter distribution, 1.6 was used in this study. From that scaled space-time distance it identifies unusually close events that can be categorized as clusters. When applied to southern California seismicity, distinct groupings of clustered versus background seismicity was apparent (Zaliapin and Ben-Zion, 2013). In Oklahoma, there is not nearly as clear groupings of background and clustered events (Figure 35). Instead, visual examination of a range of possible space-time thresholds was performed to select the optimal thresholds that best group the clusters (Figure 36). The final thresholds employed are T = 5, R = 0.009, where interevent distances below those thresholds were identified as clusters. The clustering resulted in 74 clusters with over 20 events.

Some clusters contained clearly separated delineated faults within them. To prevent overlapping diffusion signals in such clusters, these faults were manually separated, resulting in 89 total clusters. Next, aftershocks were removed within each cluster with the space-time windowing method of Gardner and Knopoff (1974). This is done to remove the space-time imprint of aftershock triggering from the diffusion

curves. The narrower windows suggested by Urhammer (1986) were used in this method, which agree better to this dataset than the Gardner and Knopoff (1974) window size. This method designs a magnitude-scaled spatial and temporal window after a mainshock, in which events are considered aftershocks. For example, a M5 event will have any event that occurs within 27 days and 20 km after it be considered an aftershock and removed, while for M3 it is only 2.3 days and 4 km. Aftershock removal eliminated 5,207 events, resulting in 81 clusters containing at least 20 events. The 20 events requirement was established as a threshold for minimum number of events for the diffusion migration analysis.

Cluster-based diffusion curve fitting

Seismicity that demonstrates diffusion triggering can be detected by fitting the migration of events with a diffusion curve. A spherical diffusion of increased pore pressure will radiate out based on the properties of the formation. The diffusion front can be approximated by:

$$r = \sqrt{4\pi D t} \qquad 13)$$

where r is the distance the pressure change has traveled by time t for a given hydraulic diffusivity D (Shapiro et al., 1997). The method of Chen and Shearer (2011) was used, that employs weighted windows to find a best-fit diffusion curve for seismicity away from the first event in the cluster. The fundamental premise of the method is the first event in a cluster is a proxy for the fluid source, and as time progresses events will occur away from the first event within the space-time window of the pressure front of equation (13). By varying the diffusivity value, a best-fitting diffusion rate can be found for each cluster (Figure 37).

The method allows a measure of confidence, where the total misfit between the observation and the diffusion curve prediction are calculated based on a residual-based weighting scheme (Figure 37). To test the statistical significance of the migration, a resampled dataset was generated for each cluster, where the spatial locations are held constant, but the occurrence times are randomly shuffled for all events. This discards any time-dependent spatial migration information in the original database. This process was repeated 1000 times, and the misfits from the 1000-resampled datasets were compared with the original dataset. This tests how likely the misfit of the observed data was if the events had happened in a random order. The percent of the resampled datasets with misfit greater than the original data represents the confidence in the diffusion trend. Like Chen and Shearer (2011), I find a confidence threshold of 85% agrees well with clusters that show a clear diffusion pattern from visual examination, while clusters with confidence below 85% show weaker signs of diffusion migration. An additional step carried out in this study not done by Chen and Shearer (2011) was the removal of events that occurred in the first 10% of each clusters time series. This helps prevent false-positives caused by the weighting method applied on early events. The diffusion curve fitting was carried out to all clusters that have over 20 events after aftershock removal. Due to inherent inaccuracy in the depth location of events, the diffusion curve fitting was tested with both 2-D and 3-D distance. A grid search over 50 diffusivities that are equally logarithmic spaced from 0.0001 (10^{-4}) to 1 (10^{0}) m^2/s was performed to find the best-fitting diffusivity.

Cluster fluid migration direction

An additional part of recognizing pore pressure diffusion was identifying the direction of the fluids migration. First the strike vector of each seismic cluster was calculated by determining the primary eigenvector of the covariance matrix of demeaned earthquake locations after the methodology of Vidale and Shearer (2006). Events were then binned into 20 time bins spanning the cluster's duration. Each time bin's average location was calculated and projected onto the strike vector, with the first event set as the origin. Clusters were then categorized as unilateral versus bilateral by how many time bins were located on either side of the origin. If over two thirds of the bins were on one side of the origin, then the average location of events is migrating in that direction, i.e. unilateral. An example of the time binning and diffusion vector can be seen in (Figure 38). This method gives the direction that pressure migrates along the fault, however, it does not give information on the direction the pressure came from outside the fault. For example, if the fluid came from a source in a direction near perpendicular to the strike of a fault, then the migration arrows, be it unilateral or bilateral, will point perpendicular to the actual fluid source direction.

Contour mapping of disposal volumes

To represent the disposal well data, the injection data was binned on a grid of 0.02 degree bins (approximately 3 km) and in each bin the total volume injected between 2005 and 2017 was summed. Next, the binned data was smoothed by a circular filter out to three adjacent cells (approximately 9 km). Contours were fit to the data, representing the 50^{th}, 75^{th}, 95^{th}, and 99^{th} percentile of smoothed total injection volume. In this case, the 95^{th} percentile contour line represents bins where more fluid has been

disposed than 95% of the bins in the state, represents approximately 40% of the total smoothed volume, and represents areas that have had over 99,000 barrels of fluid injected/km^2.

Large scale diffusion patterns

The 2008–2013 Jones swarm in central Oklahoma showed a high likelihood of being triggered by far-field pore-pressure diffusion (Keranen et al., 2014). If triggered from a single source, this regional scale diffusion should show a pore-pressure diffusion curve. To test this hypothesis, a similar diffusion curve fitting algorithm was used to automatically detect a diffusion pattern for a larger scale. The OGS catalog from 2005 to 2016 is used for this portion of the study because of its lower magnitude of completeness compared to the relocated catalog. Events outside the three main pressure regions (Figure 23) were removed to prevent likely non-related seismicity affecting the curve fitting. The Woodward cluster was also removed due to its potential complication in higher probability of far-range poroelastic triggering (Goebel et al., 2017). Instead of looking at individual clusters and testing from the first event, seven large regional injection highs, each representing the 95th percentile of injection volume, were used as fluid-source points. Events were associated with the nearest fluid-source or 'seed' point and the diffusion misfit was calculated with start dates from 2005 to when 10% of the associated events have occurred in 3 month increments. A wide range of diffusivities, from 0.25 to 2.5 m^2/s in 0.25 m^2/s increments were examined and the start date and diffusivity with the lowest misfit was saved for each fluid-source point.

Results and Discussion

Cluster-based diffusion curve fitting

Fitting the 2-D distance results in 33/81 clusters with a fit confidence above 85% while the 3-D distance had 30/81 clusters with a high fit (Figure 39). This resulted in a total of 42/81 clusters above 85% confidence between the two methods and an average diffusivity of 0.03 ± 0.04 m^2/s. The statistics between 2-D, 3-D, and joined results can be found in Table 1. The data shows a narrow range of diffusivities, where even the majority of low confidence fits have diffusion rates below 0.1 m^2/s (Figure 39). The clusters with high fit show a good match to the normalized diffusion curve (Figure 39). Also noteworthy the best-fitting diffusivities for all clusters follow a normal distribution in log space (Figure 40). Examples of some large, notable clusters that show diffusion are the Guthrie cluster, the Fairview cluster, the Prague cluster, and the Logan County cluster (Figure 41). As seen in Figure 36, 28 of the high confidence diffusion clusters show unilateral migration while 14 show bilateral migration.

About 52% of the large clusters in the state show strong signs of pore pressure diffusion triggering. Although other factors could be affecting the seismicity rate in these clusters, for example Coulomb stress changes from previous events or poroelastic effects. In cases such as the Fairview (2D example) and Guthrie (3D example) clusters, numerous events are outside the diffusion front despite an apparent overall diffusion pattern (Figure 41). These events take the form of vertical spikes in the diffusion curve and likely represent embedded mainshock/aftershock sequences that were not removed by the aftershock filtering, or due to larger depth errors. These spikes likely represent combined aftershock triggering. As the fluid diffuses, events are triggered within the

pressure front, but larger triggered events can generate aftershocks sequences that will rapidly propagate through the fault. The clusters that do not show clear diffusion are either still diffusion triggered but are dominated by mainshock/aftershock sequences, clusters composed of multiple faults with their own diffusion signature, or, represent mainshock/aftershock sequences triggered by alternative mechanisms such as poroelastic stress changes from nearby injection or Coulomb stress changes from nearby earthquakes. However, the average diffusivities of all clusters is consistently around 0.05 m^2/s (Table 1), which suggests that despite a low confidence there is consistency in the likely migration rate of seismicity away from the first event.

The range of diffusion rates are lower than typical cases of induced-seismicity. A review of 22 case histories of reservoir induced seismicity found diffusivities ranging from 0.5 to 50 m^2/s (Talwani and Acree, 1984). Case studies of fluid injection induced seismicity have shown diffusion between 0.2 and 1 m^2/s (Shapiro et al., 1997; Chen and Shearer, 2011). In the case of the Guy-Greenbrier Arkansas swarm, events migrated away from nearby disposal wells within a diffusivity range of 0.01 to 0.1 m^2/s (Llenos and Michael, 2013). The Guy-Greenbrier swarm shows similar diffusivities to Oklahoma clusters, which show a range between 0.01 – 0.29 m^2/s, where 87% of the high confidence clusters are below 0.1 m^2/s. Due the majority of earthquakes being located in the crystalline basement (McNamara et al., 2015a), the diffusion is likely showing fluid pressure movement through basement faults and fractures.

The interquartile range of high confidence diffusivities, D_{25th}=0.018 m^2/s and , D_{75th}=0.067 m^2/s, can be used to estimate the fracture permeability of the basement fault

zones. Using the method of Talwani et al. (2007) the fracture permeability (k) can be approximated by:

$$k = D\mu[\Phi\beta_f + (1 - \beta_r)\Phi] \quad (14)$$

where D is the hydraulic diffusivity, μ is the viscosity of the water, Φ is the porosity of the host rock, and β_f and β_r are the compressibilities of the fluid and host rock, respectively (Bodvarsson, 1970). The average earthquake depth in Oklahoma is approximately 5 km. Assuming a geothermal gradient of 25.5 ºC/km, or 1.4º F/100 feet (Harrison et al., 1983), and a surface temperature of 20 ºC results in a temperature around 145 ºC for the average location of earthquakes. This corresponds to a viscosity of 0.190 x 10^{-3} Pa*s (TET, 2017) and a density of approximately 1000 kg/m^3. The remaining parameters were based on published estimates from Talwani et al. (1999) of Φ = 3 x 10^{-3}, β_f = 10^{-10} Pa^{-1}, and β_r = 2 x 10^{-11} Pa^{-1}. The 25th and 75th percentile of high confidence diffusion rates, D_{25th}=0.018 m^2/s and , D_{75th}=0.067 m^2/s, results in permeability range of 0.08–0.29 mD (or 0.785–2.920 x 10^{-14} m^2). This range is an order of magnitude lower than the range of compiled 'seismogenic permeabilities', a range of permeabilities in which most documented cases of fluid pressure diffusion triggering seismicity occur (Talwani et al., 2007), but still within the range of estimated crystalline basement permeabilities (Clauser, 1991). These diffusion rates support the notion of uncharacteristically low permeability basement faults exhibiting pore pressure diffusion triggering in Oklahoma. This could potentially be due to the scale of injected fluid and corresponding fluid pressures in these basement fractures being larger than in other studies of induced seismicity. However, in the analysis of seismogenic permeabilities in Talwani et al. (2007), it was hypothesized that if there were greater pressures

permeating permeabilities below the range of seismogenic permeabilities, like observed here, then the pressure should open the fractures and increase the permeability into the range of seismogenic permeabilities.

The direction of unilateral migration vectors shows spatial variations in migration orientations. Although many of the clusters appear to migrate away from regions of greater volume injection, the trend is not consistent across the state. At the Prague cluster the migration is moving to the north, seemingly away from the nearby injection high (Figure 42). In the western seismic region, the majority of clusters have migration to the west, despite for some clusters this being in the direction of the main injection high in the region (Figure 42). The variability is possibly tied to the structural dip of the basement, which dips to the southwest part of the state, and faults acting as fluid conduits and barriers, preventing a homogenous diffusion. It is also possible the discrepancy is due to triggering from smaller volume wells that are near the clusters but are not highlighted by the contours.

Large scale diffusion patterns

The large scale diffusion fitting associates events with seven high volume injection sites, regions a, b, c, d, e, f, and g, and determines the start date and diffusivity to best fit the migration pattern from each fluid-source points (Figure 43). Of the seven regions, five of the regions (a, c, e, f, and g) show large scale diffusion above 85% confidence (Figure 44, 45). Because the original diffusion curve fitting method was made for single clusters that spanned relatively short time periods, the application at this large scale is less stable. The locations chosen, although restricted to the 95th percent contour interval, represent the best locations that visually showed a diffusion pattern

and the confidence is relatively unstable for two of the five high confidence regions (regions c and e) while the other three high confidence regions (a, f, and g) consistently show a high confidence diffusion pattern.

The high confidence fits were matched with diffusivities between 0.75 m^2/s and 2.25 m^2/s, which are an order of magnitude greater than the cluster-based diffusion. The low confidence diffusion regions show no trend in the data, this is likely due to multiple large fluid sources in the area, implying smaller scale diffusion triggering, or the location of the seeds are not representing the actual fluid source. Not all the high confidence fluid-sources diffusion fits have events close in both space and time with the start point (Figure 44 a, 45 e and f). This is likely because the method treats diffusion of fluid pressure from a region of high injection as a single point source for spherical diffusion, while in reality the diffusion front would be significantly more complex. Despite this lack of early events near the fluid source, the large scale diffusion pattern is quite apparent in those five regions.

