J Seismol (2020) 24:921–935
https://doi.org/10.1007/s10950-019-09864-9

ORIGINAL ARTICLE

Temporal patterns of induced seismicity in Oklahoma revealed from multi-station template matching

Robert J. Skoumal · Michael R. Brudzinski · Brian S. Currie · Rosamiel Ries

Received: 22 April 2019 / Accepted: 13 August 2019 / Published online: 29 August 2019

Abstract Over the past decade, Oklahoma became the most seismically active region of the mid-Continental USA as a result of industry operations. However, seismic network limitations and completeness of earthquake catalogs have restricted the types of analyses that can be performed. By applying multi-station template matching on the 23,889 cataloged earthquakes in Oklahoma and Southern Kansas between late-2008 and 2016, we increased the number of detected earthquakes to 209,409 events. While the improved catalog produced an order of magnitude events than the original catalog, the frequency-magnitude distribution remains similar to the original catalog. We found that the coefficient of variation of interevent times in small spatial bins tends to spatially correlate with the location of $M \geq 4$ earthquakes. The improved catalog reveals the pervasiveness of swarm-like patterns in seismicity across the entire study region. The rapid increase in seismicity rate of these swarms in 2013 coincided with a reduction in the calculated p values (power law decay rates) before and after larger events. We also used the catalog to revisit the temporal patterns in the four $M \geq 5$ sequences, finding more active foreshock behavior than previously recognized and variations in aftershock behavior. When compared against poroelastic stress models for the Pawnee and Fairview sequences, the catalog shows an improved correlation with stress that accounts for variable-rate injection, supporting the conclusion that injection rate is an important contributor to seismic hazard.

Keywords Induced seismicity · Template matching · Oklahoma

Highlights
•We increased the number of cataloged earthquakes in Oklahoma between 2008 and 2016 by roughly an order of magnitude
•This improved catalog allows individual sequences and regional seismicity patterns to be analyzed in greater detail
•We find more foreshock activity prior to $M \geq 5$ earthquakes, pervasive swarminess, and high coefficients of variation

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s10950-019-09864-9) contains supplementary material, which is available to authorized users.

R. J. Skoumal (✉)
U.S. Geological Survey, Earthquake Science Center, 345 Middlefield Rd., Menlo Park, CA 94025, USA
e-mail: rskoumal@usgs.gov

M. R. Brudzinski · B. S. Currie · R. Ries
Department of Geology and Environmental Earth Science, Miami University, 118 Shideler Hall, 250 S. Patterson Ave., Oxford, OH 45056, USA

M. R. Brudzinski
e-mail: brudzimr@miamioh.edu

B. S. Currie
e-mail: curriebs@miamioh.edu

R. Ries
e-mail: riesr@miamioh.edu

1 Introduction

Between 1975 and 2007, the Advanced National Seismic System (ANSS) Comprehensive Earthquake Catalog (ComCat) demonstrates an average rate of ~ 20 $M \geq$ 3 earthquakes/year in the Central and Eastern USA. A decade ago, the seismicity rate began to increase, reaching ~ 1000 $M \geq 3$ earthquakes in 2015 alone. The largest rate increase occurred in Central and Northern Oklahoma. Historically, an average of ~ 1–2 $M \geq 3$ earthquakes/year occurred in Oklahoma, but the number of earthquakes rose to over 900 $M \geq 3$ earthquakes in 2015. While the seismicity rate declined in 2016, the yearly total seismic moment in Oklahoma increased largely due to three $M \geq 5$ earthquakes, including the 2016 M 5.8 Pawnee earthquake, the largest earthquake ever recorded in Oklahoma.

The increase in oil and gas recovery operations, development of new hydrocarbon plays, and change of industry practices are thought to be largely responsible for this dramatic rise in seismicity (e.g., Ellsworth 2013; McGarr et al. 2015). The recent seismicity rate increase in Oklahoma over the past decade is largely related to the disposal of large volumes of produced water into the Arbuckle Group, a Cambrian-Ordovician formation composed primarily of dolomitized carbonates, which is proximal to the Precambrian basement (e.g., Keranen et al. 2014; Walsh and Zoback 2015; Haffener et al. 2018). The disposal of produced water is likely associated with the 2011 Mw 5.7 Prague (Keranen et al. 2013; Sumy et al. 2014), 2016 Mw 5.1 Fairview (Yeck et al. 2016b; Goebel et al. 2017), 2016 Mw 5.8 Pawnee (Barbour et al. 2017; Chen et al. 2017), and 2016 Mw 5.0 Cushing earthquakes (McNamara et al. 2015a; McGarr and Barbour 2017). Hundreds of hydraulic fractured wells have also been found to have induced earthquakes during this time (Skoumal et al. 2018a). The increase in seismicity rate over the past few years has raised public concern due to the growing seismic risk to local populations and infrastructure (McNamara et al. 2015b). Recent seismic hazard assessments in Oklahoma are comparable to those in California or New Madrid (Ellsworth et al. 2015; Petersen et al. 2017).

Because of the relatively low historical seismicity rate in Oklahoma, the state had a limited seismic network when the dramatic increase in seismicity rate began which has inhibited the evaluation of the early evolution of the earthquake sequences. Additionally, while the seismic network has since been improved in some parts of the state, the magnitude of completeness of cataloged earthquakes (M_C) is between M 2–3, limiting the types of analyses that can be performed. We have attempted to address this problem by applying large-scale waveform template matching to improve the completeness of the catalog. Template matching has become an important observational tool for the characterization of repetitive seismic sources (e.g., Shelly et al. 2007) and is known to lower the earthquake detection threshold by ~ 1 magnitude unit below what standard processing detects (e.g., Schaff 2008). Template matching is particularly well-suited for characterizing induced seismicity due to the repetitive, swarm-like nature of the sequences (e.g., Cochran et al. 2018; Healy et al. 1968; Rubinstein et al. 2018a; Skoumal et al. 2015a; Walter et al. 2017a; Yeck et al. 2016a). Lowering the magnitude of completeness enables a more detailed description of the induced earthquake sequence and capability to evaluate the conditions that caused them.

2 Methods

Template matching (Fig. S1) generally has three requirements: (1) a signal of interest must be identified beforehand, (2) multiple templates if the seismic sequences are from different sources, and (3) computational capability to perform the correlations. We have dealt with the first two requirements by using existing catalogs with a M_C ~ 2–3. Parameters for template matching in this study were the optimized set previously established for regional template matching (Skoumal et al. 2014), including the use of three regional stations, 30-s template lengths, and bandpass filtered data between 5 and 15 Hz. Also consistent with this optimized template matching set, template times were offset so that a template began ~ 5 s prior to the P-wave on vertical components and ~ 5 s prior to the S-wave arrival on horizontal components. These phase arrival times were roughly approximated based on the method of Trabant et al. (2012) using the IRIS DMC iasp91 velocity model. Template matches with a Network Normalized Cross-Correlation Coefficient (NNCC) above 15 times the daily median absolute deviation (MAD) of the NNCC are determined to be successful detections, as this threshold would be expected to produce < 1 false

positive per year of data (Skoumal et al. 2014). Because it was not practical to manually inspect all detected matches due to the large number of earthquakes, a very small number of false positives in the final catalog could be expected.

In the region around Oklahoma, the closest high-quality, long-standing stations were US.WMOK, TA.TUL1, and US.MIAR (Fig. 1) with reliable uptime between 16 October 2008 and 31 December 2016. We constructed templates from all of the 21,563 earthquakes in Oklahoma cataloged by the Oklahoma Geological Survey (OGS) and all of the 2326 earthquakes from the Advanced National Seismic System (ANSS) catalog in southern Kansas that occurred during our study window (Fig. 1). Template matching this catalog across our network for the selected time frame required more than two quadrillion correlations, made feasible by the use of an optimized correlation routine on a high-performance cluster.

To evaluate the geographic extent of such a large number of earthquakes, we assumed that newly detected matched events produced by template matching have a similar location as the template that identified the event. Earthquakes were then grouped into geographical bins to examine temporal patterns on a regional scale of approximately 10 × 10 km. Other work (Skoumal et al. 2019) was able to relocate ~ 77% of the events in the template matching catalog, which allowed for more local spatial methodologies (e.g., fault mapping) to be performed. Here, we focus on using the full template matching catalog to investigate regional-scale temporal patterns.

