Characterizing Seismogenic Fault Structures in Oklahoma Using a Relocated Template-Matched Catalog

by Robert J. Skoumal, J. Ole Kaven, and Jacob I. Walter

ABSTRACT

Oklahoma is one of the most seismically active places in the United States as a result of industry activities. To characterize the fault networks responsible for these earthquakes in Oklahoma, we relocated a large-scale template-matching catalog between 2010 and 2016 using the GrowClust algorithm. This relocated catalog is currently the most complete statewide catalog for Oklahoma during this seven year window. Using this relocated catalog, we identified seismogenic fault segments by developing an algorithm (FaultID) that clusters earthquakes and then identifies linear trends within each cluster. Considering the large number of earthquakes in Oklahoma, this algorithm made the process of identifying previously unmapped seismogenic faults more approachable and objective. We identify approximately 2500 seismogenic fault segments that are in general agreement with focal mechanisms and optimally oriented relative to maximum principle stress measurements. We demonstrate that these fault orientations can be used to approximate the maximum principle stress orientations.

Supplemental Content: Relocated earthquake catalog in Oklahoma between 2010 and 2016 and a table of the seismogenic fault segments identified with the FaultID algorithm.

INTRODUCTION

The seismicity rate in the central and eastern United States increased 40-fold within the past decade predominantly as a result of human activities (e.g., Ellsworth, 2013; van der Baan and Calixto, 2017). Oklahoma, the state with the most prominent seismicity rate increases, had more cataloged M > 3 earthquakes than California during the same period between 2010 and 2018. The vast majority of these induced earthquakes in Oklahoma occurred along unmapped strike-slip faults in the upper Precambrian basement (e.g., McNamara et al., 2015; Keranen and Weingarten, 2018). Although the seismicity and corresponding faults of some of the more prominent sequences have been identified by prior work (e.g., Keranen et al., 2013; Chen et al., 2017), the majority of the numerous smaller seismogenic faults that are ubiquitous throughout the state remained uncharacterized.

Schoenball and Ellsworth (2017) relocated the previously cataloged earthquakes in northern Oklahoma and southern Kansas between May 2013 and November 2016 using standard location (HYPOINVERSE-2000; Klein, 2014) and relocation tools (hypoDD; Waldhauser and Ellsworth, 2000). Because phase picks from the Oklahoma Geological Survey (OGS) were not published at the time, Schoenball and Ellsworth (2017) estimated phase arrivals for the OGS-cataloged events using an automatic phase detection approach and were able to refine the locations of 11,997 earthquakes in Oklahoma.

In this study, we improved the statewide seismicity catalog by relocating a template-matching catalog using manual phase picks from the OGS between 2010 and 2016 and the GrowClust (Trugman and Shearer, 2017) relocation algorithm. We then developed an algorithm (FaultID) to identify linear clusters of seismicity and applied it to our relocated catalog to characterize previously unmapped seismogenic faults. Because nearly all of the seismicity occurred on previously unmapped faults, these improved earthquake and fault catalogs could help improve our understanding of tectonic variations, seismic hazard estimates, and processes that induce seismicity in Oklahoma.

METHODS

Earthquake Relocation

Previous work relocated template-matched earthquake catalogs to identify fine-fault networks (e.g., Shelly and Hill, 2011; Shelly et al., 2013, 2016). Here, we use the improved Oklahoma earthquake catalog (SBC16; Skoumal et al., 2016) that was generated by applying large-scale template matching to all 23,889 earthquakes cataloged by the OGS (master events) in Oklahoma between 16 October 2008 and 31 December 2016 using three regional seismometers. This catalog used a conservative detection threshold of 15 times the daily median absolute deviation of the network normalized cross-correlation coefficients. This improved earthquake catalog contains 209,409 earthquakes with...
Figure 1. All 64,236 relocated Oklahoma earthquakes from this study (pluses). Rectangles denote areas used in Figures 3-6, 8. Triangles represent seismometers used for earthquake location within this region. Filled triangles denote the three stations used to generate the original template-matched earthquake catalog (SBC16; Skoumal et al., 2016). The color version of this figure is available only in the electronic edition.