Support for large-scale fluid pressure diffusion comes from the previous analysis. Each fluid-source point is associated to events up to 50 km away, which is close to the spatial analysis results of 40 km minimum distance to observe a stable volume-seismicity scaling relationship. The earlier temporal analysis saw high correlations with association distances around 20 km (0.3 degree grids) for the central parts of the pressure regions, represented primarily by Figure 44 b and d. However, the edges of the seismic regions, such as around the Fairview, Pawnee, and Prague clusters, showed low correlations at the small grid size but not at the large grid size, supporting the idea that the edges of the pressure regions are being triggered by far field

pressurization. Similar findings also come from pressure modeling of the Jones swarm (Keranen et al., 2014) and the western seismic region (Goebel et al., 2017) where both studies found large-scale pressure diffusion triggering from areas of high injection.

The higher diffusion rates of the large-scale diffusion are also more in line with the permeability range of the Arbuckle Group. Following the methods of Goebel et al. (2017), the permeability k can be estimated from the diffusivity D by:

$$k = \frac{\mu D S_s}{\rho_f g} \quad (15)$$

where μ and ρ_f are the viscosity and density of the fluid, g is the acceleration due to gravity, and S_s is the specific storage (Wang, 2000). This equation estimates permeability based on radial diffusion through a porous medium, while equation (14) was based on diffusion through fractures. To test the magnitude of possible permeabilities, typical values for water-filled sandstone/limestone makeup like the Arbuckle are used. The average Arbuckle disposal well reaches depths of approximately 2 km. Again, assuming a geothermal gradient of 25.5 ºC/km, or 1.4º F/100 feet for Oklahoma (Harrison et al., 1983), and a surface temperature of 20 ºC results in an average temperature around 70 ºC in the Arbuckle. This corresponds to a viscosity of 0.404 x 10^{-3} Pa*s (TET, 2017) and a density of approximately 1000 kg/m^3. A typical specific storage for a sandstone/limestone formation is around S_s = 10^{-6} m^{-1} (Wang, 2000). Testing the average diffusivity of the regional diffusion, 1.8 m^2/s, results in permeability of 74 mD (or 7.4 x 10^{-14} m^2). This value is in good agreement with the median and 75th percentile horizontal permeabilities of 6.62 and 18.83 mD for core measurements of the upper Arbuckle and median and 75th percentile horizontal

permeabilities of 20.34 and 55.86 mD for the lower Arbuckle (Morgan and Murray, 2015).

Conclusions

About 52% of large seismic clusters show the signature of pore pressure being the dominant triggering mechanism with at least 85% confidence. Clusters that don't have high confidence either don't have a clear diffusion pattern due to mainshock/aftershock, or maybe triggered by an alternative mechanism. However, the majority of both diffusive-migration clusters and all clusters have diffusivities under 0.1 m^2/s and similar diffusivity distributions, supporting consistent heterogeneities in the basement diffusivities. The large scale diffusion curve fitting shows five high volume regions that are potential fluid sources for far field triggering. These five injection highs are located near the edges of seismically active regions, while seismicity within the seismic regions shows no clear large scale diffusion trend and is likely associated with multiple, smaller scale high volume regions. The large scale diffusion has best fitting diffusivities between 0.75 and 2.25 m^2/s. Assuming some simple Arbuckle fluid properties this equates to a permeability around 74 mD. Because the commonly high permeabilities in the Arbuckle and extensive injection interval of disposal wells it seems likely the large scale diffusion is showing the movement of fluid pressure through the Arbuckle, while the cluster diffusion is showing the movement of fluid pressure within the basement fractures and faults.

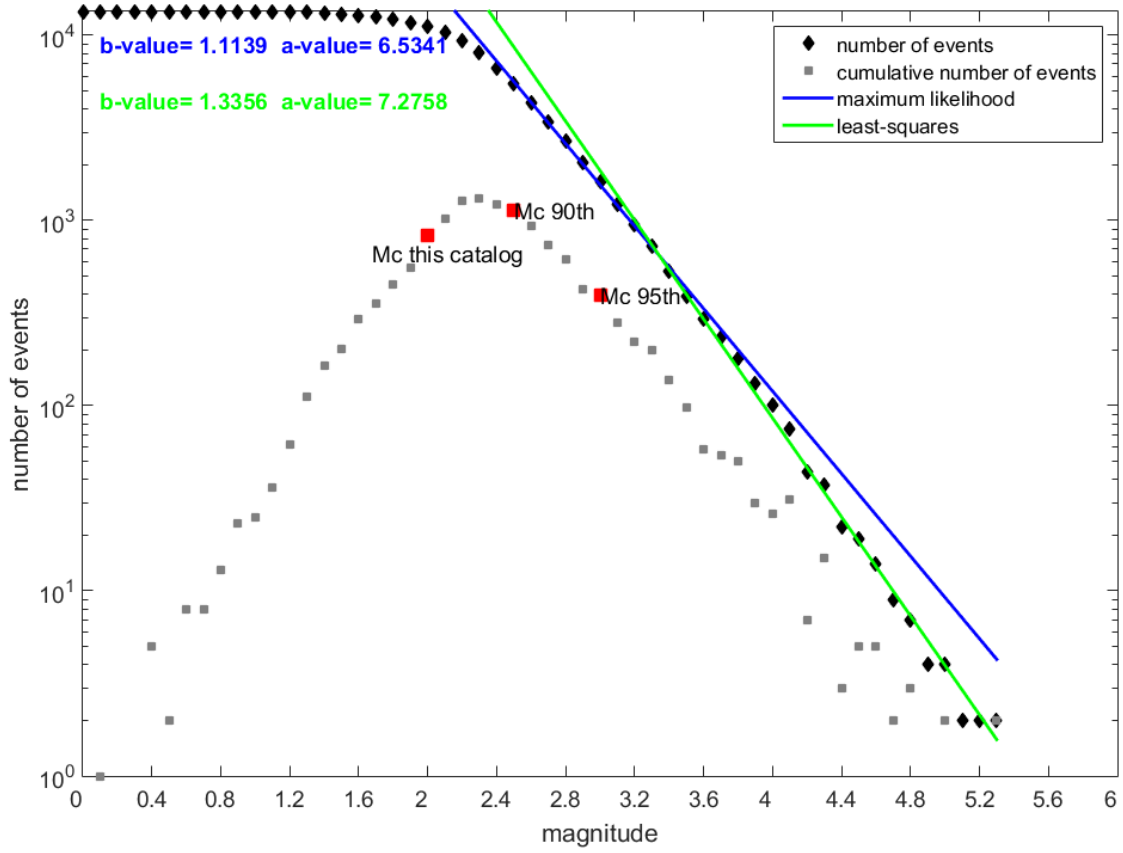

Figure 34. Cumulative magnitude frequency plot of the relocated OGS catalog. The 90th and 95th percentile goodness of fit Mc estimates are larger than the maximum curvature of about M = 2.1. An Mc of 2.5 is used in this study, despite being clearly too low to accurately represent a Gutenberg-Richter distribution across the state, to increase the number of viable events for the diffusion fitting.

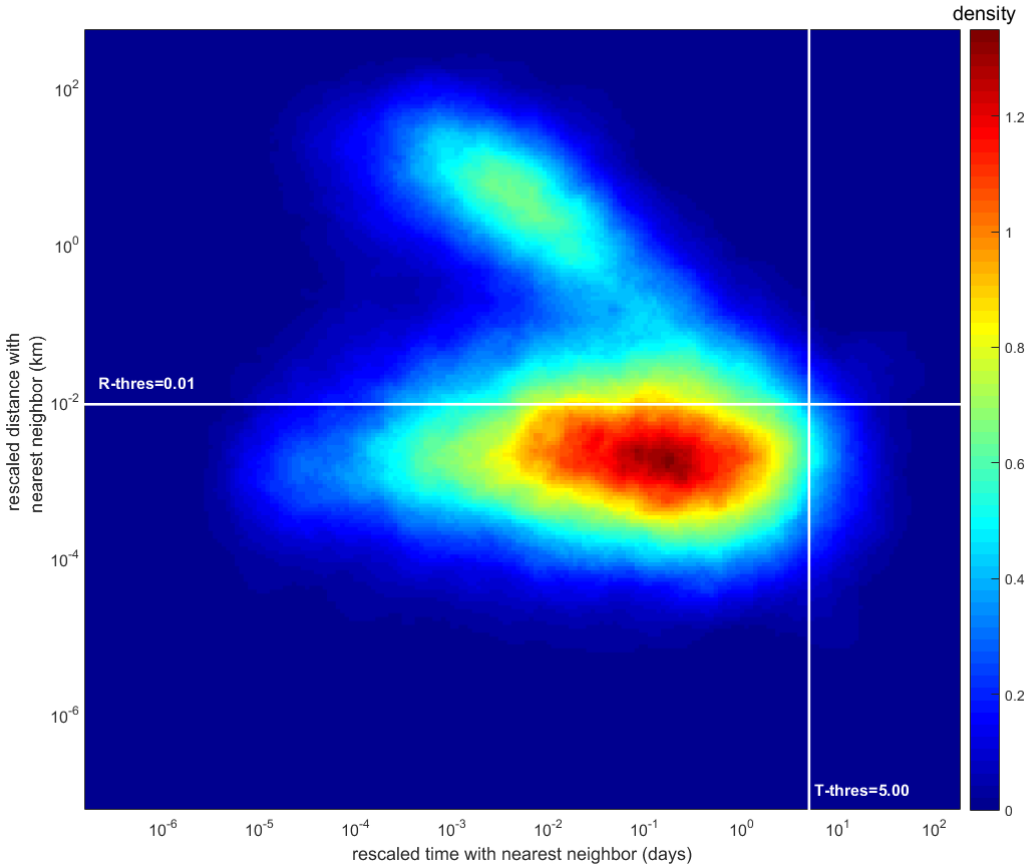

Figure 35. Heat map of rescaled nearest neighbor time and distance. A weak grouping of further distance nearest neighbors may represent background seismicity. The time and distance thresholds dictate the interevent distances that are considered a clustering relationship.

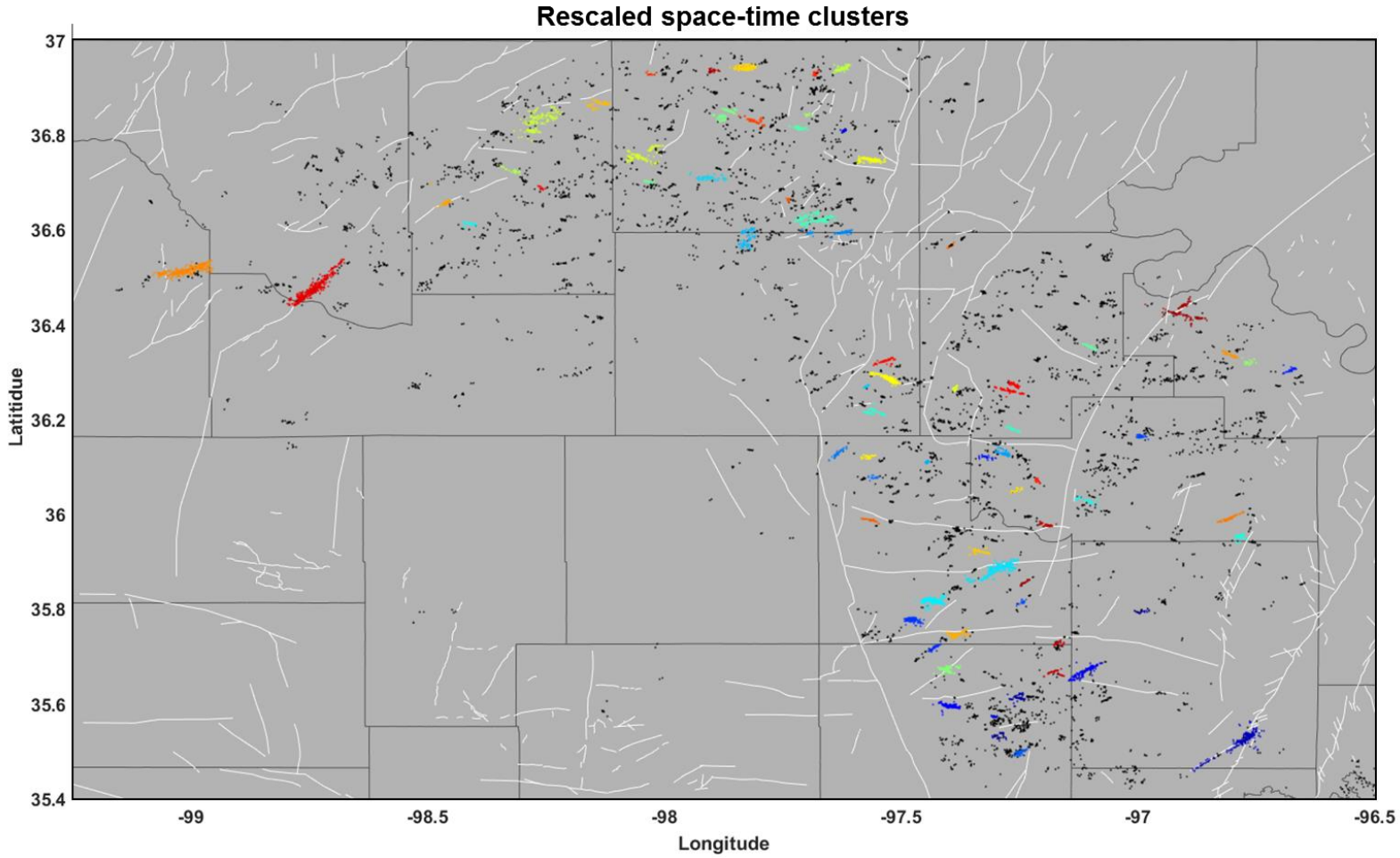

Figure 36. Rescaled space-time clusters (colored dots), non-clustered earthquakes (black dots), known sedimentary faults (white), and county lines (black). Although there are still visible clusters of seismicity, the interevent distance clustering method clustered all of the larger seismic clusters that can be used for the diffusion analysis.