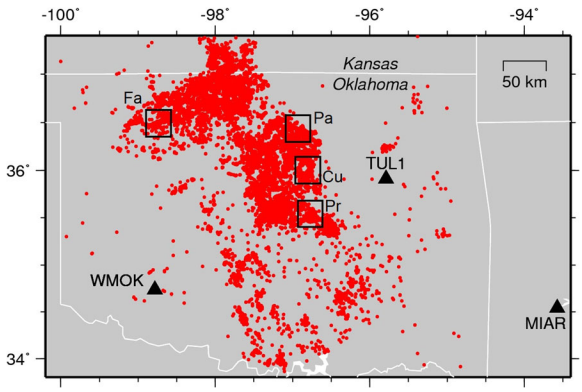

Fig. 1 Map showing locations of earthquakes (red dots) 16 October 2008–31 December 2016 from the OGS catalog in Oklahoma and from the ANSS catalog for Kansas. Triangles represent the three seismic stations used for template matching. Boxes represent the (Cu)shing, (Fa)irview, (Pa)wnee, and (Pr)ague regions used in Fig. 7

Local magnitudes were estimated through a Richter scale approach:

$$M_L = \log_{10}[A/A_0]. \quad (1)$$

The scale factor (A_0) was calculated using the single station US.WMOK with peak-to-peak filtered S-waveform amplitudes (A) and catalog magnitudes for each event. During two data gaps on US.WMOK (in February 2010 and December 2015), station TA.TUL1 was used for the magnitude calculation. The high-quality US.WMOK station was preferentially used to estimate magnitudes due to a better signal-to-noise ratios than the stations TA.TUL1 and US.MIAR. Magnitudes for matched events were determined through relative comparisons to the template event that produced the largest NNCC. A small subset of earthquakes in the OGS catalog have unusually low reported magnitudes relative to other sequences in the same region, so for cataloged events with $M \leq 1.5$, we computed the magnitude directly from the seismograms using a calibrated attenuation-distance relationship tuned to Oklahoma (e.g., Walter et al. 2016, 2017b). We estimated the reliability of the OGS catalog magnitudes by comparing them with common events in the NEIC catalog and found that the mean difference was less than 0.01 magnitude units with a standard deviation of 0.11. We further investigated the uncertainty of our Richter magnitudes by recalculating the magnitude of events in the OGS catalog using scale factors for other cataloged events in a small (< 4 km) nearby region. This approach identified a standard deviation of 0.25 magnitude units.

The Gutenberg-Richter b-values (Gutenberg and Richter 1944) were estimated using the maximum likelihood estimate (Aki 1965) and the corrected maximum curvature algorithm (Wiemer and Wyss 2000; Woessner and Wiemer 2005). Gutenberg-Richter b-values reflect the relationship between the frequency occurrence of earthquake magnitudes. Global earthquake catalogs and catalogs in seismically active regions tend to have a b-value ≈ 1.0, but b-values can vary between regions or over time. Gutenberg-Richter b-values > 1 represent a larger proportion of smaller magnitude events to large earthquakes. We estimated uncertainties in the b-value estimations by performing bootstrap resampling

removing 10% of the dataset and applying a random magnitude perturbation consistent with a normal distribution with a standard deviation from the combined magnitude uncertainties.

Omori-Utsu aftershock decay p values (Utsu et al. 1995) were calculated using both the improved and OGS catalogs for the Mw 5.7 Prague, Mw 5.1 Fairview, Mw 5.8 Pawnee, and Mw 5.0 Cushing earthquake sequences. An Omori-Utsu p value represents the power law decay rate of an aftershock sequence and can also vary between regions. Omori-Utsu p values were calculated over 50-day periods after each mainshock, with uncertainties estimated by the Kolmogorov-Smirnov test for goodness of fit (Massey Jr 1951; Miller 1956). We also calculated Omori-Utsu p values for each geographical bin by setting the largest magnitude earthquake in the bin as the mainshock, and then examining the rate decay both before and after (e.g., Schoenball and Ellsworth 2017). We restricted the aftershock selection window to cases where the seismicity rates can be fit by a similarly sloped straight line in log-log space over different time windows and can be used as a proxy for whether choosing different time windows will substantially affect the p value. This helps to restrict the influence of increased background seismicity rate that would increase the rate at larger times, creating a curved distribution with shallower slopes (apparent lower p values) at larger times. We also created stacked event rates by combining all events across spatial bins by normalizing time relative to the mainshock. We then explored temporal variations in the p values by only stacking events with a mainshock during a given year. For comparison, we also determined p values for specific Oklahoma sequences that were identified as being induced by hydraulic fracturing (Skoumal et al. 2018a) by using the largest event in each sequence as the mainshock.

The temporal pattern of seismicity was also evaluated by calculating the coefficient of variation (CoV), which is defined as the ratio of the standard deviation of interevent times to the average interevent time of the earthquakes (Kagan and Jackson 1991) in each of our selected geographical bins. If earthquakes are randomly distributed in time, we would expect $\mathrm{CoV} \approx 1$, while $\mathrm{CoV} < 1$ represents periodic seismicity and $\mathrm{CoV} > 1$ indicates temporal clustering (Kagan and Jackson 1991; Schoenball and Ellsworth 2017; Cochran et al. 2018).

3 Results and discussion

3.1 Regional template matching catalog ("SBCR19")

Utilizing the 23,889 earthquakes from the combined OGS and ANSS catalogs, our template matching algorithm detected 209,409 unique events above the 15 × MAD threshold from late 2008–2016 (Fig. 2). Using the corrected maximum curvature algorithm (Wiemer and Wyss 2000; Woessner and Wiemer 2005), we found that this more complete catalog ("SBCR19") has a lower magnitude of completeness ($M_C = 1.6$) than the starting catalog ($M_C = 2.3$). The Gutenberg-Richter b-values of the SBCR19 catalog ($b = 0.87 \pm 0.04$) and of the pre-existing catalog ($b = 0.92 \pm 0.06$) are similar (Fig. S2), indicating the template matching process largely preserved the frequency-magnitude distribution.

While there is roughly an order of magnitude more events in the SBCR19 catalog, the relative spatiotemporal distribution of matched events is still similar to the initial catalog (Figs. 2c and 3). At the start of our study in late-2008, the seismicity rate exceeded the average background rate by more than an order of magnitude. As has been shown in previous studies (e.g., Walsh and Zoback 2015; Weingarten et al. 2015; Langenbruch and Zoback 2016), the shape of the seismicity rate increase followed the increase of fluid disposal into the Arbuckle Group.

The spatial density of earthquakes throughout the 2008–2016 study window is shown in Fig. 4a. While a majority of the earthquakes occurred in north-central Oklahoma and southern Kansas where wastewater disposal wells were more prevalent (e.g., Weingarten et al. 2015), productive earthquake sequences in parts of the southern-half of the state far from large-volume disposal wells were also found. The increased completeness of the SBCR19 catalog has previously been used to demonstrate that many sequences in the southern and western areas of Oklahoma are predominantly induced by hydraulic fracturing stimulations rather than wastewater disposal (Skoumal et al. 2018a).

The maximum magnitude of each geographical bin is shown in Fig. 4b. The largest magnitude earthquakes were located in Northern and Central Oklahoma and have been associated with large volume disposal wells (e.g., Keranen et al. 2014; Weingarten et al. 2015; Yeck et al. 2017). As $M > 3$ earthquakes are contained in nearly every geographical bin in this area, it appears that critically stressed faults are prevalent throughout a

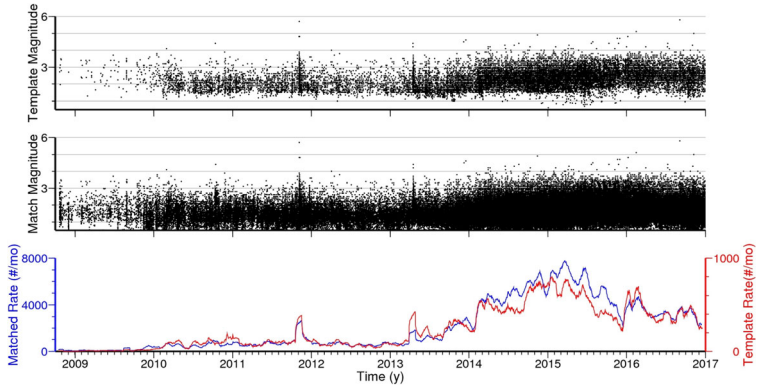

Fig. 2 Magnitudes of (a) template (cataloged) earthquakes and (b) matched earthquakes identified by template matching over time. (c) Seismicity rates per month for templates (red) and matched events (blue)

broad geographic area in the shallow crust. The proximity of injection depth to the Precambrian basement has been found to have a strong correlation between the likelihood of induced seismicity in Oklahoma and other basins in the USA (Zhang et al. 2013; Hincks et al. 2018; Skoumal et al. 2018b), and could help explain the pervasiveness of seismicity in this area.

The CoV observed in each geographic bin is shown in Fig. 4d. We found that > 97% of spatial bins had CoV values > 1, with CoV values up to ~ 20, indicating pervasive temporal clustering of earthquakes across our study area (Fig. S6). Our CoV results are similar to previous work in Kansas with CoV values that ranged between 1 and 8 (Cochran et al. 2018), but we found that ~ 10% of our spatial bins had CoV > 8. Intriguingly, we found that these larger CoV values generally coincide with the areas that have had $M > 4$ earthquakes (Fig. 4d), potentially suggesting stress interactions from these large earthquakes (Schoenball and Ellsworth 2017).