Using these lag and correlation coefficients between event pairs, earthquakes are relocated with the GrowClust algorithm (Trugman and Shearer, 2017) using the OGS 1D velocity model (Darold et al., 2015). GrowClust has a similar objective to other relative relocation algorithms (e.g., hypoDD; Waldhauser and Ellsworth, 2000), but has some algorithmic advantages (Trugman and Shearer, 2017). GrowClust utilizes a hierarchical cluster algorithm that relocates earthquakes within similar clusters that allow larger earthquake catalogs, such as our improved Oklahoma catalog, that allows large data sets to be processed more efficiently. The algorithm uses the L1 norm, allowing for more robust misfit criteria than the standard least-squares approach. A maximum station distance of 80 km and a maximum root mean square (rms) differential time residual of 0.2 s for cluster merger is used in the GrowClust algorithm. To determine location uncertainties, 100 bootstrap iterations of GrowClust’s nonparametric uncertainty estimation algorithm are performed. To be considered in later analyses, relocated earthquakes must have P- and S-phase rms residual differential times less than 0.2 s, contain five or more events in its respective GrowClust branch, and five or more phase differential times used in the relocation. Our final relocated catalog (SKW19) contains 64,236 events with reported 1σ horizontal and vertical relative location uncertainties of ~600 and 570 m, respectively. The relocated Oklahoma seismicity catalog (SKW19) and the linear fault segments that were identified with FaultID in this study are included in the supplemental content to this article.

Identification of Seismogenic Faults

Using our SKW19 relocated catalog of 64,236 earthquakes, we identify linear trends of seismicity that we then interpret to represent seismogenic faults. To make the process of identifying fault segments more easily accessible and objective, we develop an algorithm (FaultID) that iteratively identifies clusters of seismicity and distinguishes linear features within each cluster (Fig. 2). Previous work focused on approaches to identify seismogenic fault planes using a variety of methods ranging in complexity and amount of manual parameter selection (e.g., Ouillon et al., 2008; Kaven and Pollard, 2013; Wang et al., 2013). Here, our goal was to create a simpler method for rapidly identifying linear trends of seismicity in Oklahoma using popular, efficient clustering and point-fitting algorithms.

Earthquakes are clustered based on their horizontal spatial location using the density-based spatial clustering of applications with noise (DBSCAN) algorithm (Esté et al., 1996). With this approach, an earthquake is classified as a core event if there are at least N other earthquakes within the Euclidean distance D. Any earthquakes that are within distance D from a core point and had fewer than N neighbors within distance D are considered border events. If a core event is within distance D from another core event, they are considered to be in the same cluster along with any corresponding border events. We repeat this clustering approach five times with the number of neighbors N represented by the numbers in the set [1000, 500, 250, 100, 50].
Focal Mechanisms and Stress Orientations

We attempt to relate the fault orientation to the local stress orientations by comparing fault trends to focal mechanisms and borehole stress measurements. As described previously, the earthquake locations are the only data used as input in the FaultID algorithm. Because the FaultID results are created independently from focal mechanisms and stress orientations, we use these other independent measures to evaluate the identified seismogenic fault segments. We focus this evaluation on two areas in Oklahoma: (1) the area around the Cushing earthquake and (2) the area around Grant County. The Cushing area is selected because it is the only 5.0 earthquake that has not yet been well characterized by other seismological studies. The Grant County area is selected because it contains the highest density of identified faults as well as numerous fault orientations that were poorly oriented to a reported principal stress $S_{H\text{max}}$ value (described in the Seismogenic Fault Identification section).

We compute focal mechanisms with the HASH algorithm (Hardebeck and Shearer, 2002) using manually picked P-wave polarities and the same velocity model that was used in the relocations. Only focal mechanisms for events with a minimum of eight phase picks and a maximum azimuthal gap of 90° are considered. This produces 12 focal mechanisms in the Cushing area and 22 focal mechanisms for the southern Grant County area.

We compare our fault orientations against nearby $S_{H\text{max}}$ orientations from Alt and Zoback (2016). Only stress orientations determined from drilling-induced tensile fractures (DITFs) with reported qualities of A or B were considered. For B quality $S_{H\text{max}}$ orientations, at least six DITFs with a combined length $>100$ m and standard deviations of orientations $\leq 20^\circ$.
had to have been identified in a single well; for A quality orientations, DITFs had to have a combined length >300 m with a standard deviation of ≤12° (Alt and Zoback, 2016).

We also estimate the approximate stress orientations using the seismogenic faults identified by FaultID in the Grant County and west Stillwater areas (Fig. 1). With this approach, we assume that seismogenic faults would be preferentially distributed 30° from $S_{H_{max}}$, which is what we would expect for vertical strike-slip faults and a coefficient of friction of $\sim 0.6$.

Using 0.1° geographical bins with a horizontal step interval of 0.0125°, we calculate the median fault trend weighted by the size of the respective faults in each bin. For the fault trend to be calculated in a geographical bin, we require a minimum of 10 fault segments with a combined length of at least 4 km. To estimate the fault trend uncertainties with this approach, we perform 100 jackknife iterations and randomly remove 10% of the faults in each trial.