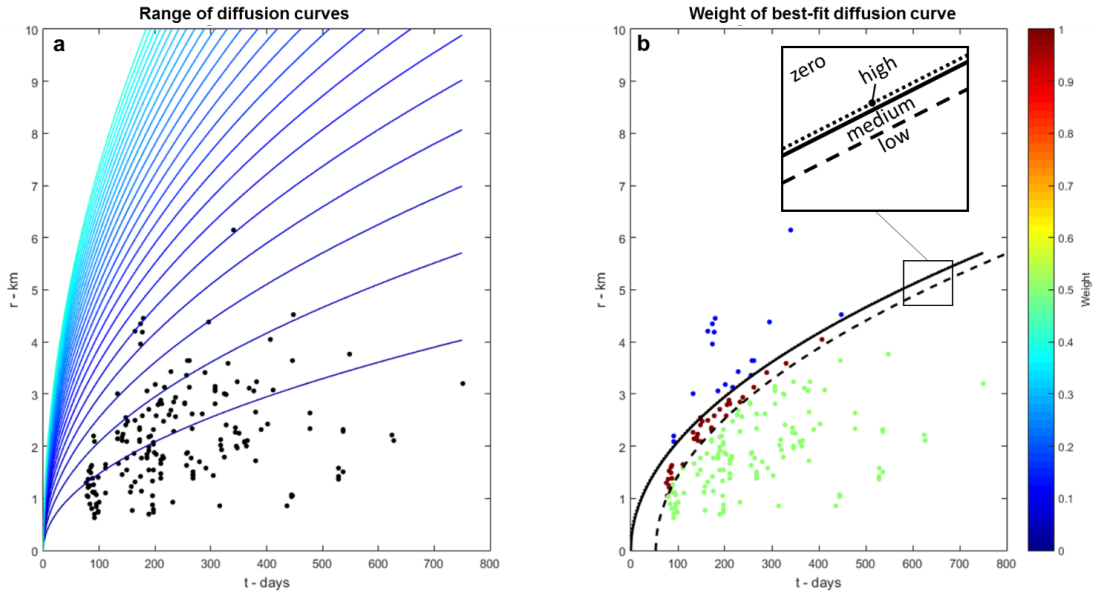

Figure 37. Example of a range of diffusivity curves applied to a diffusion pattern of events spatial migrating away from the first event over time (**a**). For the same cluster, the best fitting diffusion curve with events color coded by the weights applied to each event to calculate the total misfit (**b**). The width of the high weight curve is based on the number of events with negative misfit (to the left of the curve) and becomes larger if the diffusivity is too low. The weighting scheme results in the tightest fitting curve that allows a maximum of around 10% of the events to have negative misfit.

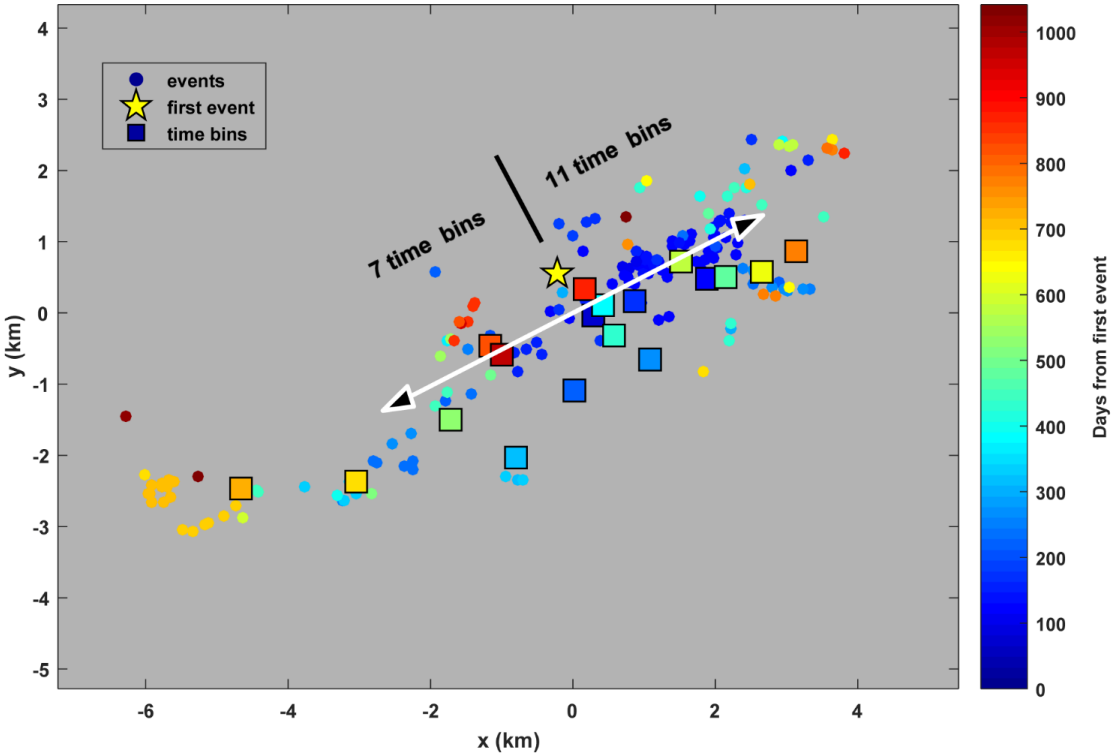

Figure 38. Example of migration vector calculation. Events (circles) and the average location of time bins (squares) are color coded by their time after the first event (star). If over 2/3 of the time bins are on one side of the strike vector, it is unilateral migration. In this case there are 7 time bins on one side and 11 time bins on the other, so the cluster is categorized as bilateral migration.

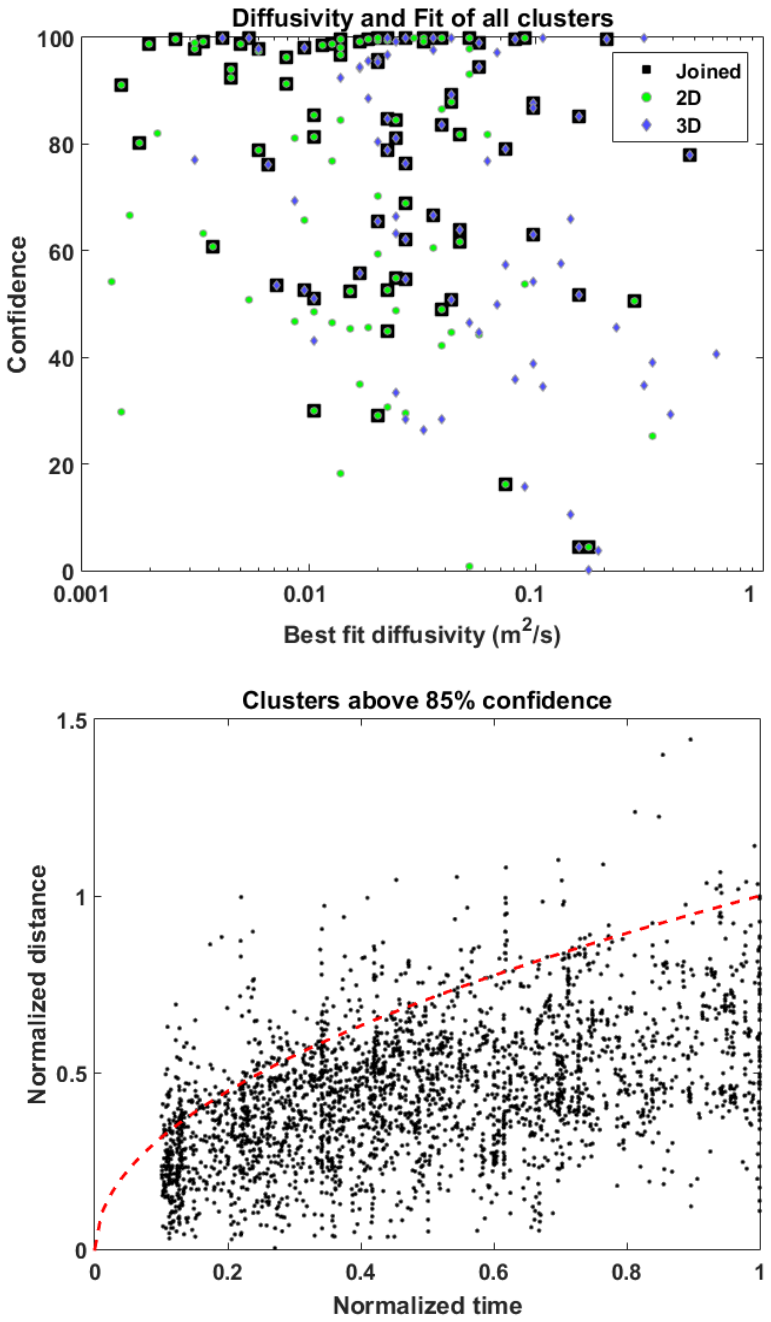

Figure 39. Diffusivity and fit confidence of the 2D distance, 3D distance, and joined runs (top). The majority of high and low fitting clusters have a best fitting diffusivity below 0.1 m^2/s. Normalized diffusion fit of all clusters above 85% fit (bottom). Events in the first 10% of each clusters time period was removed to prevent false-positive fits. The overall fits are good, with some mainshock/aftershock spikes present in the data appearing outside the diffusion curve.

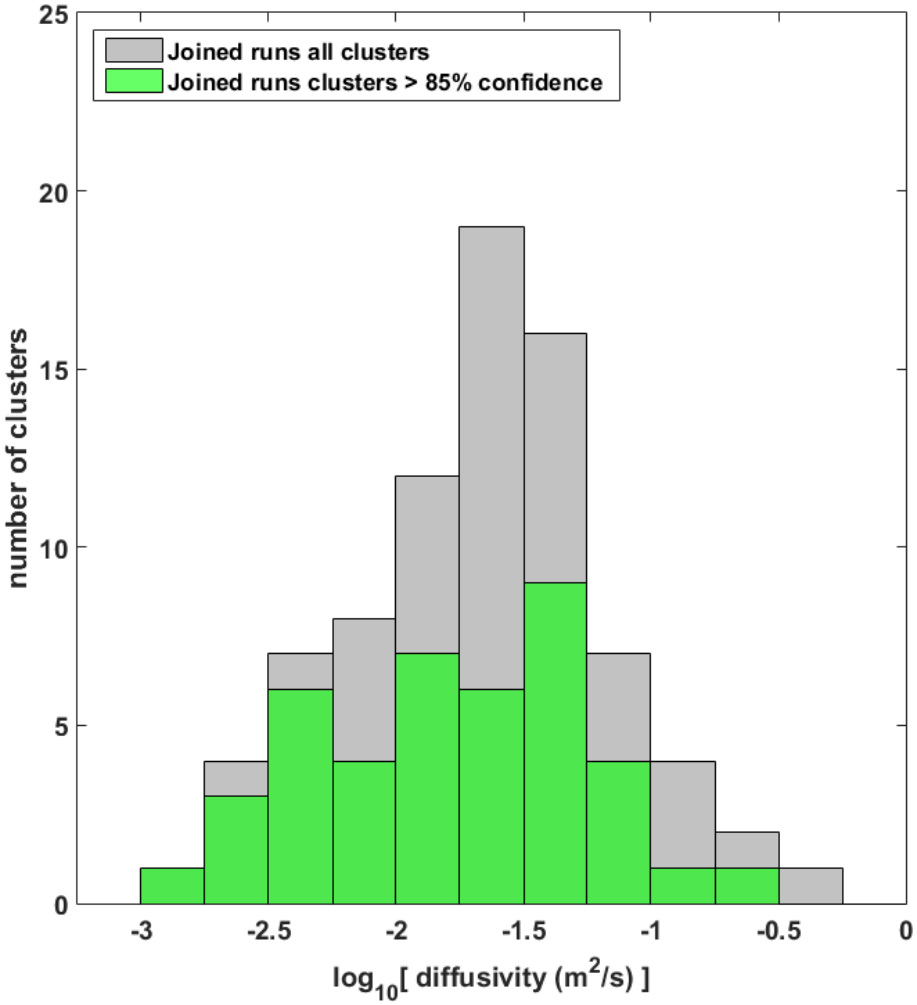

Figure 40. Histogram of best fitting diffusion rates for all cluster and high confidence clusters. The histogram is plotted along a logarithmic scale and shows a normal distribution to both the high confidence and all clusters diffusion rates.

Table 1. Statistics of diffusion curve fitting for 2-D, 3-D, and joined runs

	Above 85%			All		
	Mean	Std	Range	Mean	Std	Range
2-D	0.02	0.02	0.09	0.03	0.05	0.33
3-D	0.05	0.07	0.29	0.08	0.11	0.62
Joined	0.03	0.04	0.20	0.04	0.07	0.47