3.2 Swarm behavior

An effort to distinguish between earthquakes swarms and mainshock-aftershock sequences was previously proposed by comparing the maximum magnitude in a sequence to the total number of earthquakes above a given magnitude threshold (Vidale and Shearer 2006; Holtkamp and Brudzinski 2011). This proposed classification boundary was later extended to small magnitude ($M < 3$) earthquakes and used with template matching to

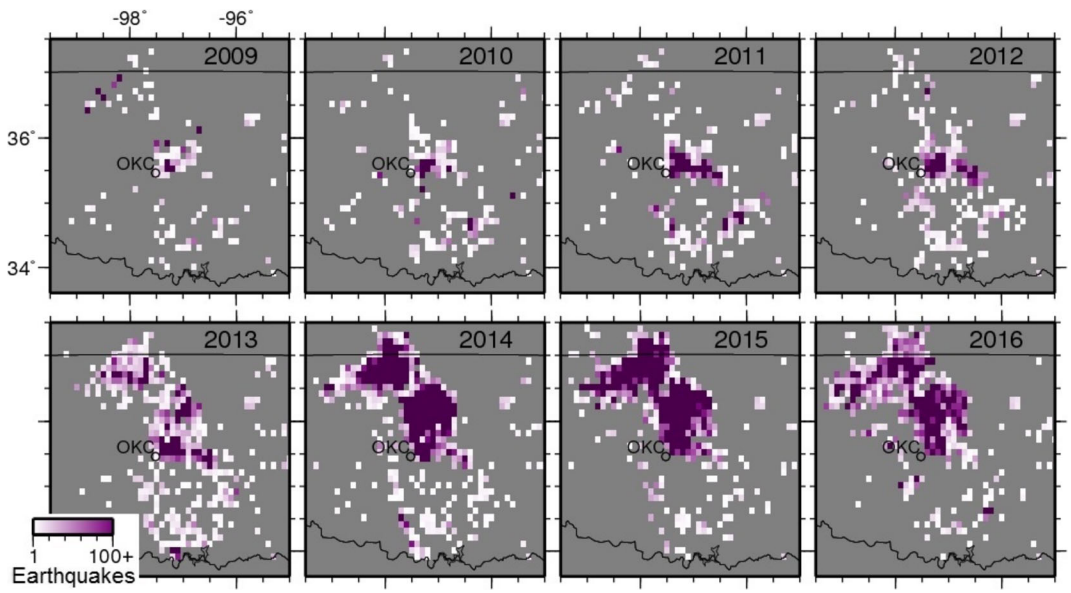

Fig. 3 Map of the SBCR19 catalog showing the number of events in 0.1° × 0.1° bins by year. State borders (black) and Oklahoma City (black circle) are shown

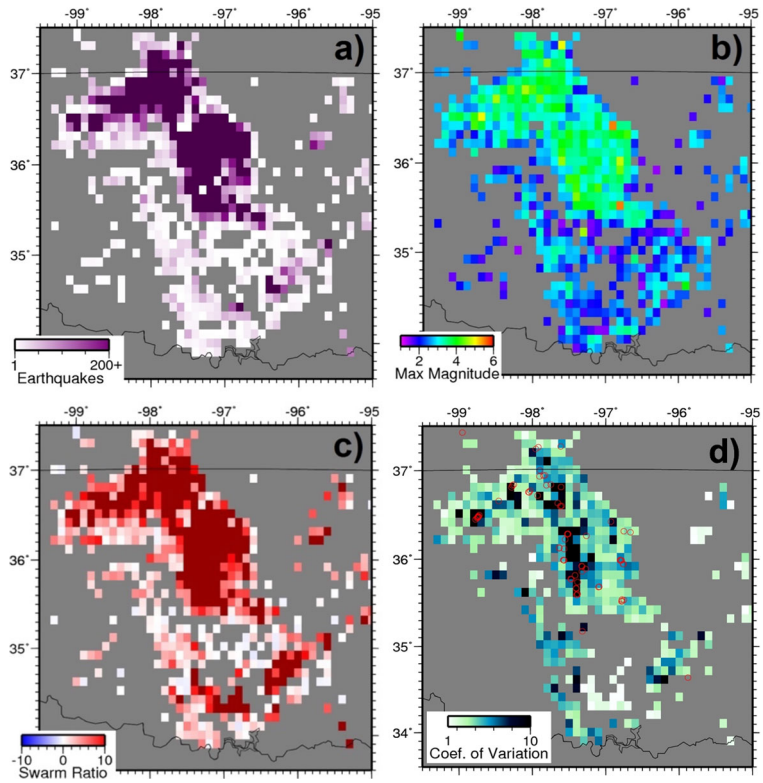

Fig. 4 Maps showing earthquake patterns in each 0.1° × 0.1° bin using the SBCR19 catalog. **a**) Total number of events with $M > 1.5$. **b**) Maximum magnitude. **c**) Swarm ratio, where colored shading indicates number of events above (positive) or below (negative) the red line separating swarms and aftershocks in Fig. 5 divided by the number of events expected based on that line. **d**) CoV values and the location of $M \geq 4$ earthquakes (red circles)

help distinguish induced seismicity from naturally occurring intraplate seismicity (Skoumal et al. 2015a, 2018b) (Fig. 5a). Here, we apply this swarm classification to the SBCR19 catalog. Since a specific, individual earthquake sequence is harder to define in Oklahoma given the much larger amount of seismicity, we used the geographical bins to define sequences and then compared them to the classification boundary (Fig. 5b, red line). In addition to the 0.1° × 0.1° geographic bins, we explored that a variety of spatial bin sizes to see whether our characterizations of swarm behavior were a result of spatial binning (Fig. S3). We also explored subdividing seismicity in each bin into multiple sequences if there was a temporal gap in seismicity of windows. We also investigated whether eschewing spatial bins and focusing on individual clustered sequences following the explicit spatial and temporal criteria of Vidale and Shearer (2006) would have an influence on our results (Fig. S4). We found that vast majority of seismic sequences in our study fell into the swarm category regardless of the selection parameters.

Although we refer to the distinction between swarms and aftershocks as a boundary for the purposes of classification, in actuality, the transition from aftershocks to swarms should be thought of as a spectrum. The number of events away from this boundary, termed the "residual events," is an attempt to quantify the degree to which seismicity is more or less swarm-like. The confidence in a swarm/mainshock-aftershock classification (the "swarminess") is proportional to the magnitude of separation from the boundary line (i.e., the residual events); sequences that further from this boundary can be defined more confidently as either a swarm or mainshock-aftershock sequence, whereas sequences near this boundary would be more ambiguous.

Using the number of earthquakes (Fig. 4a) and the maximum magnitudes (Fig. 4b) for each geographical bin, the swarminess for each bin was determined (Fig. 4c). Over 98% of the geographical bins that have at least 30 $M > 1.5$ earthquakes would be classified as swarms (Figs. 4c and 5b). This swarm-like pattern has been atypical for intracontinental seismicity, but it has been commonly observed in fluid-injection-induced seismicity (Fig. 5a) (e.g., Horton 2012; Skoumal et al. 2015b; Walter et al. 2016, 2017a; Skoumal et al. 2018b). This indicates an interpretation similar to previous studies which concluded that nearly all of the seismicity in Oklahoma and southern Kansas appears to be related

