Seismicity During the Initial Stages of the Guy-Greenbrier, Arkansas, Earthquake Sequence

Clara E. Yoon[1], Yihe Huang[1,2], William L. Ellsworth[1], and Gregory C. Beroza[1]

[1]Department of Geophysics, Stanford University, Stanford, California, USA.

[2]Department of Earth and Environmental Sciences, University of Michigan, Ann Arbor, Michigan, USA.

Key Points:

- We detected and located microearthquakes from first 3 months of the Guy-Greenbrier sequence in 2010 with a sparse 3-station seismic network
- Most events ($-1.5 \leq M_L \leq 2.9$) in June to September 2010 were induced by hydraulic fracturing at some but not all stimulated wells
- Initial seismic activity on the Guy-Greenbrier Fault was induced by wastewater injection starting in July 2010

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/2017JB014946

Corresponding author: Clara E. Yoon, ceyoon@stanford.edu

Abstract

We analyze the background seismicity, initiation, and earliest stages of the Guy-Greenbrier, Arkansas, earthquake sequence, which was potentially induced by wastewater injection starting in July 2010, during the 3-month time period 2010-06-01 to 2010-09-01. High-resolution observations of low-magnitude seismicity, and the high-quality Arkansas public well database, facilitate detailed analysis of spatial and temporal correlations between earthquakes, wastewater injection, and hydraulic fracturing. We detected 14,604 earthquakes, with magnitudes $-1.5 \leq M_L \leq 2.9$, using two sensitive, waveform-similarity-based event detection methods in parallel: Fingerprint And Similarity Thresholding (FAST), and template matching. We located the 1,740 largest earthquakes that form 16 spatially compact clusters, using P and S phases from 3 stations with the double-difference relocation algorithm and an improved velocity model constrained by the location of quarry blasts. We enhanced the temporal resolution of these event clusters by assigning smaller unlocated events to a cluster based on waveform similarity. Most clustered earthquakes during this time were both spatially and temporally correlated with hydraulic fracturing stimulation at several production wells. For one cluster, microseismicity was correlated with individual stages of stimulation. Many other wells had no detectable nearby seismicity during stimulation. We found a smaller number of events located on the Guy-Greenbrier Fault that were likely induced by wastewater injection. The concurrent presence of seismicity induced by hydraulic fracturing and wastewater injection presents a challenge for attribution and seismic hazard characterization, but the combination of precision seismology and high-quality well information allows us to disentangle the effects of these two processes.

1 Introduction

Since 2009, the central and eastern United States, an intraplate region with historically low levels of seismicity, has experienced a striking increase in earthquake activity, including several damaging earthquakes greater than magnitude 5 [*Ellsworth et al.*, 2013; *Rubinstein and Mahani*, 2015]. Many of these earthquakes, especially the larger ones, are thought to have been induced by deep injection of large volumes of wastewater produced by oil and gas operations over several years [*Ellsworth et al.*, 2013; *Rubinstein and Mahani*, 2015; *Walters et al.*, 2015]. Increased pore fluid pressure from injection can reduce the effective normal stress across a preexisting fault close to failure, unclamping it and allowing it to slip [*Healy et al.*, 1968; *Raleigh et al.*, 1976]. The Guy-Greenbrier area in central Arkansas (Figure 1), where hydraulic fracturing was used to increase natural gas production in the Fayetteville Shale, experienced several moderate strike-slip earthquakes: M_w 4.0 in October 2010, then M_w 4.1 on 2011-02-19, and finally the largest earthquake with M_w

4.7 (yellow star) on 2011-02-27 [*Horton*, 2012]. These earthquakes were part of an intense sequence that lasted over a year. They were reported to start in July 2010 following injection of wastewater at Well 1 (Figure 1, inverted triangle), and migrated southwest over the next few months, illuminating a previously unknown ~13-km long, near-vertical fault with strike ~N30°E, subsequently named the Guy-Greenbrier Fault for the nearby towns [*Horton*, 2012]. After the M_w 4.7 earthquake, injection stopped at the wells nearest the fault in March 2011 on an emergency order from the Arkansas Oil and Gas Commission (AOGC) [*Horton*, 2012]. The seismicity promptly decreased but remained higher than the background seismicity rate before the sequence for at least the next 7 months [*Huang and Beroza*, 2015].