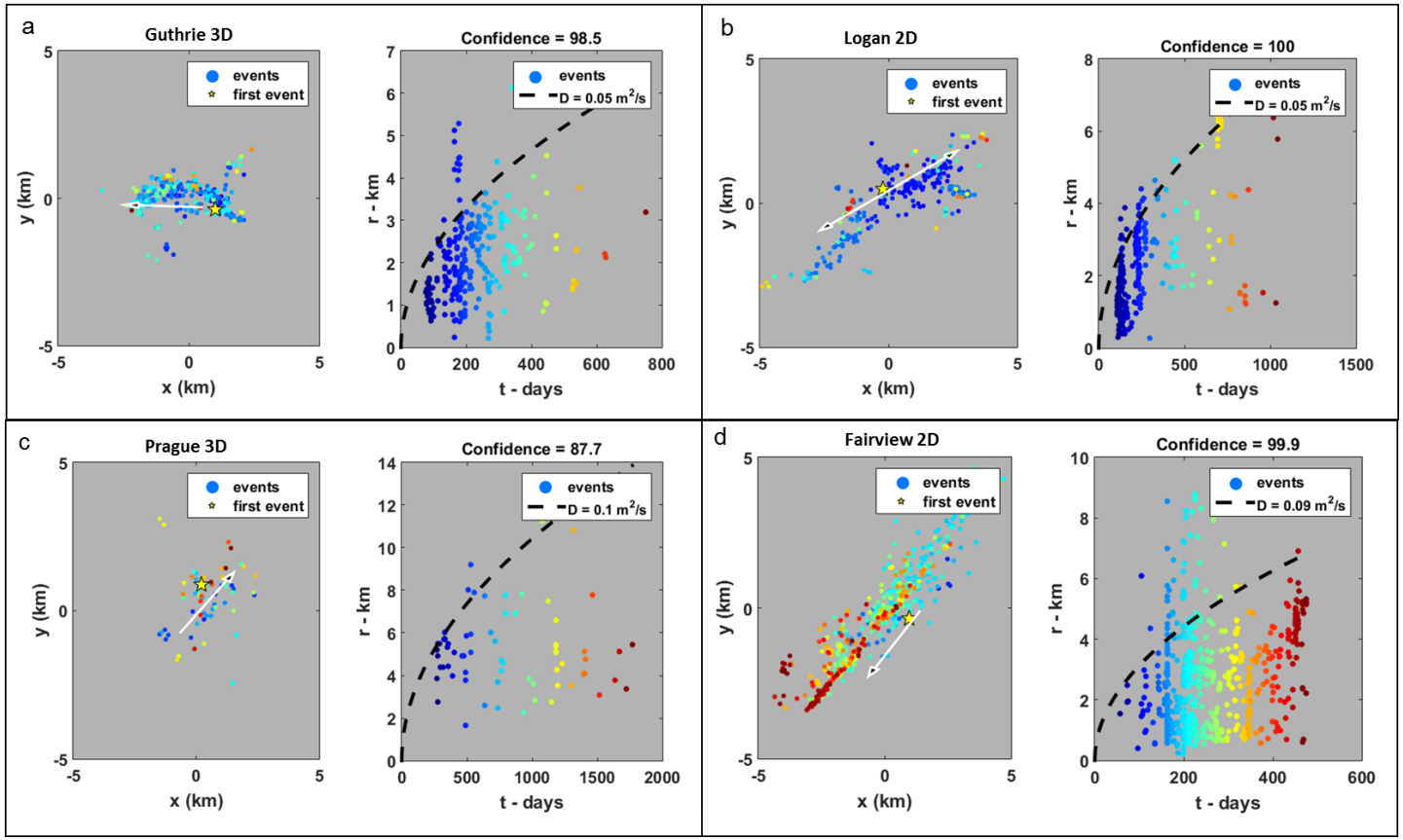

Figure 41. Map view of events color coded by time of occurrence relative to first event with arrow showing unilateral or bilateral migration direction (insert left) and best fitting diffusion curve for the spatial migration of events away from the first event over time (insert right) for Guthrie (**a**), Logan (**b**), Prague (**c**), and Fairview (**d**) clusters. All four clusters have a high confidence fit to the diffusion pattern. Guthrie and Prague are shown with the fit to the 3D distance data, while Logan and Fairview had a higher confidence fit with the 2D distance data.

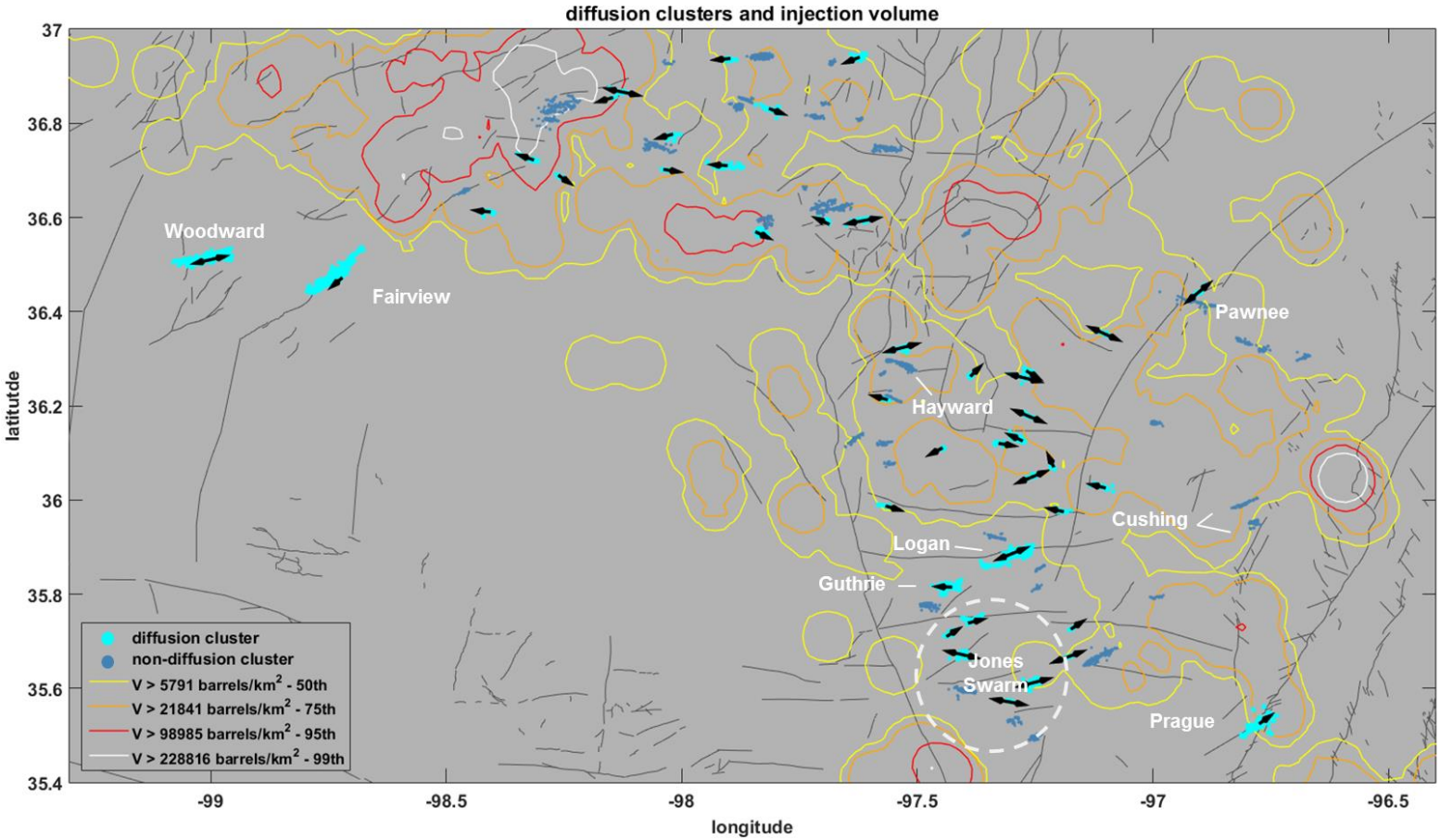

Figure 42. Map view of diffusion and non-diffusion clusters based on 85% confidence threshold, migration direction arrows on diffusion clusters, contour lines of a range of disposal volume percentiles, and known faults (grey lines). Some well-studied clusters such as the Fairview and Prague sequences show diffusion migration while others such as the two Cushing sequences do not. Although in many areas the migration direction appears to be moving away from injection highs, the trend is not consistent across the entire study area.

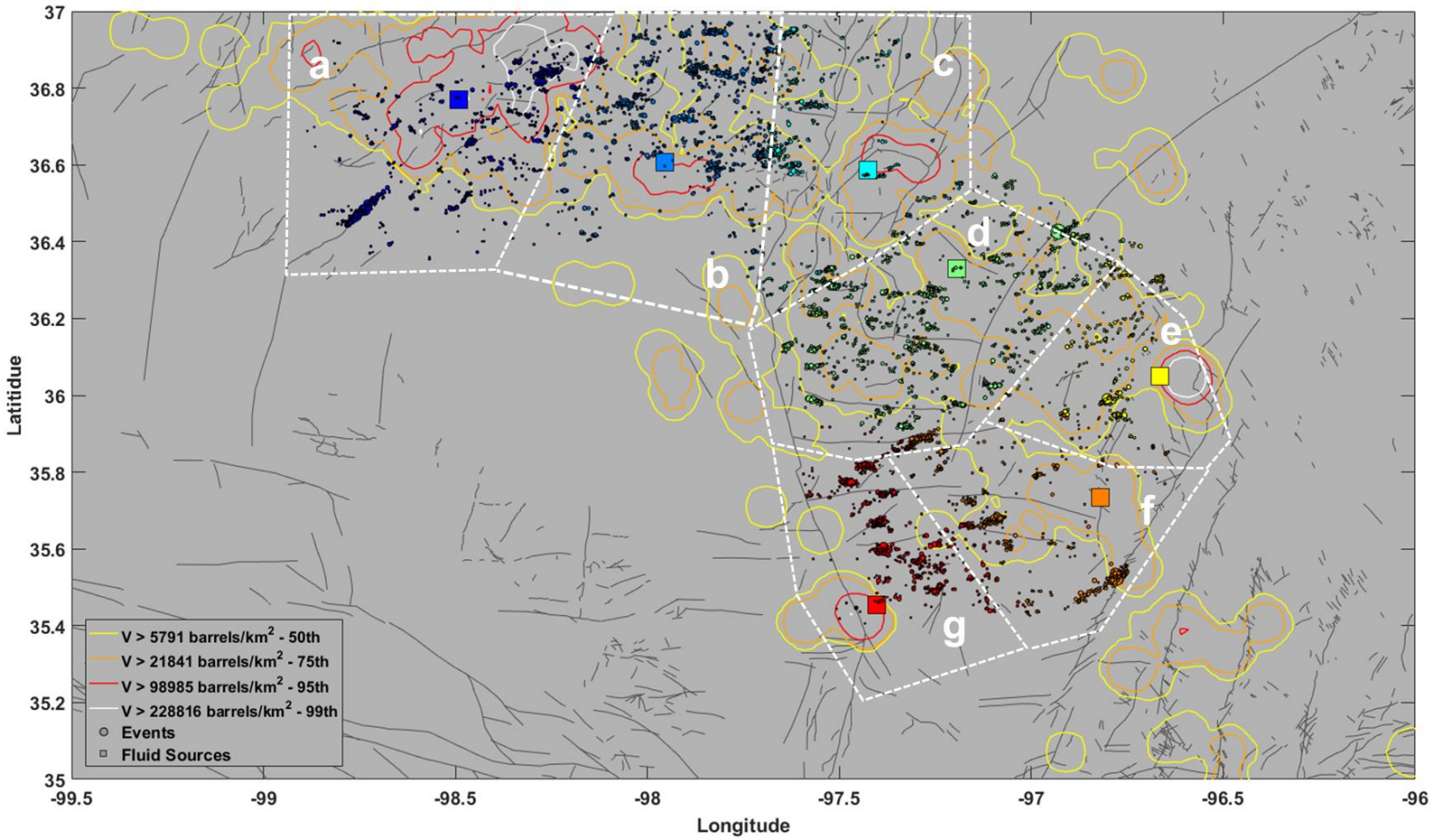

Figure 43. Map view of events within the main pressure regions, minus the Woodward cluster, color coded by the nearest fluid-source seed point as well as contour lines representing a range of total injection percentiles and known faults (grey lines). Based on this division of seismicity, the Guthrie clusters is tied to the red seed (**g**), the Prague cluster to the orange seed (**f**), and the Fairview cluster to the dark blue seed (**a**).

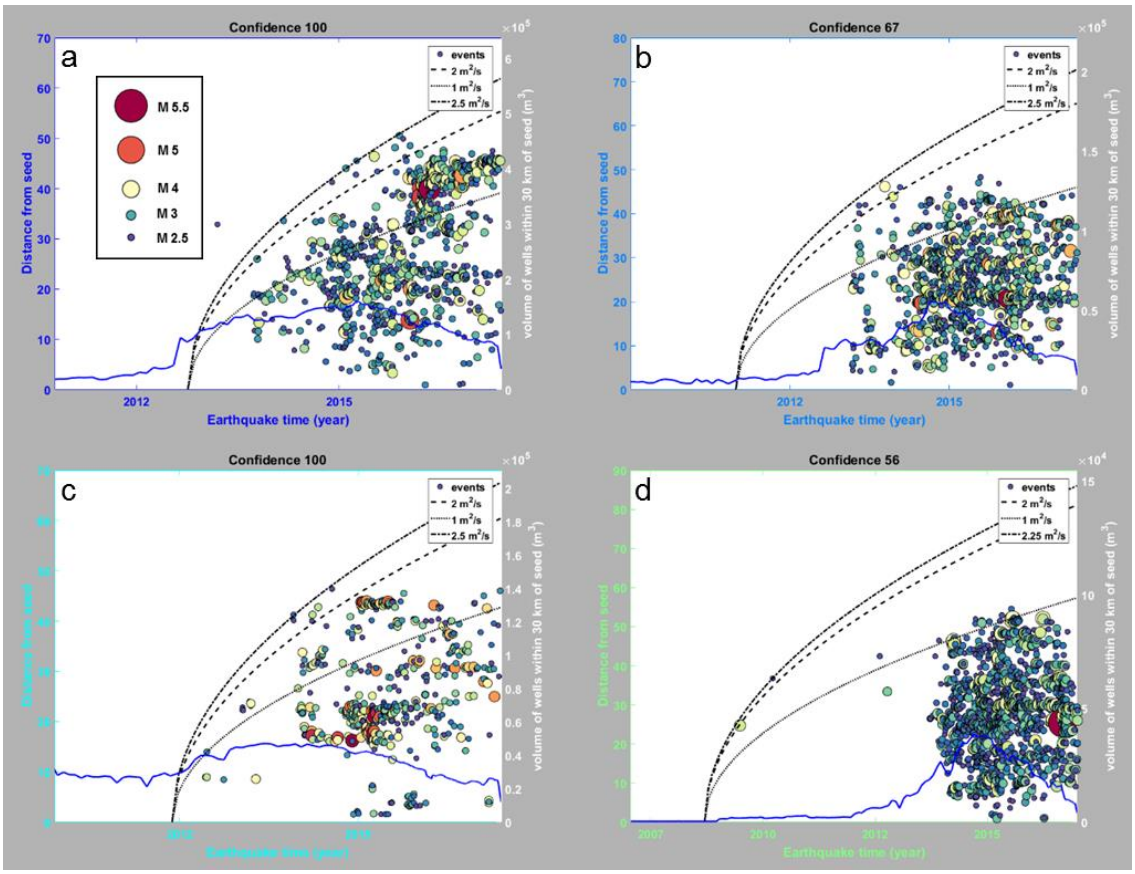

Figure 44. Diffusion pattern for the events associated with four of the seven seed points (continues on next page). The axis color corresponds to the seed point color. For each plot there are the best fitting diffusivity starting from the start time with the least misfit, as well as a 1 and 2 m^2/s diffusion curve for consistency across plots. Also on each plot is the daily disposal volume injected within 30 km of each seed point. The high confidence fits (**a**,**c**) appear to show large scale diffusion out to 50 km. Many of the high confidence fits have minimum misfit start dates very close to periods of increased injection volumes.