**RESULTS AND DISCUSSION**

**Pawnee Catalog Comparison**

Using both local and regional seismometers, Chen et al. (2017) relocated all OGS-cataloged earthquakes in the Pawnee area (Fig. 1) along with some additional events that were manually identified. These earthquakes were cross correlated to improve differential phase picks and relocated using a 3D velocity model with hypoDD. The final catalog includes 950 earthquakes between 13 December 2013 and 19 October 2016 with a reported relative horizontal and vertical location uncertainties of $\sim 35$ and 200 m, respectively. Within this time window, we would expect the Chen et al. (2017) catalog to be the most accurate locations currently available for the Pawnee area. We compare our catalog to (1) the OGS, (2) Chen et al. (2017), and (3) Schoenball and Ellsworth (2017) catalogs using the same time window and spatial area considered by Chen et al. (2017) (Fig. 3). We find that the number of earthquakes in their relocated template-matching catalog was comparable to both the OGS and Schoenball and Ellsworth (2017) catalogs prior to the mainshock (Fig. 3e). Our catalog contains over twice the number of earthquakes as Chen et al. (2017), albeit with higher location uncertainties. Despite using a simple 1D velocity model, our regional relocated catalog still does a comparable job at identifying the similar earthquake locations as seen in refined catalog of Chen et al. (2017).

**Cushing Catalog Comparison**

Approximately 3 km west of Cushing, Oklahoma, a $M_w$ 5.0 earthquake occurred on 7 November 2016. In the Cushing area (Fig. 1), the Schoenball and Ellsworth (2017) catalog contains 155 earthquakes between 31 August 2014 and 22
November 2016. We compare the catalogs from (1) the OGS catalog, (2) Schoenball and Ellsworth (2017), and (3) this study during this time window (Fig. 4). To our knowledge, these are the three most complete earthquake catalogs of this area and no detailed analysis has previously been done to characterize the corresponding faults in this area. Our catalog contains over four times as many located earthquakes as the other two catalogs (Fig. 4d). Similar to the Pawnee example, the refined catalogs constrain the relatively diffusive OGS-cataloged locations and allow fault planes to be identified. The earthquake locations in our catalog are similar to Schoenball and Ellsworth (2017), but our catalog allows additional faults to be identified, illuminating a more complex fault system that was not previously documented.

Seismogenic Fault Identification

In this study, 2492 seismogenic fault segments are identified by FaultID using the SKW19 catalog. These fault segments represent a total length of ∼826 km (average segment length of ∼0.33 km), 49,302 earthquakes (∼77% of the SKW19 earthquake catalog) were associated with these faults, and faults were represented by an average of ∼19 earthquakes.

Here, we present examples of the seismogenic faults identified by the FaultID algorithm in the Cushing (Fig. 5) and the southern Grant County areas (Fig. 6). The focal mechanisms determined in our analysis are in general agreement with the orientations of our algorithmically identified faults. Although the automatically identified faults are imperfect, and even erroneous in some cases, the majority of fault segments represent...
the trends of cataloged seismicity that we would have manually selected. In general, the orientations of seismogenic strike-slip faults identified by FaultID are \(\sim 30^\circ\) away from the regional stress field (Fig. 7). This result is expected because injection-induced seismicity has widely been observed along optimally oriented faults (e.g., Skoumal et al., 2015; Keranen and Weingarten, 2018), although a minority of suboptimally oriented faults have also hosted induced earthquakes (e.g., Frohlich et al., 2014; Keranen and Weingarten, 2018).

Although most of our geographically binned seismogenic faults are optimally oriented with respect to \(S_{\text{H max}}\) measurements, the southern Grant County area (centroid of approximately \(-97.7^\circ\ E, 36.7^\circ\ N\)) is the major exception (Fig. 6); an A quality \(S_{\text{H max}}\) measurement of N79\(^\circ\)E is similar to the median fault trend azimuth of N69\(^\circ\)E in the surrounding area. Local stress rotations in this area may help explain this disagreement. Two additional A quality \(S_{\text{H max}}\) reports of N68\(^\circ\)E and N59\(^\circ\)E were \(\sim 20\) km to the east and southeast, respectively (Fig. 8a). In support of these \(S_{\text{H max}}\) measurements, we observed a similar rotation in the seismogenic fault azimuths (Fig. 8a) of \(\sim 20^\circ\) over the same 20 km area (Fig. 8a,c). This rotation is in contrast with the majority of stress measurements and seismogenic faults elsewhere in the state where little variation or deviation is observed (Fig. 8b,d). Jackknife uncertainty estimates (Fig. 8c,d) tended to be \(< 5^\circ\), suggesting reasonably consistent fault orientations in individual geographical bins, although some bins with relatively few faults had larger errors (\(< 10^\circ\)). The cause of this local stress rotation could potentially be associated with the nearby Nemaha ridge. Another potential explanation could be that there are elevated pore pressures due to the Nemaha fault acting as a no-flow boundary for injected fluids. If this was the case, larger pore pressures may have allowed for slip along faults that were suboptimally oriented, although this has yet to be observed in this particular case.