We perform a retrospective analysis to understand how the Guy-Greenbrier earthquake sequence initiated, and to determine whether it was induced by wastewater injection. *Ogwari et al.* [2016] detected and located earthquakes in the first 4 months of the sequence, starting from the onset of wastewater injection at Well 1 on 2010-07-07, to 2010-10-20. Their improved catalog, complete down to M 0.2 and containing events down to M -0.6, revealed seismicity that started in the shallow (2-4 km depth) sedimentary formation below injection Well 1, and migrated southwest and down into the basement (deeper than 4 km) from September to October 2010. However, *Ogwari et al.* [2016] found only scattered seismicity without any particular spatial or temporal characteristics during the time immediately following injection, from 2010-07-07 to the end of August 2010. We chose to study the 3-month time period from 2010-06-01 to 2010-09-01. This includes the month before injection started at Well 1, which should help us understand background seismicity in the region, as well as the two months right after the start of injection, so that we can characterize the earliest stages of seismicity occurring in response to injection. The Advanced National Seismic System (ANSS) catalog contains only 75 events during these 3 months (Data Set S1), with uncertain locations, and few of them near the soon-to-be activated Guy-Greenbrier Fault (Figure S1). Similar off-fault locations were seen for events located with the regional Cooperative New Madrid Seismic Network (CNMSN) during this time [*Horton*, 2012]. We detect and locate as many small earthquakes as possible from continuous seismic data for these 3 months using a sparse 3-station network (Figure 1, black triangles), then explore spatial and temporal correlations between the seismicity and unconventional hydrocarbon development.

We also consider the possibility that earthquakes in the Guy-Greenbrier sequence may have been induced by hydraulic fracturing itself, instead of deep disposal of the by-product wastewater. Hydraulic fracturing injects fluids at high pressure in order to increase natural gas production at wells that are oriented horizontally within the target rock formation. This process creates small fractures

in the formation, increasing its permeability and facilitating flow of the natural gas [*Davies et al.*, 2013; *Rubinstein and Mahani*, 2015]. In a process called stimulation, fluid injection is carried out in stages along different sections on the horizontal section of the production well, over a period of several days. The first stage is usually located near the toe (furthest point) of the well and subsequent stages move progressively back to the heel (where the well turns from horizontal to vertical). In each stage, which typically lasts several hours, a slurry containing a mixture of fluid and solid proppant is injected at a pressure high enough to fracture the rock, overcoming the minimum compressive stress. In hydraulic fracturing, the volume and duration of fluid injection are lower, but the pressure is much higher, compared to wastewater disposal; therefore, they have different potential risks for inducing earthquakes [*Walters et al.*, 2015]. Hydraulic fracturing is expected to generate microearthquakes with magnitude $-3 < M < 0$, since the intent is to create fractures restricted to the target formation [*Warpinski et al.*, 2012; *Maxwell*, 2013; *Rubinstein and Mahani*, 2015]; however, several studies have reported the occurrence of $M > 1$ earthquakes induced by hydraulic fracturing in Oklahoma [*Holland*, 2013], Ohio [*Friberg et al.*, 2014; *Skoumal et al.*, 2015a,b], United Kingdom [*Clarke et al.*, 2014], and western Canada, in northeast British Columbia and northwest Alberta [*BCOGC*, 2012, 2014; *Farahbod et al.*, 2015; *Schultz et al.*, 2015a,b, 2016; *Atkinson et al.*, 2016; *Bao and Eaton*, 2016; *Wang et al.*, 2016], including a M 4.6 event in British Columbia [*Atkinson et al.*, 2016]. These events are likely caused by reactivation of nearby critically stressed faults that are well-oriented to slip in the local stress field [*Maxwell*, 2013]. *Ogwari et al.* [2016] found a cluster of seismicity west of the Guy-Greenbrier Fault that was probably induced by hydraulic fracturing from 2010-09-29 to 2010-10-04. We search for spatial and temporal correlations between seismicity and the many production wells with hydraulic fracturing stimulation (Figure 1, small red-orange triangles with black lines) during 2010-06-01 to 2010-09-01.

2 Methods and Results

2.1 Data

The permanent seismic network in Arkansas is sparse, but includes a 3-component broadband seismic station WHAR (Figure 1, black triangle) recording 100 Hz data continuously since May 2010, located close to the Guy-Greenbrier Fault and in the area being prepared for production [*Horton*, 2012; *Ogwari et al.*, 2016]. ARK1 and ARK2, two temporary 3-component stations that started recording on 2010-06-11, were the only other available local seismic stations operating during 2010-06-01 to 2010-09-01; they are also known as CH1 and CH2 [*Ogwari et al.*, 2016; *Mousavi et al.*, 2017] or CHKGRS and CHKGUY [*Huang et al.*, 2016], respectively. We first detect earthquakes on

the single station WHAR, then use data from all 3 stations to confirm these detections and to locate and estimate magnitudes of the newly detected earthquakes.