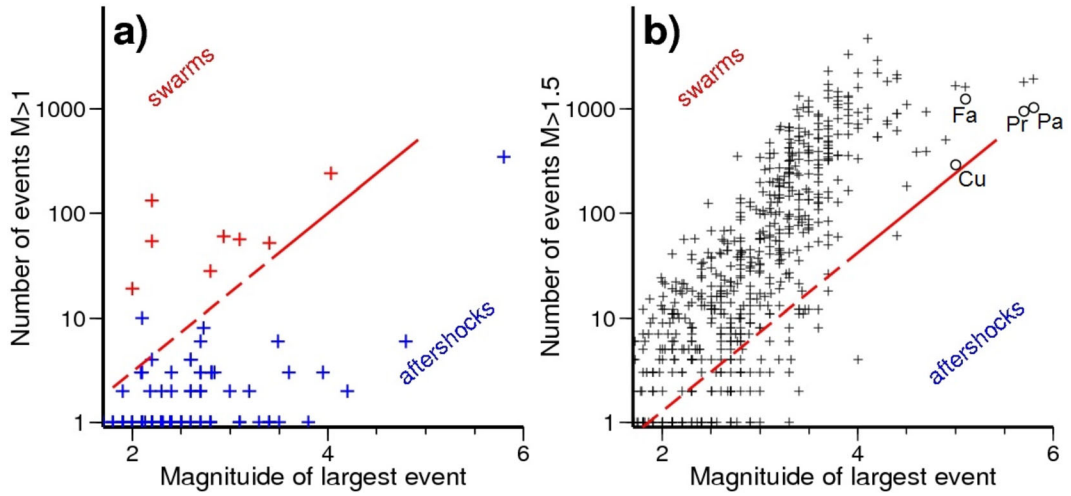

Fig. 5 Quantification of swarm behavior for earthquake sequences in **a**) the Appalachian, Illinois, and Williston Basins (Skoumal et al. 2015, Skoumal et al. 2018b); and **b**) Oklahoma. In **a**), cross color denotes sequences proposed to have been induced (red) or natural (blue). In **b**), each cross represents the Oklahoma and Kansas earthquakes found in each 0.1° × 0.1° bin. Circles are 50-day subsets for specific large ($M \geq 5$) event sequences: (Cu)shing, (Fa)irview, (Pa)wnee, and (Pr)ague. Solid red line is the boundary between swarms and mainshock-aftershock sequences (Vidale and Shearer 2006; Holtkamp and Brudzinski 2011; Skoumal et al. 2015), and the dashed line is an extension following the same slope. The red line is adjusted in (**b**) to account for a higher magnitude threshold in the Oklahoma catalog than that in previous work (Skoumal et al. 2015, Skoumal et al. 2018b)

to fluid injection (Walsh and Zoback 2015; Rubinstein et al. 2018b). When we isolate the four $M \geq 5$ sequences to just their mainshock and 50 days of aftershocks (the maximum available for the Cushing earthquake due to the end of the study window), we find that three of the four sequences fall near the dividing line between mainshock-aftershocks and swarms. The exception is the Fairview sequence which appears to be significantly more swarm-like than the others. This suggests that while induced seismicity in Oklahoma generally follows a swarm-like pattern, the larger magnitude sequences tend to represent a combination of both swarm and mainshock/aftershock behavior and is consistent with our CoV results described previously.

To further examine the aftershock behavior of events in this study, we use the stacked mainshock decay rate analysis. Using the largest event in each of the ~600 spatial bins as the mainshock, we found that the SBCR19 and OGS catalogs produce similar aftershock decay rates with relatively low p values of 0.55 and 0.59, respectively (Fig. 6b). Applying the analysis to seismicity preceding each mainshock, we found similar but even lower p values of 0.45 and 0.46 for the SBCR19 and OGS catalogs, respectively (Fig. 6a). Given our methodology of analyzing seismicity spatial bins, the p values appear to be consistent between both the OGS and our SBCR19 catalog. These results are also generally consistent with similar analysis performed by Schoenball and Ellsworth (2017) on the OGS catalog, but we do not find unusual rate variations in the month prior to mainshocks as they suggested. They interpreted that the lower p values were a result of the sequence stacking process which tends to show smaller values than well-developed single mainshock-aftershock sequences (Utsu et al. 1995). However, when we isolated the calculations to stack 172 sequences of hydraulic fracturing-induced seismicity (Skoumal et al. 2018a), we found higher p values of 0.95 prior to and 0.98 following the largest magnitude event in those sequences (Fig. 6c,d). While there are some cases of hydraulic fracturing-induced seismicity lasting months following stimulation (e.g., Skoumal et al. 2015b), the majority of earthquakes induced by hydraulic fracturing occur during the timing of the stimulation (typically lasting days to weeks) and quickly decay afterwards.

To help investigate whether the lower p values might be due to changing earthquake rates over time, we calculated yearly p values by building stacks from mainshocks restricted to a year time frame at a time. Temporally, p values prior to 2013 were > 0.7, while the dramatic increase seismicity around 2013 coincides with a decrease in p value < 0.7 (Fig. 6e). This trend of

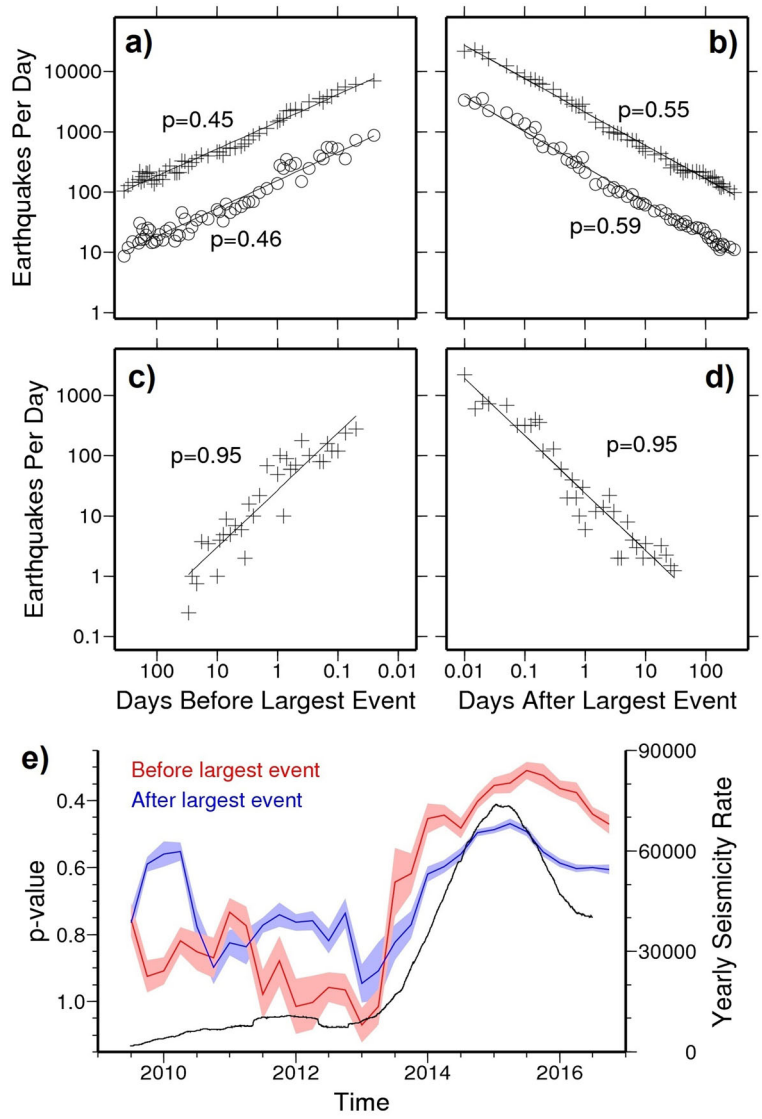

Fig. 6 Omori-Utsu p values. p values of geographic regions prior to the largest magnitude event using the improved catalog (crosses) and OGS catalog (circles) are shown for the time period **a**) prior to and **b**) following the occurrence of the maximum magnitude in the respective geographical bin. p values of hydraulic fracturing induced sequences (Skoumal et al. 2018a) in Oklahoma in the time period **c**) prior to and **d**) the occurrence of the maximum magnitude earthquake. **e**) Yearly seismicity rate (black line) compared with the temporal evolution of p value before (red) and after (blue) the largest magnitude event. Shaded regions represent the σ confidence bounds of the p value calculations determined from bootstrapping

decreasing p values correlates with the rapid increase in seismicity rate and can be observed regardless of the geographic bin size or temporal subdivisions that were tested (Fig. S5). Considering that there were an increasing number of sequences in each time window (up to ~ 150 in 2015) that could potentially reduce the p value over time, we also performed a version of the bootstrapping by restricting each iteration and time window stack to a randomly selected 40 sequences. The resulting p values did not vary from that shown in Fig. 6e by more than 0.1 with an average change of 0.01, similar to the standard deviation range shown in Fig. 6e. Thus, we have interpreted the reduction in p values as a real feature of the seismicity.