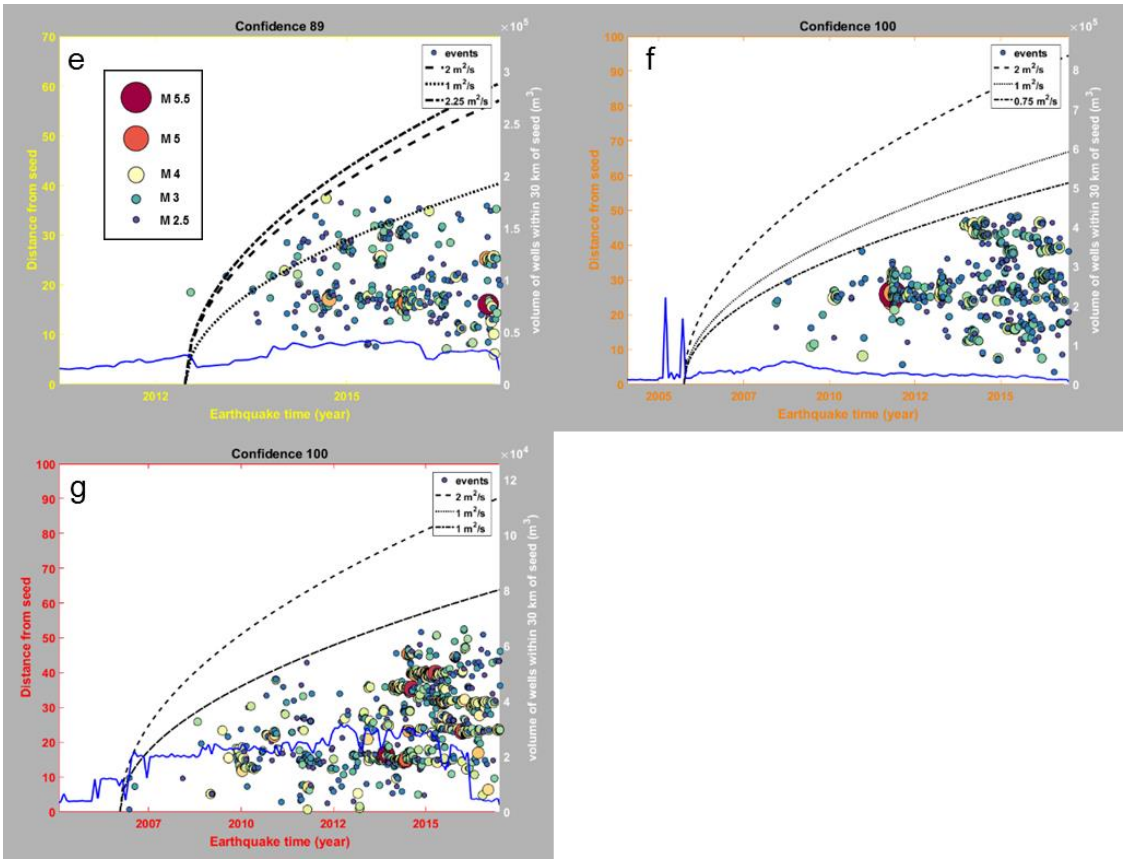

Figure 45. Diffusion pattern for the events associated with three of the seven seed points (continues on previous page). The axis color corresponds to the seed point color. For each plot there are the best fitting diffusivity starting from the start time with the least misfit, as well as a 1 and 2 m^2/s diffusion curve for consistency across plots. Also on each plot is the daily disposal volume injected within 30 km of each seed point. The high confidence fits (**e**,**f**,**g**) appear to show large scale diffusion out to 50 km. Many of the high confidence fits have minimum misfit start dates very close to periods of increased injection volumes.

Chapter 5: Seismogenic Fault Analysis

Introduction

An important part of reducing the induced seismicity in the state has been preventing disposal wells from injecting near known faults. One of the earliest steps in the OCC's actions to reduce seismicity was the traffic light system that went into effect in 2013 (OCC, 2015). In that system, any new disposal well required extra review to be placed within 3 miles of a stressed fault (OCC, 2015). Information on stressed faults at that time came from compiled fault data of the Oklahoma Geological Survey (OGS) and the regional state of stress derived from 152 focal mechanism solutions (Holland, 2013b). Based on the regional stress field, different fault orientations were assigned risk levels of failing, with faults striking between 40º–60º and 130º–150º considers optimally oriented for failure (Holland, 2013b). An updated optimally oriented fault assessment using 688 focal mechanism solutions from 2010–2015 data assigned optimally orientation to faults striking 40º–60º, 105º–120º and 135º–150º E and a mean and median maximum horizontal stress direction around 84º E (Darold and Holland, 2015). More recent work has incorporated well bore stress data with focal mechanism stress orientations and found a similar range of maximum horizontal stress directions between 80º–90º E (Alt and Zoback, 2017).

Although the fault maps and associated hazard appear valuable to regulators, the locations of the vast majority of clustered seismic activity are not aligned with known faults. This is partially due to mapped faults being reported are mostly from the sedimentary section, while the recent seismicity has been on shallow basement, likely re-activated crustal faults that extend into the Arbuckle Group (McNamara et al.,

2015a). Furthermore, many of the focal mechanism orientations are more similar to faulting in southwest Oklahoma than the mapped faults of the seismically active regions (Qi, 2016). In this chapter I compare seismogenic faults based on the earlier clustering analysis and compare these faults orientations to the known sedimentary faults as well as published, optimally orientated fault orientations. I hypothesize that although focal mechanisms show consistent orientations in the state, the fault zones represented by seismic clusters have different orientations and assigning a hazard level to known faults may underestimate the hazard of sub-optimally oriented faults.

Method

From the earlier cluster analysis there were 89 clusters with at least 20 events after grouping clusters by nearest neighbor rescaled distance and separating clearly delineated faults within each cluster. Aftershocks highlight the primary fault plane and were kept in this chapter. Next, the three eigenvalues and corresponding eigenvectors were calculated for the covariance matrix of each clusters demeaned earthquake location after the methodology of Vidale and Shearer (2006). The eigenvector of the largest eigenvalue represents the strike direction of the cluster, while the dip and planarity can be calculated with the addition of the other two eigenvectors (Vidale and Shearer, 2006). The planarity is a measurement of how planar a cluster is, where regardless of length, a seismic cluster with a planarity of 0 would have near identical width and height (Vidale and Shearer, 2006).

For each pressure region (Figure 46), the seismogenic fault azimuths represented by the clusters' strikes were compared to the mapped faults near the clusters from the OGS preliminary fault map (Holland, 2015). The seismogenic faults were also

compared to the Frontal Wichita fault system in southwestern Oklahoma that is related to the state scale deformation of the Anadarko Basin and Wichita-Amarillo uplift (Crone and Luza, 1990). Each fault group is then linearized and the strikes of individual faults were calculated using FracPaQ software (Healy et al., 2017). The FracPaQ software uses a Hough transform method to find linear patterns in the fault image and also allows the rejection of short fault segments (Healy et al., 2017). An example of the original fault geometries and the linearized result of FracPaQ can be seen in Figure 47. The strikes of the mapped and seismogenic faults in each pressure region were then visualized with an area-weighted rose diagram, which gives a more accurate representation of azimuthal data (Nemec, 1988).

Results and Discussion

The seismogenic faults in Oklahoma are primarily oriented east-west with strikes ranging between 30° and 120° E (Figure 48). The seismogenic faults are also for the most part steeply dipping and planar (Figure 48). The clusters showing low planarity are likely multiple faults that were not visually separable and as such likely represent a group of synthetic faulting. Although the steep dip of the clusters agrees well with moment tensor studies (Darold and Holland, 2015; McNamara et al., 2015a), the strike of the seismogenic faults do not. McNamara et al. (2015a) found very few focal mechanism solutions with a 75° to 120° E or the reciprocal 225° to 300° E azimuth, which is the strike direction of over half of the seismogenic faults. Darold and Holland (2015) had a higher probability for events in the 75° and 120° E range but had numerous solutions between 135° and 150° E, which does not match any seismogenic faults, likely represents secondary nodal planes. Although some of the seismic clusters

striking NEE and SEE (the orientations McNamara et al. (2015a) and Darold and Holland (2015) show no focal mechanisms) may be aligned synthetic faulting, multiple clusters show sharply delineated faults in those orientations (Figure 49).

These discrepancies between focal mechanism solutions and seismogenic fault orientations likely stems from the over representation of large clusters. Moment tensor solutions represent individual earthquakes, and because of this represent larger clusters more than small clusters. This is reflected in the fault risk classification ranges of Darold and Holland (2015) applied to the seismogenic faults. Using said classification there are 12 seismogenic faults with sub-optimal orientations, 42 with moderately optimal orientations, and 35 with optimal fault orientations based on the focal mechanism solutions. The failure of this risk assessment is also visible in the case of the 2016 M5.8 Pawnee earthquake. In that case, the mainshock and majority of aftershocks resolved a fault striking between 107º E, where the mainshock occurred, to 98º E, where the end of the 6km fault depicted by aftershocks ends (Chen et al., 2017, under review). This means the largest instrumented earthquake in Oklahoma's history occurred in only a moderately-optimal fault orientation. The problem with assigning a risk to known faults based on focal mechanism solutions is that the focal mechanisms preferentially represent larger seismic clusters.

The orientation of the seismogenic faults represented by seismic clusters in the western seismic region is fairly dissimilar to the strike of the known faults in the area (Figure 50). The mapped faults primarily strike anywhere from 0° and 105° E with a dominant orientation of 60° E while a third of the seismogenic faults strike between 105° and 135° E. Notably, seismogenic faults at this orientation share a similar

orientation to the west-northwest faulting in southwest Oklahoma. The eastern pressure region shows similar trends (Figure 51). A large percentage of the seismogenic faults strike between 105° and 135° E but very few of the mapped faults share that orientation. As noted by Qi (2016), this discrepancy is also present in the orientation of mapped faults compared to moment tensor solutions. It therefore seems possible that these clusters represent reactivated faulting tied to the Frontal Wichita fault system of southwestern Oklahoma that formed during Paleozoic tectonism (Crone and Luza, 1990). This means earthquakes are not only occurring on faults not visible on known fault maps, but also they are often occurring on faults that are part of a separate structural regime than the mapped faults.

Conclusions

The seismogenic faults in Oklahoma strike primarily between 30° and 120° E with dominant trends of 60° and 110° E and show significant differences to the orientation of focal mechanism solutions in that state. This discrepancy is likely due to focal mechanism's preferential representation of large clusters, which underestimates the hazard on faults oriented in the direction of smaller clusters. Furthermore, most of the seismicity occurs on seismogenic faults not represented on fault maps. The differences in seismogenic fault and mapped fault orientations in the western and eastern pressure region suggest many of the re-activated seismogenic faults are tied to the Frontal Wichita fault system of southwest Oklahoma. Not only does assigning risk to known faults using focal mechanism solutions underestimates the chance of failure on sub-optimally oriented faults, it also allows an assumption of safety when avoiding

mapped fault even though most seismicity occurs along seismogenic faults and many show an orientation of a different structural regime.

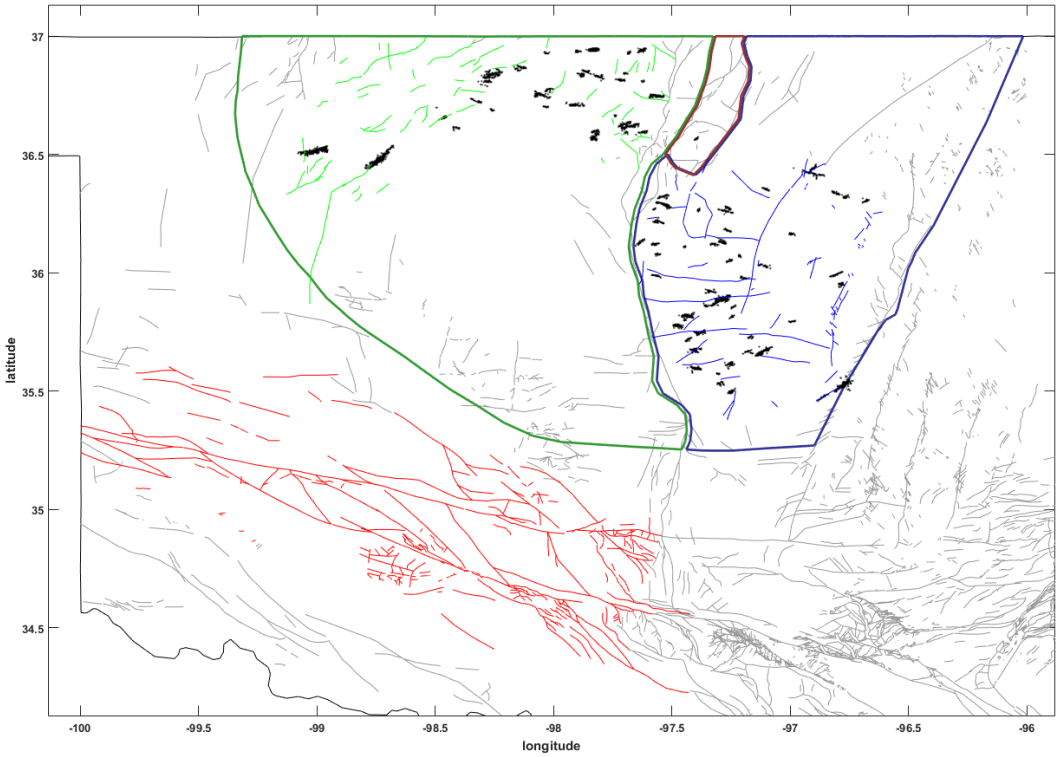

Figure 46. The western (green) and eastern (blue) pressure regions with seismic clusters (black dots) and faults. Faults near the clusters in each region are color coded as well as the basement faults in southwestern Oklahoma.