2.2 Earthquake detection

To characterize fully the beginning stages of the Guy-Greenbrier earthquake sequence, we first detect as many earthquakes as possible. *Huang and Beroza* [2015] used single-station template matching on WHAR to detect up to 100 times more earthquakes than were recorded in the ANSS catalog between June 2010 and October 2011 in this earthquake sequence. Template matching, which cross-correlates known catalog template waveforms with continuous data to detect previously unknown low-magnitude events, exploits waveform similarity to improve detection sensitivity, and has often been used to resolve details of induced seismicity [e.g., *Holland*, 2013; *Friberg et al.*, 2014; *Skoumal et al.*, 2015a,b; *Schultz et al.*, 2015a,b, 2016].

The Fingerprint And Similarity Thresholding (FAST) earthquake detection method [*Yoon et al.*, 2015] adapts data-mining algorithms to perform a comprehensive search for similar earthquake waveforms within long duration continuous seismic data. It is especially useful in situations where template waveforms are not available or are not representative of all earthquake sources in an area. FAST assumes that every time window in continuous data is a potential template, and searches for time windows with similar waveforms in a computationally efficient way. FAST trades off speed for accuracy: instead of directly comparing waveforms, it computes fingerprints that replace waveforms with key discriminative features, and compares fingerprints for similarity in a probabilistic manner.

We use FAST with parameters in Table S1 to detect earthquakes in continuous data from station WHAR, bandpass filtered from 1-20 Hz, during the 3-month study period 2010-06-01 to 2010-09-01. First, we ran the single-channel detection algorithm in *Yoon et al.* [2015] independently on each component of data at WHAR. The runtime was about 5 days per component on a single processor. The output of FAST on a single component, which we can view as a sparse matrix (Figure 2), is a list of pairs of times within the continuous data with their associated FAST similarity score, where the fingerprints (and therefore waveforms) are similar. Earthquake signals should maintain similarity in time on all 3 components, so we expect the FAST similarity to add coherently at times when similar earthquakes occur. We sum the FAST similarity matrix from each component to get the total 3-component FAST similarity, on which we empirically set an event detection threshold of 0.33 by inspection (Table S1). After removing near-duplicate pairs and events within 4 s (Table S1) as described in *Yoon et al.* [2015], we find 28,675 events above this threshold.

FAST also detects non-earthquake signals with similar waveforms, so we need to remove these during post-processing. This is less of a concern for template matching, which only finds matches to a known earthquake waveform. Many of the similar non-earthquake signals are extremely narrowband (Figure S2), and we classify them as noise if they exceed the empirically determined threshold where at least 56% percent of the total signal power is within 1.5 Hz of the peak frequency on any one component. After removing 10,738 events classified as narrowband noise, we visually inspect the remaining 17,937 events and retain only the 13,026 events with a clear earthquake signal (containing P, S, and coda waves), preferably on at least two stations: WHAR and ARK2 or ARK1 (Figure 1, black triangles).

We compare the detection performance of FAST against that of template matching from *Huang and Beroza* [2015] during the 3-month study period. Templates, taken from ANSS catalog event waveforms at WHAR between May and October 2010, were 4 seconds long and bandpass filtered from 1-20 Hz. These templates were cross-correlated with continuous data at WHAR every 0.05 seconds. A different correlation coefficient (CC) threshold was used for each template, and for each hour of data. Event detection for template matching requires exceeding a CC threshold of 8 times the median absolute deviation. FAST detects a total of 13,026 events, while template matching found 13,946 events; most (12,368) events are detected by both methods (Figure 3, blue). In contrast, the ANSS catalog has only 75 events during this time (Data Set S1). Template matching detected 74 out of 75 catalog events; the remaining catalog event was not detected because it happened during a time gap in the continuous data at WHAR. FAST detected only 55 out of 75 catalog events, which suggests that the fingerprints may be less similar for the larger events, emphasizing the value of applying multiple detectors.

Figure 3 shows the local magnitude M_L (Section 2.3) as a function of time for all 14,604 events detected by either FAST, template matching, or both methods (Data Set S2). These events are microearthquakes, with the largest magnitude $M_L \leq 2.9$. FAST detects an additional 658 events that template matching did not find (Figure 3, cyan), which are lower in magnitude and clustered in time, demonstrating that a comprehensive search for similar earthquakes in continuous data finds unknown small events that would otherwise be overlooked. On the other hand, template matching found more (1,578) events that FAST fails to detect (Figure 3, magenta), which are lower in magnitude than most events but more evenly distributed in time. FAST is unable to detect every single event because it makes approximations in both representing waveforms and in searching for similar waveforms, but this tradeoff allows us to search thoroughly and efficiently a 3-month continuous seismic data set,

and still find 13,026/14,604 ≈ 89% of all detected events. Template matching successfully captures most small earthquakes (13,946/14,604 ≈ 95% of all detected events) in this data set.