3.3 Characterizing the $M \geq 5$ Oklahoma sequences

We used the SBCR19 catalog to re-examine the four Mw ≥ 5 events in Oklahoma that occurred during our study: the 2011 Mw 5.7 Prague, 2016 Mw 5.1 Fairview, 2016 Mw 5.8 Pawnee, and 2016 Mw 5.0 Cushing earthquakes. To help examine these sequences, we identified areas that approximate the faults associated with the larger ruptures. These regions were determined by examining the first 6 h of aftershocks, a time period found to minimize the number of triggered earthquakes while including a representative number of aftershocks (Fig. 7). To provide some regional context for temporal patterns in the area where these larger events occurred,

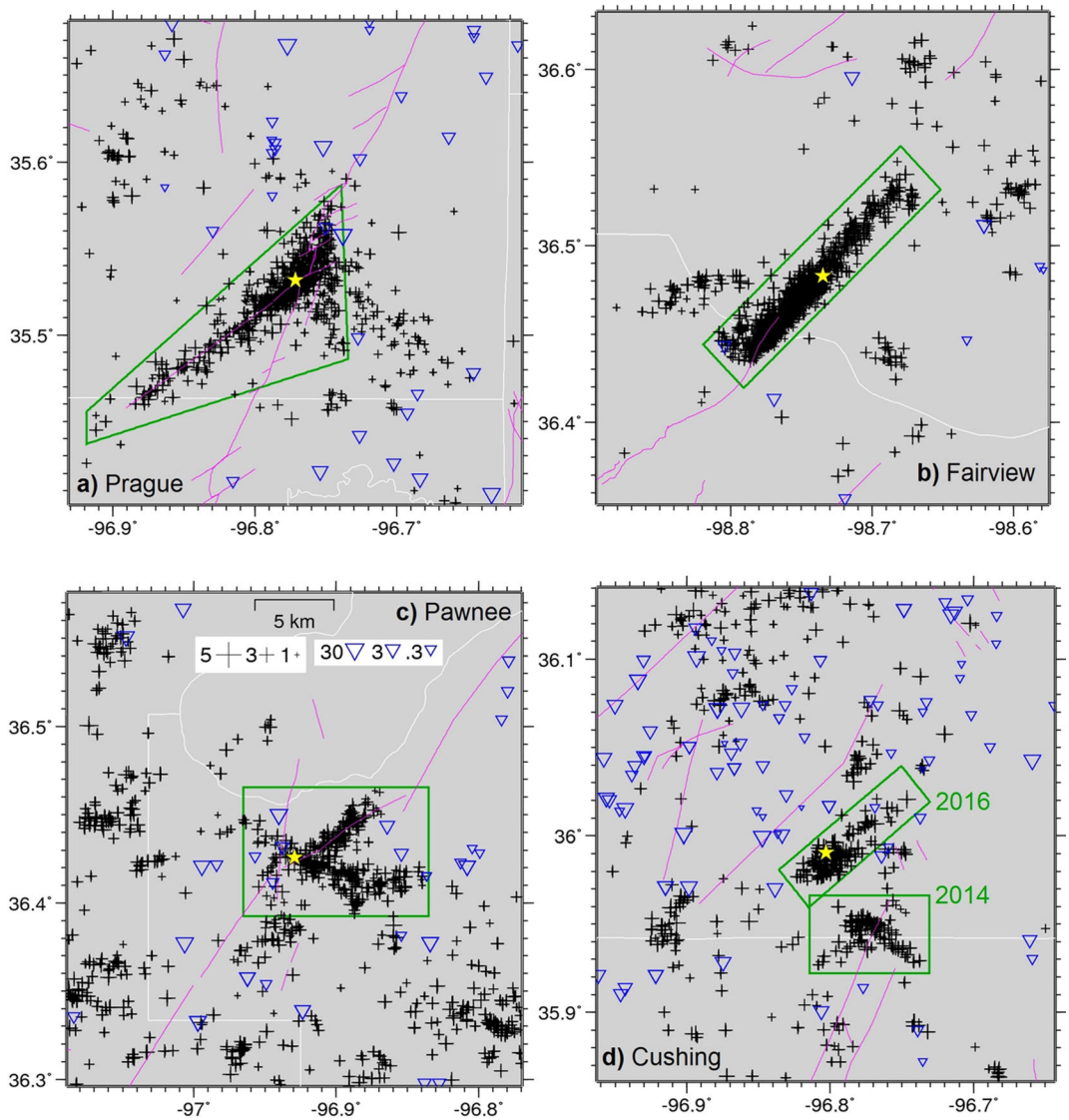

Fig. 7 Maps of regions with a $M \geq 5$ earthquake (gold star). **a**) Prague. **b**) Fairview. **c**) Pawnee. **d**) Cushing. Maps show seismicity scaled by magnitude (black crosses), faults (magenta lines) (Marsh and Holland 2016), county lines (white lines) wastewater disposal wells scaled by volume (1000s of bbl/mo, blue triangles), and wells that have been hydraulically fractured (red circles). Green boxes show the "on fault" regions for Figs. 8, 9, 10, and 11

we also compared these proximal events with seismicity in the surrounding geographical area (Fig. 8).

As noted in previous studies (Van der Elst et al. 2013; Keranen et al. 2014; Sumy et al. 2014), the 6 November 2011 Mw 5.7 Prague sequence included a Mw 5.0 foreshock on the same day that was preceded by several months of seismic quiescence. Our template matching supports this pattern (Fig. 9a). However, the SBCR19 catalog better identified the flurry of seismicity associated with the 27 February 2010 Mw 4.1 that occurred only ~1 km from the 2011 Mw 5.7 epicenter. The SBCR19 catalog identified five small ($M < 1.5$) foreshocks in the month before that mainshock and 240 aftershocks in the following month, while the initial catalog only has 15 aftershocks (Fig. 8a). Aftershocks from that event had died down in the months leading up to the 2011 Mw 5.7 mainshock, but the seismicity rate following the large Prague event remained high, regularly displaying more than 10 events per week through 2016 (Fig. 8a).

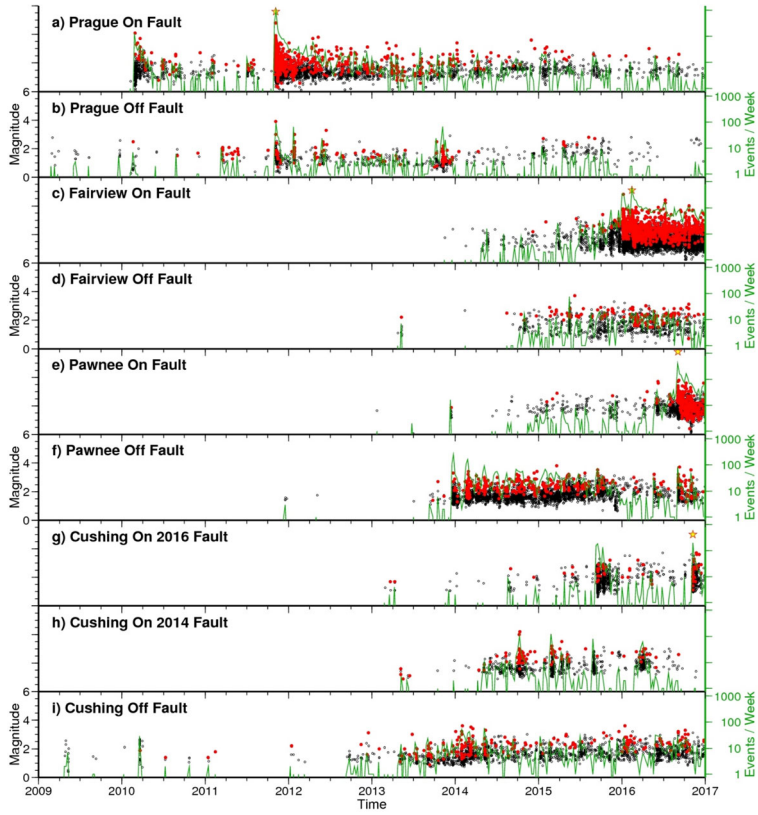

Fig. 8 Seismicity over time in regions with $M \geq 5$ earthquakes (stars). Seismicity is distinguished between events that occurred either on the same fault as the $M \geq 5$ event on or in the surrounding area (Fig. 7). Templates are represented as red dots, newly detected earthquakes are black dots, and log-scale seismicity rates are green lines

The 13 February 2016 Mw 5.1 Fairview earthquake had a robust foreshock sequence that led up to the mainshock (Fig. 9b). Much of this seismicity appears to be the aftershock sequence of the Mw 4.8 event on 7 January 2016. There were 12 earthquakes in the first week of January in the OGS catalog, as well as 24 in the year prior. In contrast, the SBCR19 catalog detected 664 earthquakes before the 7 January 2016 event, with more than 70 prior to the first cataloged event (Fig. 8c). We found that the seismic rates grew near the mainshock fault during 2014, followed months later by off-fault seismicity (Fig. 8d). The SBCR19 catalog also revealed an unusually slow rate of aftershock decay following the Mw 5.1 sequence from 12 to 10 events per day between 30 and 120 days after the mainshock (Fig. 9b). This observation is consistent with this being the most swarm-like large event sequence in Fig. 5b.

The 3 September 2016 Mw 5.8 Pawnee earthquake has been reported to have "nearly no foreshock activity" using the $M \geq 3$ earthquakes in the ANSS catalog (Yeck et al. 2017). While the OGS catalog only included three events in the month before the mainshock (Chen et al. 2017), the SBCR19 catalog identified over 100 events (Fig. 9c). Template matching results for events in the proximal fault area revealed ongoing seismicity as early as 2014 (Fig. 9e). Events away from the fault initiated in late 2013 and had an order of magnitude higher rates all the way through 2015 (Fig. 8f). In 2016, the off-fault seismicity rates decreased and then the on-fault seismicity rates exceeded those by June 2016.