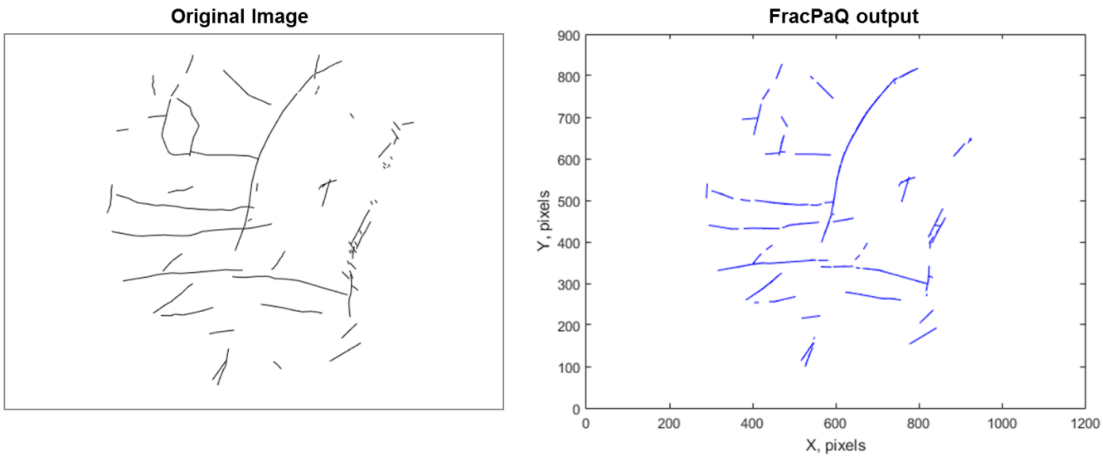

Figure 47. Original image of the faults near clusters in the eastern pressure region as well as the linearized image after the Hough transform in FracPaQ (Healy et al., 2017). The output fits straight lines to curved segments and removes faults below a certain size.

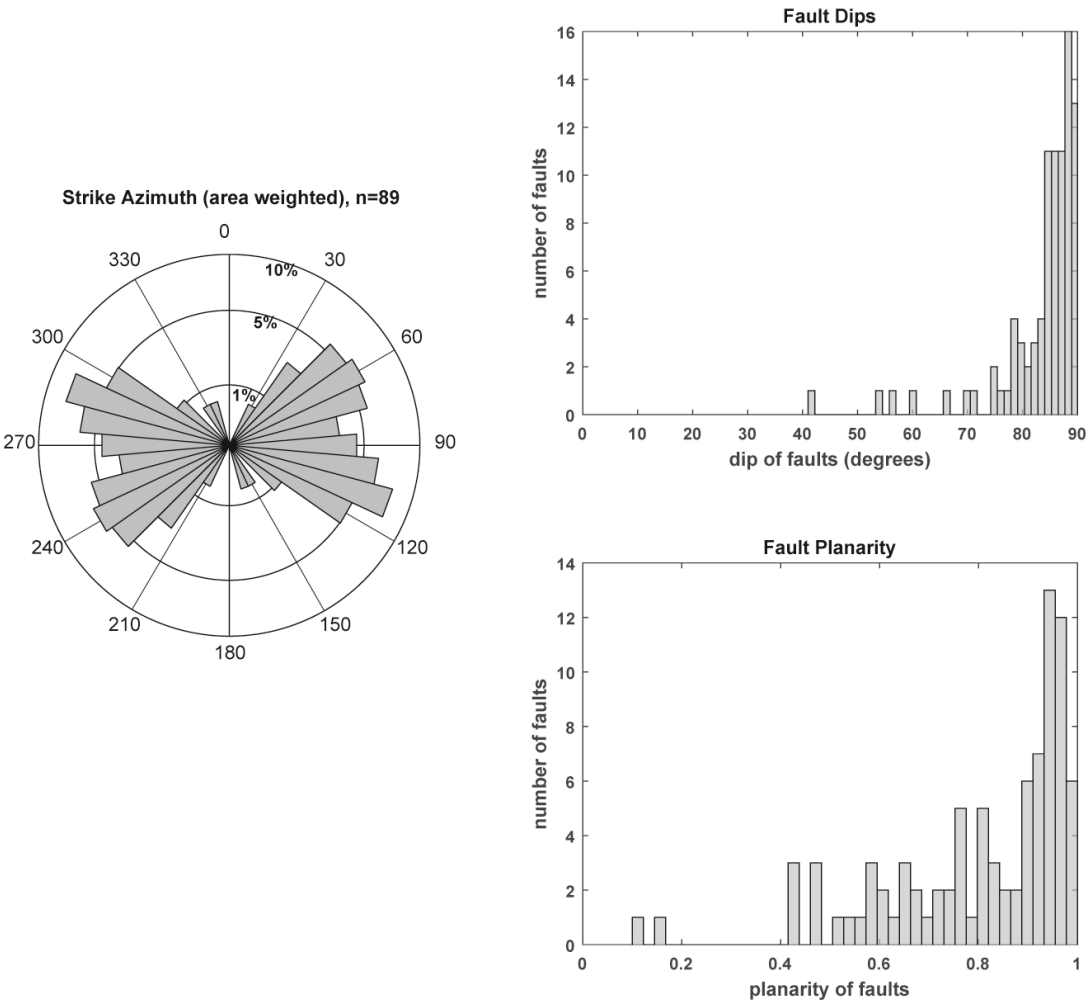

Figure 48. The strike azimuth (left), dip (top right), and planarity (bottom right) of the 89 seismogenic faults. The majority of faults are striking between 30° and 120° E, and are steeply dipping. The faults with low planarity are likely multiple faults that were not well enough resolved to separate.

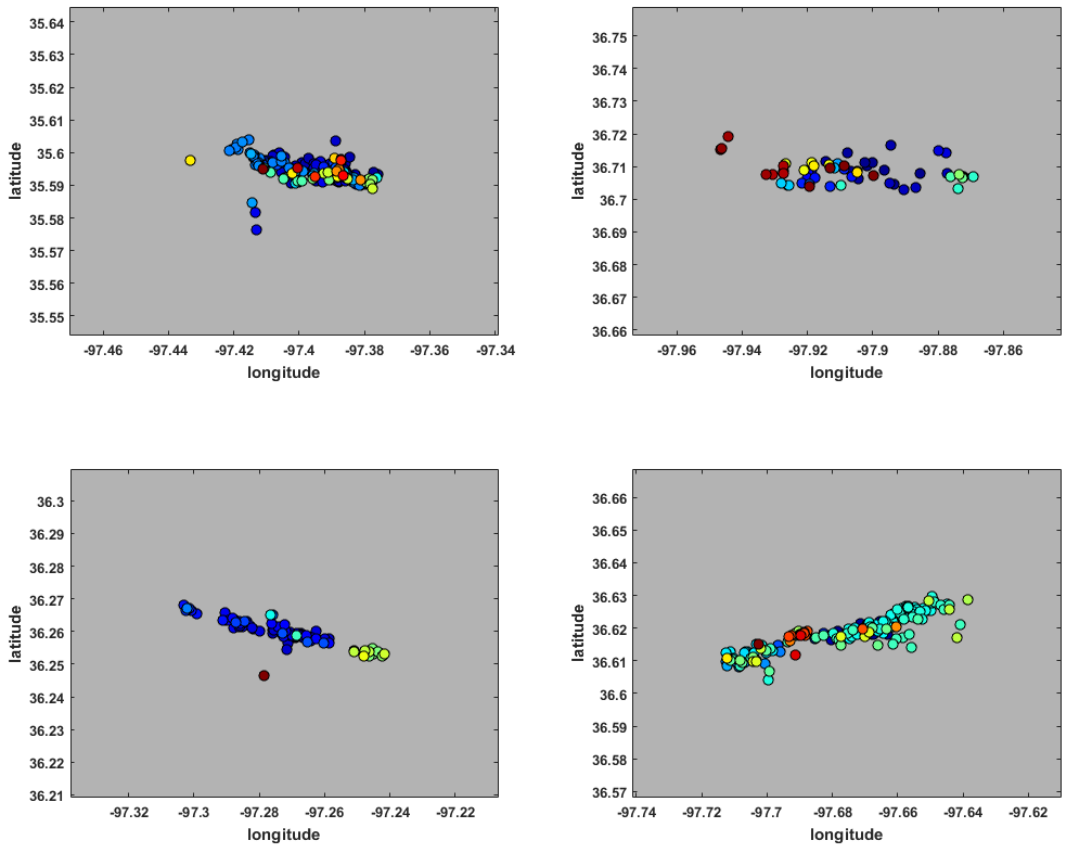

Figure 49. Example of four approximately east west striking clusters with sharply delineated faults. All four figures have the same spatial scale of around 15km east west and north south.

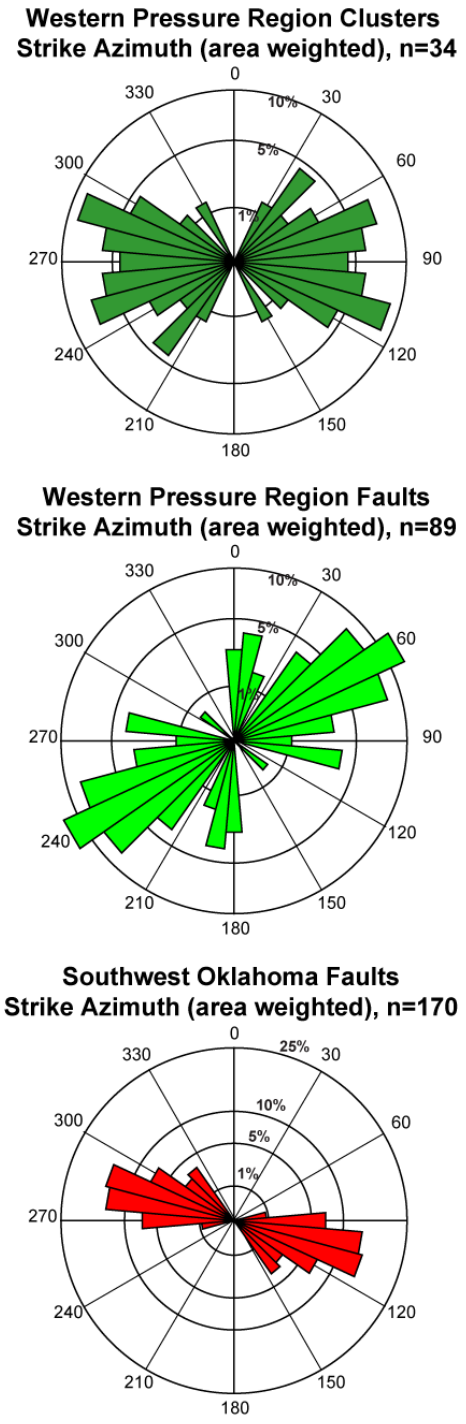

Figure 50. Rose diagram for the seismogenic faults (top) and mapped faults (middle) in the western pressure region as well as the basement faults located in southwest Oklahoma (bottom). The seismogenic faults are between 30° and 120° E while the known faults are primarily within 0° and 100° E in the western pressure region and 90° and 145° in southwest Oklahoma.

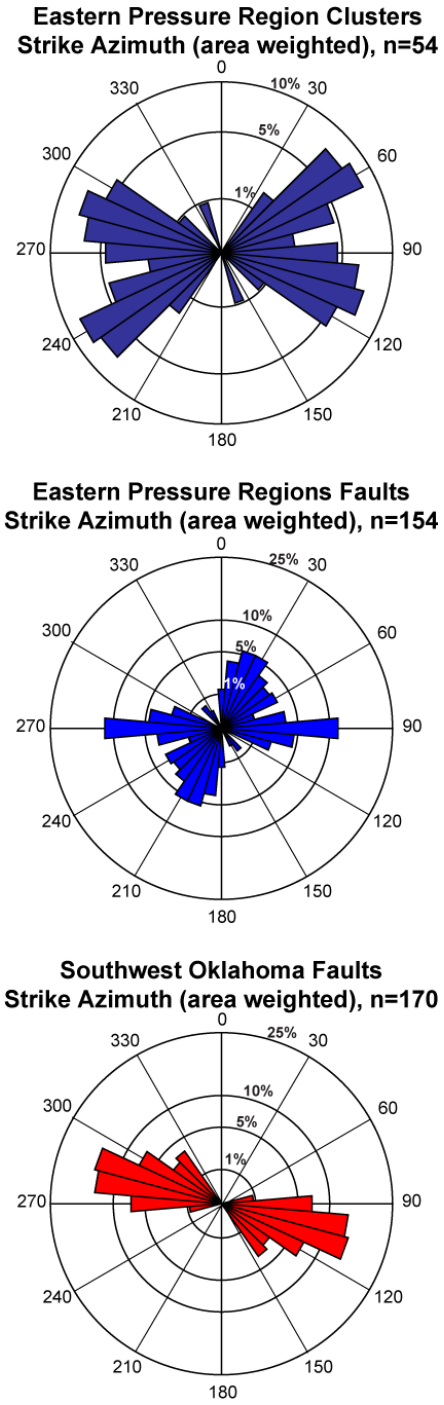

Figure 51. Rose diagram for the seismogenic faults (top) and mapped faults (middle) in the eastern pressure region as well as the basement faults located in southwest Oklahoma (bottom). The seismogenic faults are between 45° and 120° E while the known faults are primarily within 0° and 120° E in the eastern pressure region and 90° and 145° in southwest Oklahoma.

References

Aki, K., 1965. Maximum likelihood estimate of b in the formula logN= a-bM and its confidence limits.

Alt, R.C. and Zoback, M.D., 2016. In Situ Stress and Active Faulting in Oklahoma. Bulletin of the Seismological Society of America.

Barbour, A.J., Norbeck, J.N., and Rubinstein, J.L., 2017. The Effects of Varying Injection Rates in Osage County, Oklahoma on the 2016 M5.8 Pawnee Earthquake, Seismol. Res. Lett., in press.