2.3 Magnitude estimation

We estimate local magnitude M_L for all 14,604 detected events, which ranges from -1.5 to 2.9 (details in Section S1). In order to calibrate the M_L estimate, we first calculate the moment magnitude M_w for a selected group of 54 larger events with high-quality waveforms, located at different distances from station WHAR (Figure S3a). We obtain M_w by calculating seismic moment in the time domain from displacement waveforms at WHAR [*Prejean and Ellsworth*, 2001]. Next, we measure peak amplitudes on horizontal-component Wood-Anderson seismograms at all 3 stations for these 54 events, and invert for the distance correction parameters in the M_L estimate (Figure S3b). Finally, we apply this distance correction to peak Wood-Anderson amplitudes for all detected events to determine M_L [*Bormann*, 2012]. For the ANSS catalog events, the catalog magnitudes M_d computed from the coda duration are reasonably consistent with our M_L values (Figure S4a).

2.4 Initial earthquake location and refined velocity model

First, we determine absolute locations for 1,229 events with high-quality *P* and *S* phase arrivals on all 3 stations (Section S2). We estimate locations with VELEST [*Kissling et al.*, 1994] using the 1D velocity model from *Ogwari et al.* [2016], which was itself derived using VELEST as an improvement over the original 1D velocity model for this area [*Chiu et al.*, 1984].

Three of our events, all located near each other, have similar waveforms with high-amplitude surface waves characteristic of quarry blast sources [*Kafka*, 1990]. They occurred on 2010-06-24, 2010-07-02, and 2010-08-10 (Figure 4a). These events were detected by template matching but missed by FAST because the fingerprints of their waveforms at WHAR were not highly similar. Google Maps shows that the Greenbrier Quarry, owned by Rogers Group Inc., is located 1-2 km from our initial locations for these events. Inspection of Google Earth satellite imagery near the quarry location before (2009-07-23) and after (2010-09-15) the quarry blast times (Figure 4b) reveals a notch (red circle) in the southeast corner of the quarry in the post-blast image that was not in the pre-blast image. We therefore infer that all three blasts occurred on the surface (depth 0 km) at this notch location: 35.2928° N, 92.3973° W.

We use the notch location as ground truth for the 3 quarry blasts and solve for an updated 1D velocity model in VELEST, starting with the *Ogwari et al.* [2016] velocity model (Section S3). Table

1 lists the resulting improved 1D velocity model constrained by the quarry blast location; the V_p/V_s ratio deviates significantly from $\sqrt{3}$. Figure 5, which compares the new velocity model (solid lines) against the starting *Ogwari et al.* [2016] model (dashed lines), shows that the new model is slower at shallow depths where most events are located. We calculated this new model in order to refine velocity estimates in the shallowest layers within this small local area. We do not necessarily intend this model to replace the *Ogwari et al.* [2016] velocity model for the entire CNMSN.

We use the new velocity model (Table 1) to locate the events again in VELEST, starting with the same initial hypocenter location for all 1,229 events (1,226 earthquakes and 3 quarry blasts), equally weighting *P* and *S* travel times, and completing 50 iterations. Free locations of the 3 quarry blasts (Figure 1, nearby red circles) differ from the actual quarry location (Figure 1, red diamond) by as much as 2 km, which indicates a remaining absolute location error. Using the new quarry-constrained velocity model, the total root-mean-square (rms) residual for the 1,229 VELEST-located events is 0.0306, which is lower than the residual of 0.0347 for the *Ogwari et al.* [2016] velocity model.

We do not use the *Chiu et al.* [1984] velocity model for two reasons. First, the total rms residual from the resulting earthquake locations is higher than that from the *Ogwari et al.* [2016] model. Second, earthquake locations from the *Chiu et al.* [1984] model at the north end of the Guy-Greenbrier Fault in Box B1 (Figure 1) are inconsistent with the back-azimuth of these events calculated from P-wave polarization analysis [*Havskov and Ottemoller*, 2010] at station ARK2 (Figure S5).

2.5 High-precision earthquake location

The 1,229 events located by VELEST form several spatially compact clusters (Figure 1). To resolve the internal structure of each cluster, we use double-difference earthquake relocation [*Waldhauser and Ellsworth*, 2000], specifically hypoDD version 2.1b that allows as input the 1D quarry-constrained velocity model (Table 1) with variable V_p/V_s ratios in different layers.