Concern regarding the Cushing area grew following the Mw 4.0 and Mw 4.2 earthquakes on 7 and 10 October 2014 due to its proximity to critical oil and gas infrastructure (McNamara et al. 2015a). An improved seismic network was installed in the region and was able to help identify aftershocks of the 7 November 2016 Mw 5.0 event. However, the OGS catalog showed no events in the 4 months before the Mw 5.0 event (Fig. 9d), while the SBCR19 catalog was able to identify 43. Despite having access to the local network data, the OGS catalog produced a more irregular rate of aftershocks than the SBCR19 catalog that was strictly based on three regional stations (Fig. 9d). The OGS catalog included a swarm of 65 earthquakes with $M < 3$ approximately 1 year before the Mw 5.0, but the SBCR19 catalog found 1111 events during that swarm (Fig. 8g).

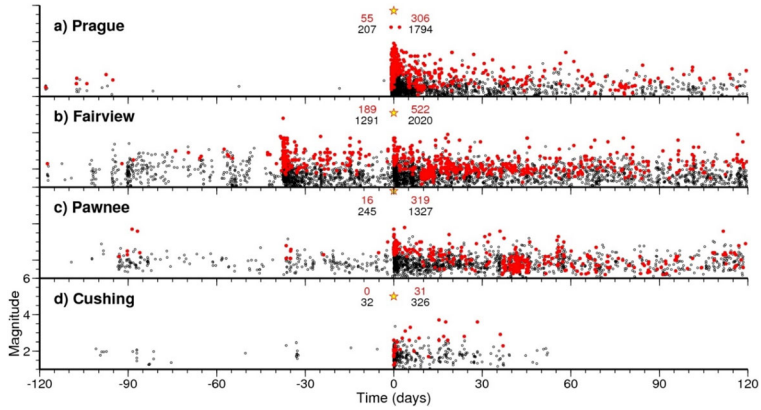

Fig. 9 Seismicity in regions associated with $M \geq 5$ earthquakes (Fig. 7) in the 4 months before and after the mainshock (stars). The original catalog is represented by red circles, and the SBCR19 catalogs are black circles. Numbers show the number of foreshocks (left) and aftershocks (right) for each catalog during the time frame plotted

Focusing on the 2014 fault region (Fig. 7d), we found three other similar but shorter swarms in addition to the 2014 sequences that produced the two Mw 4 events (Fig. 8h). The off-fault seismicity began earlier, with nine OGS catalog events and 82 events from the SBCR19 catalog prior to 2013 (Fig. 8i).

We then compared the aftershock patterns for the on-fault $M \geq 5$ sequence seismicity and calculated the Omori decay parameters. The total number of earthquakes in the SBCR19 catalog was approximately an order of magnitude larger than the OGS catalog, but the relative seismicity rates for the Prague and Cushing sequences were similar, resulting in p values ≈ 1 (Fig. 10). For the Fairview sequence, the SBCR19 catalog demonstrates a smoother aftershock decay that was more consistent with the Omori-Utsu prediction, but the maximum likelihood method still produces p values ≈ 0.6 for both catalogs. The largest p value discrepancy between the two catalogs was observed for the Pawnee sequence, where the SBCR19 catalog produced a p value ≈ 0.6, while the OGS catalog resulted in significantly different p value ≈ 0.9. We suggest that the SBCR19 catalog has captured more swarm behavior following the Pawnee sequence that resembles the Fairview sequence more than the Prague sequence. This demonstrated the potential for the SBCR19 catalog to produce more robust aftershock studies.

3.4 Comparing the SBCR19 catalog with poroelastic stress models

Understanding the influence that high-rate injection wells have on the stress state along faults is an important component of mitigating earthquake hazard, and numerous poroelastic stress models (e.g., Deng et al. 2016; Barbour et al. 2017; Goebel et al. 2017) have been proposed to a variety of induced earthquake sequences as a result. Considering that in situ stress measurements are uncommon, the validation of these models is frequently dependent on observed seismicity. A more complete earthquake catalog can be used to produce better models that offer greater insight into the poroelastic stress changes associated with fluid injection.

Figure 11 compares the OGS catalog and SBCR19 catalog with two recent stress models that investigated the Mw 5.1 Fairview and Mw 5.8 Pawnee earthquake sequences. Poroelastic stresses were estimated along the Fairview fault considering various diffusivities using the OGS catalog, which had events beginning in late 2014 (Goebel et al. 2017). We find that the SBCR19 catalog identified earthquakes back to early 2013, consistent with modeled poroelastic stress changes for all diffusivities the authors considered (Fig. 11a). Poroelastic stresses considering the influence of variable- and constant-rate injection were estimated along the Pawnee fault using a depth-dependent diffusivity model (Barbour et al. 2017). Using the SBCR19 catalog, the observed correlation between variable-rate injection and seismicity was further improved, supporting Barbour et al.'s (2017) conclusion that injection rate is an important contributor to seismicity rate (Fig. 11b).

3.5 Limitations of the SBCR19 catalog

As we primarily intended to investigate the seismicity on a statewide scale in this study, we associated newly detected events with the cataloged earthquake location. Using a high detection threshold of 15 times, the daily median absolute deviation has ensured that template matching can discern events from different source

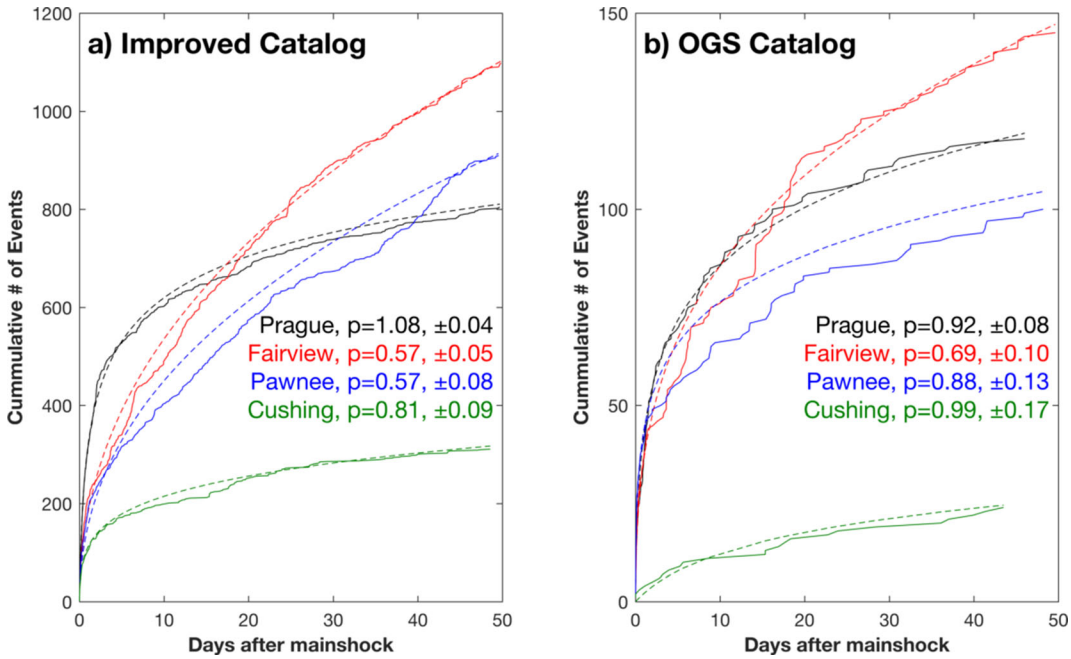

Fig. 10 Omori-Utsu aftershock decay p values showing the observed (solid lines) and predicted (dashed lines) for the Mw 5.7 Prague, Mw 5.1 Fairview, Mw 5.8 Pawnee, and Mw 5.0 Cushing sequences for the **a**) SBCR19 and **b**) OGS earthquake catalogs. Earthquakes used for each sequence are from the corresponding geographical areas in Fig. 7

regions (Skoumal et al. 2014). Using hundreds of stations in the Northern California Seismic Network, Schaff and Waldhauser (2005) analyzed the cross-correlation coefficients (CC) of earthquakes at each station and found that the CC values decrease precipitously as interevent distances exceed 1 km, due to the breakdown in waveform similarity with increasing differences in path. As a result, the majority of newly