Bodvarsson, G., 1970. Confined fluids as strain meters. Journal of Geophysical Research, 75(14), pp. 2711-2718.

Brizendine, T.A., 2017. Regional porosity evaluation of the Arbuckle Group within the Cherokee Platform (Doctoral dissertation, The University of Tulsa).

Carr, J.E., McGovern, H.E., Gogel, T. and Doveton, J.H., 1986. Geohydrology of and potential for fluid disposal in the Arbuckle aquifer in Kansas (No. 86-491). US Geological Survey.

Chang, K.W. and Segall, P., 2016. Injection-induced seismicity on basement faults including poroelastic stressing. Journal of Geophysical Research: Solid Earth, 121(4), pp. 2708-2726.

Chen, X. and Shearer, P.M., 2011. Comprehensive analysis of earthquake source spectra and swarms in the Salton Trough, California. Journal of Geophysical Research: Solid Earth, 116(B9).

Chen, C., 2016. Comprehensive analysis of Oklahoma earthquakes: from earthquake monitoring to 3D tomography and relocation (Doctoral dissertation, University of Oklahoma).

Chen, C., Nakata, N., Pennington, C., Haffener, J., Change, J.C., He, X., Zhan, Z., Ni, S., Walter, J., 2017 (under review). The Pawnee earthquake as a result of the interplay among injection, faults and foreshocks. Submitted to Nature.

Clauser, C., 1992. Permeability of crystalline rocks. Eos, Transactions American Geophysical Union, 73(21), pp. 233-238.

Crone, A.J. and Luza, K.V., 1990. Style and timing of Holocene surface faulting on the Meers fault, southwestern Oklahoma. Geological Society of America Bulletin, 102(1), pp. 1-17.

Darold, A.P. and Holland, A.A., 2015. Preliminary Oklahoma Optimal Fault Orientations Oklahoma Geological Survey Open-File Report, OF4-2015.

Davis, S.D. and Frohlich, C., 1993. Did (or will) fluid injection cause earthquakes?-criteria for a rational assessment. Seismological Research Letters, 64(3-4), pp. 207-224.

Deichmann, N., 2017. Theoretical Basis for the Observed Break in ML/Mw Scaling between Small and Large Earthquakes. Bulletin of the Seismological Society of America.

Dieterich, J.H., Richards-Dinger, K.B. and Kroll, K.A., 2015. Modeling injection-induced seismicity with the physics-based earthquake simulator RSQSim. Seismological Research Letters

Deng, K., Liu, Y. and Harrington, R.M., 2016. Poroelastic stress triggering of the December 2013 Crooked Lake, Alberta, induced seismicity sequence. Geophysical Research Letters, 43(16), pp. 8482-8491.

Ellsworth, W.L., 2013. Injection-induced earthquakes. Science, 341(6142), p. 1225942.

Franseen, Evan K., and Alan P. Byrnes, 2012. Arbuckle Group platform strata in Kansas: A synthesis, in J. R. Derby, R. D. Fritz, S. A. Longacre, W. A. Morgan, and C. A. Sternbach, eds., The great American carbonate bank: The geology and economic resources of the Cambrian–Ordovician Sauk megasequence of Laurentia: AAPG Memoir 98, p. 1031–1047.

Fritz, R.D., Medlock, P., Kuykendall, M.J. and Wilson, J.L., 2012. The Geology of the Arbuckle Group in the midcontinent: Sequence stratigraphy, reservoir development, and the potential for hydrocarbon exploration.

Frohlich, C., Hayward, C., Stump, B. and Potter, E., 2011. The Dallas–Fort Worth earthquake sequence: October 2008 through May 2009. Bulletin of the Seismological Society of America, 101(1), pp. 327-340.

Gardner, J.K. and Knopoff, L., 1974. Is the sequence of earthquakes in Southern California, with aftershocks removed, Poissonian?. Bulletin of the Seismological Society of America, 64(5), pp. 1363-1367.

Ghofrani, H., and Atkinson, G. M. 2016. A preliminary statistical model for hydraulic fracture-induced seismicity in the Western Canada Sedimentary Basin, Geophys. Res. Lett., 43, 10,164–10,172.

Goebel, T.H.W, Weingarten, M., Chen, X., Haffener, J., Brodsky, E. E., 2017 (under review). The 2016 Mw5.1 Fairview, Oklahoma earthquakes: Evidence for long-

range poroelastic triggering at > 40 km from fluid disposal wells. Submitted to Earth and Planetary Science Letters.

Hamerly, G., and C. Elkan, 2002. Alternatives to the k-means algorithm that find better clusterings, Proc. Elev. Int. Conf. Inf. Knowl. Manag., 4(9), pp. 600–607.

Harrison, W.E., Luza, K.V., Prater, M.L. and Chueng, P.K., 1983. Geothermal resource assessment of Oklahoma, Special Publication 83-1, Okla. Geological Survey.

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

Healy, D., Rizzo, R.E., Cornwell, D.G., Farrell, N.J., Watkins, H., Timms, N.E., Gomez-Rivas, E. and Smith, M., 2017. FracPaQ: A MATLAB™ toolbox for the quantification of fracture patterns. Journal of Structural Geology, 95, pp. 1-16.

Holland, A.A., 2013a. Earthquakes triggered by hydraulic fracturing in south-central Oklahoma. Bulletin of the Seismological Society of America, 103(3), pp. 1784-1792.

Holland, A.A., 2013b. Optimal fault orientations within Oklahoma. Seismological Research Letters, 84(5), pp. 876-890.

Holland, A.A., 2015. Preliminary fault map of Oklahoma. Oklahoma Geol. Surv. Open File Rep., OF3.

Horton, S., 2012. Disposal of Hydrofracking Waste Fluid by Injection into Subsurface Aquifers Triggers Earthquake Swarm in Central Arkansas with Potential for Damaging Earthquake, Seismol. Res. Lett., 83(2).

Keller, G.R. and Holland, A., 2013. Oklahoma Geological Survey evaluation of the Prague earthquake sequence of 2011. Oklahoma Geological Survey.

Kelson, K.I. and Swan, F.H., 1990. Paleoseismic history of the Meers fault, southwestern Oklahoma, and implications for evaluations of earthquake hazards in the central and eastern United States (No. NUREG/CP--0105-VOL. 2).

Keranen, K.M., Savage, H.M., Abers, G.A. and Cochran, E.S., 2013. Potentially induced earthquakes in Oklahoma, USA: Links between wastewater injection and the 2011 Mw 5.7 earthquake sequence. Geology, 41(6), pp. 699-702.

Keranen, K.M., Weingarten, M., Abers, G.A., Bekins, B.A. and Ge, S., 2014. Sharp increase in central Oklahoma seismicity since 2008 induced by massive wastewater injection. Science, 345(6195), pp. 448-451.

Kim, W.Y., 2013. Induced seismicity associated with fluid injection into a deep well in Youngstown, Ohio. Journal of Geophysical Research: Solid Earth, 118(7), pp. 3506-3518.

Langenbruch, C. and Zoback, M.D., 2016. How will induced seismicity in Oklahoma respond to decreased saltwater injection rates?. Science Advances, 2(11), p.e1601542.

Llenos, A.L. and Michael, A.J., 2013. Modeling earthquake rate changes in Oklahoma and Arkansas: Possible signatures of induced seismicity. Bulletin of the Seismological Society of America, 103(5), pp. 2850-2861.

McGarr, A., 1976. Seismic moments and volume changes. Journal of geophysical research, 81(8), pp. 1487-1494.

McGarr, A., Simpson, D. and Seeber, L., 2002. 40 case histories of induced and triggered seismicity. International Geophysics, 81, pp. 647-661.

McGarr, A., 2014. Maximum magnitude earthquakes induced by fluid injection. Journal of Geophysical Research: solid earth, 119(2), pp. 1008-1019.

McNamara, D.E., Benz, H.M., Herrmann, R.B., Bergman, E.A., Earle, P., Holland, A., Baldwin, R. and Gassner, A., 2015a. Earthquake hypocenters and focal mechanisms in central Oklahoma reveal a complex system of reactivated subsurface strike-slip faulting. Geophysical Research Letters, 42(8), pp. 2742-2749

McNamara, D.E., Hayes, G.P., Benz, H.M., Williams, R.A., McMahon, N.D., Aster, R.C., Holland, A., Sickbert, T., Herrmann, R., Briggs, R. and Smoczyk, G., 2015b. Reactivated faulting near Cushing, Oklahoma: Increased potential for a triggered earthquake in an area of United States strategic infrastructure. Geophysical Research Letters, 42(20), pp. 8328-8332.

Morgan, B.C., and Murray, K.E., 2015. Characterizing Small-Scale Permeability of the Arbuckle Group, Oklahoma Open-File Report (OF2-2015).

Murray, K.E., 2015. Class II Saltwater Disposal for 2009–2014 at the Annual-, State-, and County- Scales by Geologic Zones of Completion, Oklahoma. Oklahoma Geological Survey Open-File Report (OF5-2015), pp. 18.

Nemec, W., 1988. The shape of the rose. Sedimentary Geology, 59(1-2), pp. 149-152.

Oklahoma Corporation Commission, 2015. Media advisory – ongoing OCC earthquake response.

Oklahoma Corporation Commission, 2016a. Media advisory - regional earthquake response plan for western Oklahoma.

Oklahoma Corporation Commission, 2016b. Media advisory - regional earthquake response plan for central Oklahoma and expansion of the area of interest.

Pennington, C.N., and Chen, X., 2017. Coulomb Stress Interactions During the M5.8 Pawnee Sequence, Seismol. Res. Lett, in press.

Petersen, M.D., Moschetti, M.P., Powers, P.M., Mueller, C.S., Haller, K.M., Frankel, A.D., Zeng, Yuehua, Rezaeian, Sanaz, Harmsen, S.C., Boyd, O.S., Field, Ned, Chen, Rui, Rukstales, K.S., Luco, Nico, Wheeler, R.L., Williams, R.A., and Olsen, A.H., 2014. Documentation for the 2014 update of the United States national seismic hazard maps: U.S. Geological Survey Open-File Report 2014–1091, 243 p.

Qi, W., 2016. Stress analysis of recent earthquakes in Oklahoma (MS thesis at the University of Oklahoma).

Schorlemmer, D., Wiemer, S and Wyss, M., 2005. Variations in earthquake-size distribution across different stress regimes. Nature, 437, pp. 539-542.

Segall, P., and Lu, S., 2015. Injection-induced seismicity: Poroelastic and earthquake nucle- ation effects, J. Geophys. Res. Solid Earth, 120, pp. 5082–5103

Shah, A.K. and Keller, G.R., 2017. Geologic influence on induced seismicity: Constraints from potential field data in Oklahoma. Geophysical Research Letters.

Shapiro, S.A., Huenges, E. and Borm, G., 1997. Estimating the crust permeability from fluid-injection-induced seismic emission at the KTB site. Geophysical Journal International, 131(2), pp. F15-F18.

Shapiro, S.A., Dinske, C., Langenbruch, C. and Wenzel, F., 2010. Seismogenic index and magnitude probability of earthquakes induced during reservoir fluid stimulations. The Leading Edge, 29(3), pp. 304-309.

Shelly, D. R., Ellsworth, W. L., and Hill, D. P., 2016. Fluid-faulting evolution in high definition: Connecting fault structure and frequency-magnitude variations during the 2014 Long Valley Caldera, California, earthquake swarm, J. Geophys. Res. Solid Earth, 121, pp. 1776–1795.

Simpson, D.W., Leith, W.S. and Scholz, C.H., 1988. Two types of reservoir-induced seismicity. Bulletin of the Seismological Society of America, 78(6), pp. 2025-2040.

Talwani, P. and Acree, S., 1985. Pore pressure diffusion and the mechanism of reservoir-induced seismicity. Earthquake Prediction, pp. 947-965. Birkhäuser Basel.

Talwani, P. and Cobb, J.S., 1999. In situ measurements of hydraulic properties of a shear zone. Journal of geophysical research, 104(B7), pp.-14-993.

Talwani, P., Chen, L. and Gahalaut, K., 2007. Seismogenic permeability, ks. Journal of Geophysical Research: Solid Earth, 112(B7).

The Engineering Toolbox (TET), 2017. Water – Thermodynamic Properties. Web accessed April 2017. <http://www.engineeringtoolbox.com/water-thermal-properties-d_162.html>

Uhrhammer, R.A., 1986. Characteristics of northern and central California seismicity. Earthquake Notes, 57(1), p. 21.

van der Elst, N.J., Savage, H.M., Keranen, K.M. and Abers, G.A., 2013. Enhanced remote earthquake triggering at fluid-injection sites in the midwestern United States. Science, 341(6142), pp. 164-167.

van der Elst, N. J., M. T. Page, D. A. Weiser, T. H. W. Goebel, and S. M. Hosseini, 2016. Induced earthquake magnitudes are as large as (statistically) expected, J. Geophys. Res. Solid Earth, 121, pp. 4575–4590.

Vidale, J.E. and Shearer, P.M., 2006. A survey of 71 earthquake bursts across southern California: Exploring the role of pore fluid pressure fluctuations and aseismic slip as drivers. Journal of Geophysical Res. Solid Earth, 111(B5).

Walsh, F.R. and Zoback, M.D., 2015. Oklahoma's recent earthquakes and saltwater disposal. Science advances, 1(5), p.e1500195.

Walsh, F.R. and Zoback, M.D., 2016. Probabilistic assessment of potential fault slip related to injection-induced earthquakes: Application to north-central Oklahoma, USA. Geology, 44(12), pp. 991-994.

Wang, H., 2000. Theory of linear poroelasticity with applications to geomechanics and hydrogeology. Princeton University Press.

Wiemer, S. and Wyss, M., 2000. Minimum magnitude of completeness in earthquake catalogs: Examples from Alaska, the western United States, and Japan. Bulletin of the Seismological Society of America, 90(4), pp. 859-869.

Weingarten, M., Ge, S., Godt, J.W., Bekins, B.A. and Rubinstein, J.L. (2015), High-rate injection is associated with the increase in US mid-continent seismicity. Science, 348(6241), pp. 1336-1340.