**Limitations**

Absolute earthquake locations could be better refined using local or 3D velocity models. Although the locations are a significant improvement over the OGS and SBC16 regional catalogs, local seismicity analyses may still be inhibited by the relatively large absolute location errors. Considering the general agreement between fault orientations with \(S_{\text{H max}}\) and focal mechanisms, we suggest the fault catalog is generally reliable for regional-scale analyses. However, because the seismogenic fault segments were algorithmically determined, some erroneously identified faults should be expected. Although these automatically identified faults may inform broad fault characteristics across the region, care should be taken when using these results for local-scale assessments.
The faults identified in this study are intended to characterize the previously unidentified smaller strike-slip fault segments that are ubiquitous in the Precambrian basement throughout Oklahoma. Some of the large magnitude sequences already have been constrained by local investigations (e.g., the $M_w$ 5.8 Pawnee earthquake; Chen et al., 2017), whereas other faults hosting significant seismicity have yet to be studied in detail (e.g., the $M_w$ 5.0 Cushing earthquake; Fig. 5). The faults of productive, larger magnitude sequences could be constrained by fault planes, beyond the fault trends investigated in this study. In addition, although our method is optimized for vertically oriented faults, subvertical faults would be better characterized by planes rather than line segments and could be the focus of future work.

CONCLUSIONS

Using a large-scale template-matching catalog that was applied to cataloged earthquakes in Oklahoma (Skoumal et al., 2016), earthquakes identified between 1 January 2010 and 31 December 2016 were relocated using GrowClust. This relocated Oklahoma template-matched seismicity catalog (SKW19) is currently the most complete statewide catalog for Oklahoma during this 6 yr window. Using this relocated catalog, we identified seismogenic fault segments by developing an algorithm (which we call FaultID) that spatially clusters earthquakes and then searches for linear trends within each cluster. Considering the large number of earthquakes in Oklahoma, the FaultID algorithm made the process of identifying previously unmapped seismogenic faults more approachable and objective. The trends of these automatically identified faults were in general agreement with $S_{H\text{max}}$ measurements, and we suggest that these faults could aid our approximations of stress orientations in areas that lack borehole measurements.

DATA AND RESOURCES


ACKNOWLEDGMENTS

This study benefited from discussions with Sara Dougherty and Justin Rubinstein. Daniel Trugman provided guidance regarding the GrowClust algorithm that improved the quality of this study. Andrew Barbour, Elizabeth Cochran, Marcus Herrmann, Justin Rubinstein, and an anonymous reviewer provided constructive peer reviews that improved the article. The authors thank the Oklahoma Geological Survey (OGS, http://www.ou.edu/ogs.html). All websites were last accessed in September 2018.

Figure 7. Rose diagrams of the fault segment orientations in geographical bins. Black bars denote the frequency of a given fault orientation in 5° azimuthal bins. Within each rose diagram, $n$ and $d$ represent the number of fault segments and the summed length (in kilometers) of fault segments in the corresponding geographical bin, respectively. Only the geographical bins in this area with >20 fault segments are shown. Dashed lines represent the median $S_{H\text{max}}$ measurements (Alt and Zoback, 2016) within each bin, and solid lines represent the azimuths corresponding to optimal orientations for slip (assuming a 30° deviation from $S_{H\text{max}}$). The color version of this figure is available only in the electronic edition.
REFERENCES


Figure 8. Azimuthal deviations of the median fault trends from the reference $S_{H_{max}}$ measurement (denoted as $R$). (a) The Grant County area, with $R = N79^\circ E$, demonstrates a local stress rotation. Two additional A quality $S_{H_{max}}$ measurements (Alt and Zoback, 2016) are denoted by 1 and 2. (b) The west Stillwater area, with $R = N86^\circ E$, does not demonstrate a significant change in stress orientation. In (a) and (b), the expected azimuthal deviation of faults trends at the location of each $S_{H_{max}}$ measurement is represented by the corresponding symbols in the respective color bars. The estimated azimuthal uncertainties from 100 jackknife for the Grant County and west Stillwater areas are shown in (c) and (d), respectively. Dotted lines represent the locations of previously mapped faults (Marsh and Holland, 2016). The color version of this figure is available only in the electronic edition.


Robert J. Skoumal  
J. Ole Kaven  
U.S. Geological Survey  
Earthquake Science Center  
345 Middlefield Road  
Menlo Park, California 94025 U.S.A.  
rskoumal@usgs.gov  
okaven@usgs.gov

Jacob I. Walter  
Oklahoma Geological Survey  
University of Oklahoma  
100 E. Boyd Street  
Norman, Oklahoma 73019 U.S.A.  
jwalter@ou.edu

Published Online 29 May 2019