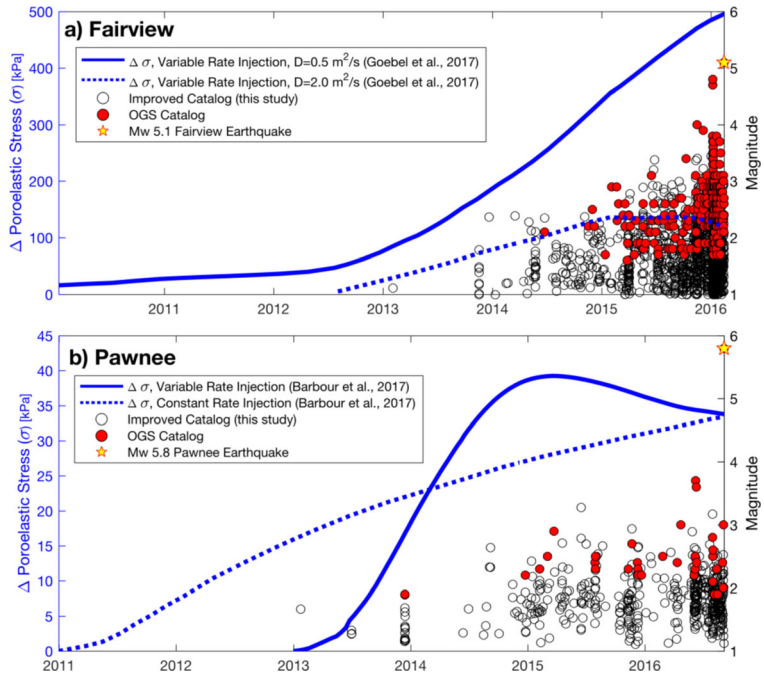

Fig. 11 Comparison of prior stress modeling work done at **a**) Fairview (Goebel et al. 2017) and **b**) Pawnee (Barbour et al. 2017) with the SBCR19 catalog and OGS earthquake catalogs. Earthquakes shown are from the corresponding geographical areas in Fig. 7. The modeled poroelastic stresses plotted were determined at the centroid of the Fairview seismicity (− 97.73 lon, 36.49 lat) in **a**) (Goebel et al. 2017) and the centroid of the Pawnee mainshock in **b**) (Barbour et al. 2017)

detected events are expected to be in relatively close proximity to the originally cataloged earthquakes and have a similar mechanism.

Our study focused on investigating the evolution of seismicity between 2008 and 2016 using the same three station network. With the installation of additional seismometers later in the study window, it is possible to use these data to further improve the completeness of the catalog. Additionally, the creation of a template for template matching relies on an earthquake being cataloged. It is very likely that small magnitude seismic sequences in Oklahoma and Kansas have lacked any cataloged earthquakes. If the maximum magnitude of a sequence is below the catalog M_C and the waveforms from these sequences are disparate from any of the waveforms from the cataloged earthquakes, these sequences would remain undetected. A solution to this issue could be the implementation of repeating earthquake detection algorithms that do not rely on a starting catalog that can process large volumes of data (Yoon et al. 2015; Skoumal et al. 2016).

Our regional swarminess, p value, and CoV analyses used geographical bins to evaluate seismicity. We found that the results were comparable after testing with multiple different spatiotemporal parameters to subdivide the seismicity catalog (Figs. S3–S4). However, spatiotemporal clusters of individual sequences could be identified in a more manual process, similar to what we did for the $M \geq 5$ sequences. This would address the limitation of our analyses that inadvertently could have included multiple sequences on significantly different faults within the same spatial bin.

4 Summary

This study has sought to utilize earthquake template matching to help investigate the seismicity induced by human activities in Oklahoma and southern Kansas. Applying template matching to earthquakes from the OGS and ANSS catalogs resulted in a ~ 10-fold increase in events. The majority of these earthquakes demonstrated swarm-like behavior, consistent with other fluid-injection-induced seismicity studies. The rapid increase in seismicity rate of these swarms in 2013 coincided with a reduction in the calculated p values both before and after larger magnitude events, and the CoV of interevent times in small spatial bins tend to spatially correlate with the location of $M \geq 4$ earthquakes. We also illustrated the importance of having a low M_C when investigating relatively large $M \geq 5$ earthquake sequences. For example, if only the existing ANSS catalog is used, the 2016 M 5.8 Pawnee sequence appears to lack foreshocks. The SBCR19 catalog detected over 2 years of productive precursory earthquakes with an abrupt seismicity rate increase beginning 3 months before the Pawnee mainshock. In addition, the aftershock productivity of the Pawnee event was more similar to the Fairview sequence with the SBCR19 catalog relative to the ANSS catalog. When compared with poroelastic modeling, the SBCR19 catalog improved the correlation between seismicity and stress accounting for variable-rate injection, supporting the conclusion that injection rate is a primary contributor to seismicity rate.

The SBCR19 catalog can be found in the electronic supplement.

Acknowledgments All waveform data were obtained through the Incorporated Research Institutions for Seismology Data Management Center (IRIS) (www.iris.edu; last accessed April 2017). Earthquake catalogs were obtained from the U.S. Geological Survey (USGS) National Earthquake Information Center (NEIC) (http://earthquake.usgs.gov/data; last accessed April 2017) and the Oklahoma Geological Survey (http://www.ou.edu/ogs.html; last accessed April 2017). Well data were retrieved from FracFocus (https://fracfocus.org; last accessed April 2017). This research benefited from discussions with A. Barbour, E. Cochran, H. DeShon, W. Ellsworth, P. Friberg, O. Kaven, A. Michael, J. Norbeck, W. Rish, J. Rubinstein, D. Shelly, and D. Sumy. This manuscript was improved thanks to G. Beroza, R. Habiger, A. Holland, A. McGarr, J. Rubinstein, and J. Shemeta. Our analysis relied heavily on Miami University's High Performance Computing, and we thank B. Koby and J. Mueller for their assistance.

Funding information Support for this work was provided by NSF (EAR-0847688) and USGS NEHRP (2015-0176, 2016-0124).

References

Aki K (1965) Maximum likelihood estimate of b in the formula log N = a - bM and its confidence limits. Bull Earthq Res Inst Univ Tokyo 43:237–239

Barbour AJ, Norbeck JH, Rubinstein JL (2017) The effects of varying injection rates in Osage County, Oklahoma, on the 2016 mw 5.8 Pawnee earthquake. Seismol Res Lett 88(5): 1040–1053

Chen X, Nakata N, Pennington C, Haffener J, Chang JC, He X, Zhan Z, Ni S, Walter JI (2017) The Pawnee earthquake as a

result of the interplay among injection, faults and foreshocks. Sci Rep 7(1):4945

Cochran ES, Ross ZE, Harrington RM, Dougherty SL, Rubinstein JL (2018) Induced earthquake families reveal distinctive evolutionary patterns near disposal wells. J Geophys Res Solid Earth 123:8045–8055. https://doi.org/10.1029/2018JB016270

Deng K, Liu Y, Harrington RM (2016) Poroelastic stress triggering of the December 2013 crooked Lake, Alberta, induced seismicity sequence. Geophys Res Lett 43:8482–8491

Ellsworth WL (2013) Injection-induced earthquakes. Science 341: 1225942. https://doi.org/10.1126/science.1225942

Ellsworth WL, Llenos AL, McGarr AF, Michael AJ, Rubinstein JL, Mueller CS, Petersen MD, Calais E (2015) Increasing seismicity in the US midcontinent: implications for earthquake hazard. Lead Edge 34(6):618–626

Goebel THW, Weingarten M, Chen X, Haffener J, Brodsky EE (2017) The 2016 Mw5. 1 Fairview, Oklahoma earthquakes: evidence for long-range poroelastic triggering at > 40 km from fluid disposal wells. Earth Planet Sci Lett 472:50–61

Gutenberg B, Richter CF (1944) Frequency of earthquakes in California. Bull Seismol Soc Am 34(4):185–188

Haffener J, Chen X, Murray K (2018) Multiscale analysis of spatiotemporal relationship between injection and seismicity in Oklahoma. J Geophys Res Solid Earth 123(10):8711–8731

Healy JH, Rubey WW, Griggs DT, Raleigh CB (1968) The Denver earthquakes. Science 161:1301–1310

Hincks T, Aspinall W, Cooke R, Gernon T (2018) Oklahoma's induced seismicity strongly linked to wastewater injection depth. Science 359(6381):1251–1255

Holtkamp SG, Brudzinski MR (2011) Earthquake swarms in circum-Pacific subduction zones. Earth Planet Sci Lett 305(1–2):215–225

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 83:250–260. https://doi.org/10.1785/gssrl.83.2.250

Kagan YY, Jackson DD (1991) Long-term earthquake clustering. Geophys J Int 104(1):117–134. https://doi.org/10.1111/j.1365-246X.1991.tb02498.x

Keranen KM, Savage HM, Abers GA, Cochran ES (2013) Potentially induced earthquakes in Oklahoma, USA: links between wastewater injection and the 2011 mw 5.7 earthquake sequence. Geology 41(6):699–702

Keranen KM, Weingarten M, Abers GA, Bekins BA, Ge S (2014) Sharp increase in Central Oklahoma seismicity since 2008 induced by massive wastewater injection. Science 345(6195):448–451