Yeck, W.L., Block, L.V., Wood, C.K. and King, V.M., 2015. Maximum magnitude estimations of induced earthquakes at Paradox Valley, Colorado, from cumulative injection volume and geometry of seismicity clusters. Geophysical Journal International, 200(1), pp. 322-336.

Yeck, W.L., Weingarten, M., Benz, H.M., McNamara, D.E., Bergman, E.A., Herrmann, R.B., Rubinstein, J.L. and Earle, P.S., 2016. Far-field pressurization likely caused one of the largest injection induced earthquakes by reactivating a large preexisting basement fault structure. Geophysical Research Letters, 43(19).

Zaliapin, I. and Ben-Zion, Y., 2013. Earthquake clusters in southern California I: Identification and stability. Journal of Geophysical Research: Solid Earth, 118(6), pp. 2847-2864.

Appendix A: Relationship between slopes of volume versus seismicity

To verify the empirical volume-seismicity equations follow a Gutenberg-Richter distribution the proportionality between the slopes of Figure 14 were derived . Equations (6), (7), and (8) take the form of:

$$D = 10^{c} * \Delta V^{S} \quad \text{A1)}$$

where *D* represents either number of earthquakes, maximum moment, or total moment, *c* is some constant, ΔV is the change in volume, and S is the power of ΔV. In the logarithmic cross-plots of Figure 14, *S* dictates the slope of the best-fit lines while 10^{c}dictates the y-intercept. First to relate max moment to moment, taking the equation to relate moment M_0 to magnitude M of Hanks and Kanamori (1979) in terms of N*m results in:

$$log_{10}M_0 = 1.5M + 9.05 \quad \text{A2)}$$

Next, the equation (A2) is combined with:

$$M_{max} = M_c + \frac{1}{b} log_{10}N_{eq} \quad \text{A3)}$$

which relates the maximum magnitude event M_{max}, to the b-value b, and the number of events N_{eq} above a reference magnitude M_c (van der Elst et al., 2016). This yields:

$$M_0max = C + N_{eq}^{1.5/b} \quad \text{A4)}$$

where maximum moment M_0max is equal to a constant C plus the number of events to the power of 1.5 divided by the b-value. Substituting in equation (A1) for the maximum moment and number of events into equation (A4) gives:

$$S_{M_0max} \propto S_{Neq}^{1.5/b} \quad \text{A5)}$$

Next, to relate the slope of maximum moment and total moment, formulation of McGarr (2014) is applied:

$$\sum M_o = \frac{B}{1-B} M_o(\max) \quad : B = \frac{b}{1.5} \qquad \text{A6)}$$

where $\sum M_0$ is the total moment, $M_0 max$ is the maximum moment. One limitation to this formulation is it solves total moment from maximum moment to zero moment, while the catalog has a magnitude cutoff of Mc = 2.5. This results in an elevated maximum moment estimation by this equation, where the difference would only be notable at smaller moment scales. Substituting in equation (A1) for total and maximum moment gave:

$$S_{M_0} \propto S_{M_0 max} \qquad \text{A7)}$$

Based on Gutenberg-Richter statistics, the slope of volume versus moment should be proportional to the slope of volume versus maximum moment and proportional to the slope of volume versus number of earthquakes to the power of 1.5 over b.

Appendix B: Data artifacts of Pearson's r coefficient

As seen in Figure 29 and 30, both the observed and synthetic datasets contain a slope to the upper bound of time lag for maximum correlation. The reason for this stems from the Pearson's r coefficient calculation. The coefficient is calculated using the Fast-Fourier Transform (FFT) which pads before and after the two times series such that each padded time series is the length of the two combined series. This means for larger time lags, more and more of the injected volume time series is being compared to the zero-pad of the event time series (Figure 52). To circumvent this, the correlation is calculated only considering the overlapping portions of the two time series (Figure 52). These two differing methods, which I will refer to as the pad and non-pad method, have pros and cons.

The pad method results in better aligned peaks (Figure 53) and on a fundamental level assumes the volume fluctuations of a long lag time series will not result in additional seismicity (Figure 52). The non-pad method better aligns the increases in volume and seismicity (Figure 53) but for long lag time series ignores potentially substantial changes in volume that are unlikely, based on Langenbruch and Zoback (2016), to result in increases in seismicity. When the two methods are compared for a 0.6 degree grid, there is a general shift into high correlations and slightly longer lags using the non-pad method, with some unstable changes to large lags that is tied to early injection activity fitting the tail end of the seismicity time series (Figure 54). Because on a fundamental level the non-pad method results in a handful of long, unlikely shifts of the volume time series, misaligns peaks between datasets, and results in only slightly more grid cells above the 0.75 correlation cutoff, the padded method is used for the

analysis. This comes with the understanding that, assuming seismicity will not increase back to 2014–2016 levels, longer term lags represent un-correlated time series.

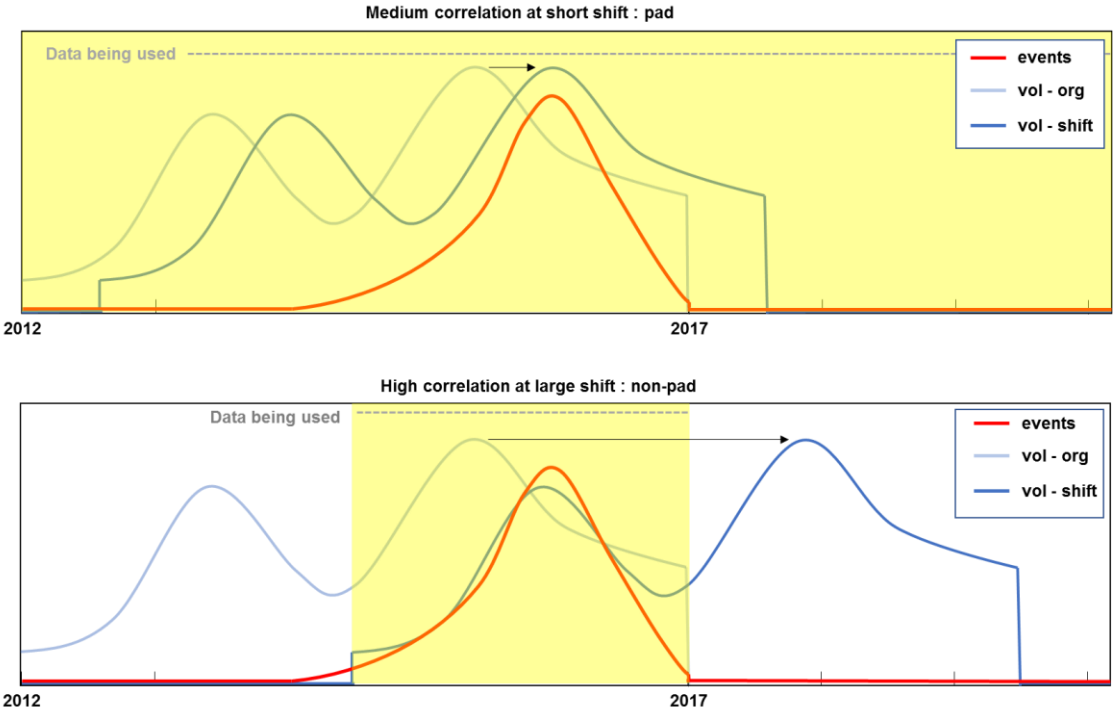

Figure 52. Example of the pad and non-pad Pearson's r calculation. The padded version penalized large amounts of the volume time series being past 2017, where only zero padded event data is present. The non-padded version only takes the overlapping data (yellow window) and does not penalized fluid past 2017. However, with this is the implication that any volume can generate seismicity in the near future. In this hypothetical for the non-padded method the second fluid peak would be expected to generate a similar level of seismicity as the first peak.

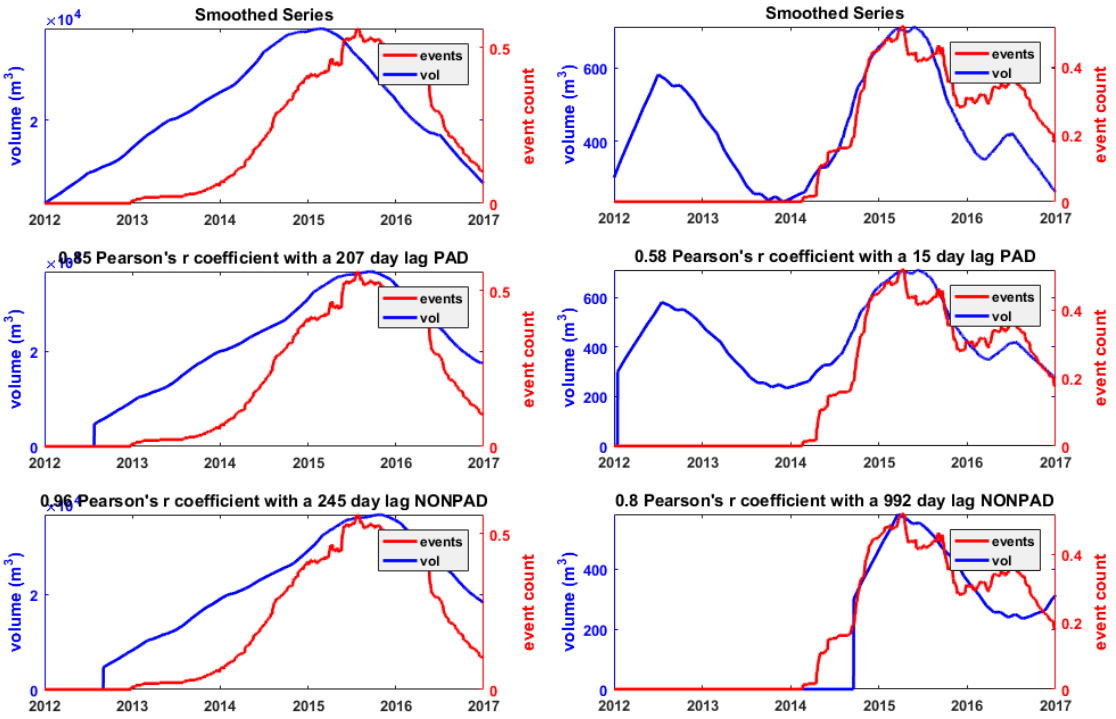

Figure 53. Two examples of the pad versus non-pad methods applied to 0.6 degree grid data. (Left) The padded version better aligns the peaks of the two time series while the non-pad shifts the volume data slightly further, as well as showing a higher correlation due to the lack of fluid data being compared to zero padded event data. (Right) The padded version lines up the nearby peak in the time series and shows a low correlation while the non-padded method does a very large shift to align the first peak with the event data. Like the example in Figure 52, this has implications that the second peak will cause seismicity in the future.

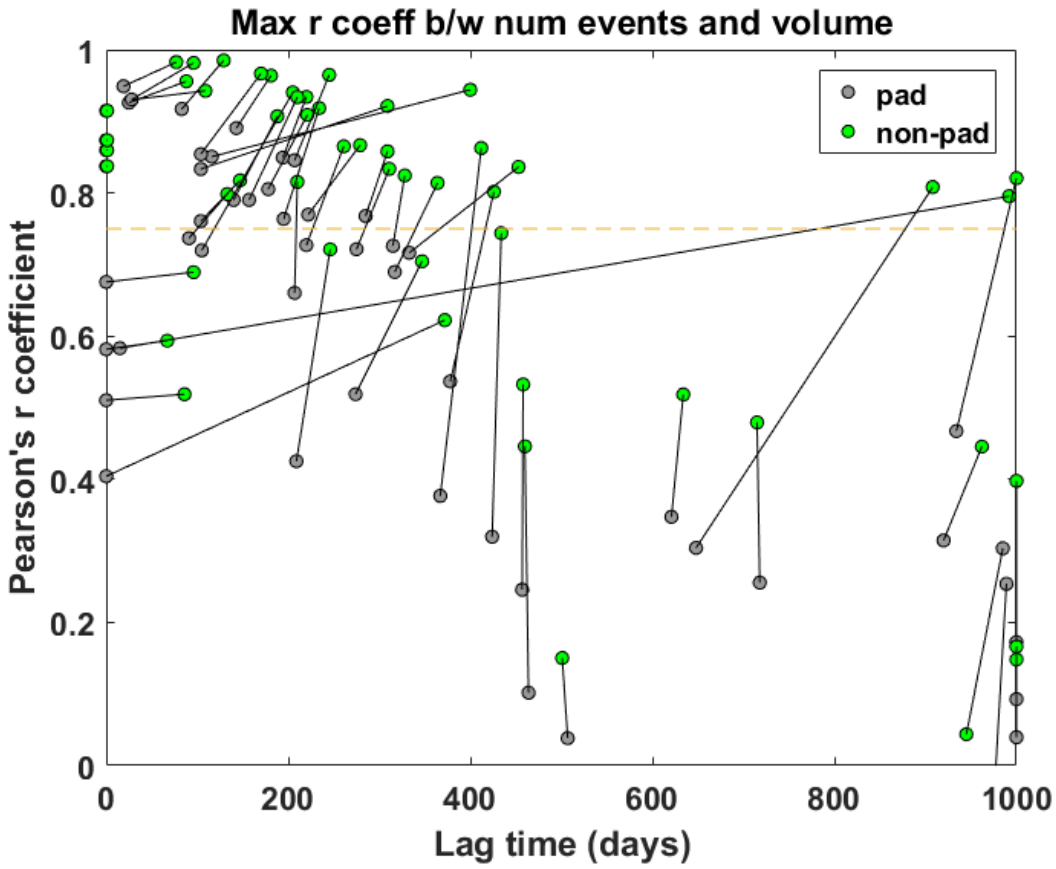

Figure 54. Differences between the pad and non-padded method for the 0.6 degree grid size. The non-padded method results in higher correlations due only using overlapping data and slightly greater time lags due to a better matching of the increases between the two times series. It also has noticeably unstable results compared to the padded method, where low correlation short time lag cells are shifted to high correlation high time lag cells similar to the example in Figure 52.