Langenbruch C, Zoback MD (2016) How will induced seismicity in Oklahoma respond to decreased saltwater injection rates. Sci Adv 2(11):e1601542. https://doi.org/10.1126/sciadv.1601542

Marsh SM and Holland AA (2016). Comprehensive fault database and interpretive fault map of Oklahoma. Oklahoma Geological Survey Open File Report OF2–2016

Massey FJ Jr (1951) The Kolmogorov-Smirnov test for goodness of fit. J Am Stat Assoc 46(253):68–78

McGarr A, Barbour AJ (2017) Wastewater disposal and the earthquake sequences during 2016 near Fairview, Pawnee, and Cushing, Oklahoma. Geophys Res Lett 44(18):9330–9336

McGarr A, Bekins B, Burkardt N, Dewey J, Earle P, Ellsworth W, Ge S, Hickman S, Holland A, Majer E, Rubinstein J (2015) Coping with earthquakes induced by fluid injection. Science 347(6224):830–831

McNamara DE, Hayes GP, Benz HM, Williams RA, McMahon ND, Aster RC, Holand A, Sickbert T, Herrmann R, Briggs R, Smoczyk G, Bergman E, Earle P (2015a) Reactivated faulting near Cushing, Oklahoma: increased potential for a triggered earthquake in an area of United States strategic infrastructure. Geophys Res Lett 42(20):8328–8332

McNamara DE, Rubinstein JL, Myers E, Smoczyk G, Benz HM, Williams RA, Hayes G, Wilson DC, Herrmann RB, McMahon ND, Aster RC, Bergman E, Holland A, Earle P (2015b) Efforts to monitor and characterize the recent increasing seismicity in Central Oklahoma. Lead Edge 34(6): 628–639

Miller LH (1956) Table of percentage points of Kolmogorov statistics. J Am Stat Assoc 51(273):111–121

Petersen MD, Mueller CS, Moschetti MP, Hoover SM, Shumway AM, McNamara DE, Williams R, Llenos AL, Ellsworth WL, Rubinstein JL, McGarr AF, Rukstales KS (2017) 2017 one-year seismic-hazard forecast for the central and eastern United States from induced and natural earthquakes. Seismol Res Lett 88(3):772–783

Rubinstein JL, Ellsworth WL, Dougherty SL (2018a) The 2013 – present induced earthquake sequence in southern Kansas. Bull Seismol Soc Am 108:674–689. https://doi.org/10.1785/0120170209

Rubinstein JL, Ellsworth WL, Dougherty SL (2018b) The 2013–2016 induced earthquakes in Harper and Sumner counties, southern Kansas. Bull Seismol Soc Am 108(2):674–689

Schaff DP (2008) Semiempirical statistics of correlation-detector performance: bull. Seismol Soc Am 98:1495–1507. https://doi.org/10.1785/0120060263

Schaff DP, Waldhauser F (2005) Waveform cross-correlation-based differential travel-time measurements at the northern California seismic network. Bull Seismol Soc Am 95(6): 2446–2461

Schoenball M, Ellsworth WL (2017) A systematic assessment of the spatiotemporal evolution of fault activation through induced seismicity in Oklahoma and southern Kansas. J Geophys Res Solid Earth 122:10,189–10,206. https://doi.org/10.1002/2017JB014850

Shelly DR, Beroza GC, Ide S (2007) Non-volcanic tremor and low-frequency earthquake swarms. Nature 446(7133):305–307

Skoumal RJ, Brudzinski MR, Currie BS, Levy J (2014) Optimizing multi-station earthquake template matching through re-examination of the Youngstown Ohio, sequence. Earth Planet Sci Lett 405:274–280. https://doi.org/10.1016/j.epsl.2014.08.033

Skoumal RJ, Brudzinski MR, Currie BS (2015a) Earthquakes induced by hydraulic fracturing in Poland township, Ohio. Bull Seismol Soc Am 105(1):189–197. https://doi.org/10.1785/0120140168

Skoumal RJ, Brudzinski MR, Currie BS (2015b) Distinguishing induced seismicity from natural seismicity in Ohio: demonstrating the utility of waveform template matching. J

Geophys Res Solid Earth 120:6284–6296. https://doi.org/10.1002/2015JB012265

Skoumal RJ, Brudzinski MR, Currie BS (2016) An efficient repeating signal detector to investigate earthquake swarms. J Geophys Res 121:5880–5897. https://doi.org/10.1002/2016JB012981

Skoumal RJ, Ries R, Brudzinski MR, Barbour AJ, Currie BS (2018a) Earthquakes induced by hydraulic fracturing are pervasive in Oklahoma. J Geophys Res Solid Earth 123:10, 918–10,935. https://doi.org/10.1029/2018JB016790

Skoumal RJ, Brudzinski M, Currie B (2018b) Proximity of Precambrian basement affects the likelihood of induced seismicity in the Appalachian, Illinois, and Williston basins. Geosphere 14(3):1365–1379. https://doi.org/10.1130/GES01542.1

Skoumal RJ, Kaven J, Walter J (2019) Characterizing seismogenic fault structures in Oklahoma using a relocated template matched catalog. Seismol Res Lett 90(4):1535–1543. https://doi.org/10.1785/0220190045

Sumy DF, Cochran ES, Keranen KM, Wei M, Abers GA (2014) Observations of static coulomb stress triggering of the November 2011 M5. 7 Oklahoma earthquake sequence. J Geophys Res Solid Earth 119(3):1904–1923. https://doi.org/10.1002/2013JB010612

Trabant C, Hutko AR, Bahavar M, Karstens R, Ahern T, Aster R (2012) Data products at the IRIS DMC: stepping stones for research and other applications. Seismol Res Lett 83(5):846–854. https://doi.org/10.1785/0220120032

Utsu T, Ogata Y, Matsu'ura RS (1995) The centenary of the Omori formula for a decay of aftershock activity. J Phys Earth 43:1–33

Van der Elst NJ, Savage HM, Keranen KM, Abers GA (2013) Enhanced remote earthquake triggering at fluid-injection sites in the midwestern United States. Science 341(6142):164–167

Vidale JE & Shearer PM (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 Research: Solid Earth, 111

Walsh FR, Zoback MD (2015) Oklahoma's recent earthquakes and saltwater disposal. Sci Adv 1(5):e1500195

Walter JI, Dotray PJ, Frohlich C, Gale JF (2016) Earthquakes in Northwest Louisiana and the Texas–Louisiana border possibly induced by energy resource activities within the Haynesville shale play. Seismol Res Lett 87(2a):285–294. https://doi.org/10.1785/0220150193

Walter JI, Chang JC, Dotray PJ (2017a) Foreshock seismicity suggests gradual differential stress increase in the months prior to the 3 September 2016 M w 5.8 Pawnee earthquake. Seismol Res Lett 88(4):1032–1039

Walter JI, Chang JC, Dotray PJ (2017b) Foreshock seismicity suggests gradual differential stress increase in the months prior to the 3 September 2016 mw 5.8 Pawnee earthquake. Seismol Res Lett 88(4):1032–1039. https://doi.org/10.1785/0220170007

Weingarten M, Ge S, Godt JW, Bekins BA, Rubinstein JL (2015) High-rate injection is associated with the increase in US mid-continent seismicity. Science 348(6241):1336–1340

Wiemer S, Wyss M (2000) Minimum magnitude of completeness in earthquake catalogs: examples from Alaska, the Western United States, and Japan. Bull Seismol Soc Am 90(4):859–869

Woessner J, Wiemer S (2005) Assessing the quality of earthquake catalogues: estimating the magnitude of completeness and its uncertainty. Bull Seismol Soc Am 95(2):684–698

Yeck WL, Sheehan AF, Benz HM, Weingarten M, Nakai J (2016a) Rapid response, monitoring, and mitigation of induced seismicity near Greeley, Colorado. Seismol Res Lett 87(4):837–847

Yeck WL, Weingarten M, Benz HM, McNamara DE, Bergman EA, Herrmann RB, Rubinstein JL, Earle PS (2016b) Far-field pressurization likely caused one of the largest injection induced earthquakes by reactivating a large preexisting basement fault structure. Geophys Res Lett 43(19):10–198

Yeck WL, Hayes GP, McNamara DE, Rubinstein JL, Barnhart WD, Earle PS, Benz HM (2017) Oklahoma experiences largest earthquake during ongoing regional wastewater injection hazard mitigation efforts. Geophys Res Lett 44(2):711–717

Yoon C, O'Reilly O, Bergen K, Beroza G (2015) Earthquake detection through computationally efficient similarity search. Sci Adv 1:e1501057

Zhang Y, Person M, Rupp J, Ellett K, Celia MA, Gable CW, Bowen B, Evans J, Bandilla K, Mozley P, Dewers T (2013) Hydrogeologic controls on induced seismicity in crystalline basement rocks due to fluid injection into basal reservoirs. Groundwater 51(4):525–538