We first compute differential travel times from both catalog *P* and *S* picks and cross-correlation for the 1,229 events where we already have initial absolute locations from VELEST (Section S4). We then compute cross-correlation differential times between each of the 1,229 initially located events and the 13,375 remaining unlocated events, which allows us to locate 511 additional events. Although these remaining events lack enough reliable *P* and *S* picks to locate with VELEST, their source locations are near already-located events such that the cross-correlation of time windows from the located and unlocated events will yield reliable relative locations in hypoDD.

We obtain precise relative earthquake locations within each cluster by running hypoDD in LSQR mode with parameters from Table S3 and weights from Table S4, using 904,354 *P* and 1,567,757 *S* cross-correlation differential travel times, as well as 72,368 *P* and 72,310 *S* catalog differential travel times. The blue bars in Figure 6 display the magnitude-frequency distribution of all 1,740 located events. 1,719 out of 1,740 events belong to one of 16 spatially compact clusters of earthquakes as defined by the latitude and longitude boundaries listed in Table S5. Figure 1 plots these event locations as circles sized by relative magnitude and colored by depth. Most of the events are located on or near the Guy-Greenbrier Fault. Profile A-A', a depth slice along the ~N30°E strike of the Guy-Greenbrier Fault [*Horton*, 2012; *Ogwari et al.*, 2016], shows only events located within 0.5 km of the fault; most events occur at the northeastern end in distinct clusters with relatively shallow depth (2-4 km), while there is a deeper (4-6 km depth) cluster of events to the southwest. In subsequent figures, Boxes B1 and B2 (red rectangles) explore 5 event clusters along or near the fault in greater detail. In addition, a significant number of events, many of them in compact clusters, are located at least 4 km away from the Guy-Greenbrier Fault. Profile B-B', a depth slice normal to the Guy-Greenbrier Fault, indicates not only the circled events along the near-vertical Guy-Greenbrier Fault, but also several event clusters located off the main Guy-Greenbrier Fault. In later figures, Box B3 (red rectangle) zooms in on 5 off-fault event clusters to the southeast, while Boxes B4 and B5 (red rectangles) closely examine 3 off-fault event clusters to the northwest. The map in Figure 1 also shows the location of 3 isolated earthquake clusters (C14, C15, C16 in blue boxes). To estimate the relative location error between pairs of closely spaced events, we run hypoDD in singular value decomposition (SVD) mode separately for 3 subsets of events: Cluster 3, Cluster 4, and Cluster 11. The relative location uncertainty is < 10 m for events within a cluster, which suggests that the structure within each cluster is real.

2.6 Improving temporal resolution of seismicity

Double-difference relocation significantly improves the spatial resolution of the 1,740 located earthquakes (Data Set S3). We are unable to locate the majority (12,864/14,604 $\approx$ 88%) of the detected events from Figure 3 because we lack quality *P* and *S* arrival picks at enough stations; however, we can improve the temporal resolution of the earthquake sequence by assigning unlocated events to Clusters 1-16 (Table S5) based on waveform similarity at station WHAR [*Cattaneo et al.*, 1999]. Also, stations ARK1 and ARK2 did not start operating until 2010-06-11, so we can only assign, instead of locate, events that occurred before this date. We represent each cluster with a stack waveform at WHAR, generated by averaging all located events belonging to that cluster. We then

cross-correlate each unlocated event with the stack waveform from every cluster, and assign it to the cluster with the highest CC. Section S5 has a detailed description of the assignment procedure.

Figure 7 verifies that the 2,525 unlocated events assigned to Cluster 1 have similar waveforms to each other and to the 667 located events in this cluster (shaded orange) at the 3 components of station WHAR. The CC between the pictured stack waveform (blue) and each of the 2,525 assigned events was at least 0.5. In all clusters, the high degree of waveform similarity gives us confidence that the assigned events originate from nearly the same source as the located events, and therefore can reliably improve the temporal resolution of the cluster.

For all clusters, the assigned events provide important information about the lower-magnitude events (Figure 6, black) and their timing within each cluster. The assigned events comprise (6,508/14,604 $\approx$ 44%) of the detected events, in addition to the (1,740/14,604 $\approx$ 12%) located events; however, 6,356 remaining detected events have waveforms that are too noisy to locate or assign (Figure 6, red), which are predominantly at the lowest magnitudes. We do not know if they are tiny events belonging to existing clusters, if they are events with different focal mechanisms in the same cluster, or if they are distinct or more distant earthquake sources that produce only small events.

2.7 Spatial and temporal correlation of seismicity with well data

Most of the 16 earthquake clusters are located near a production well stimulated by hydraulic fracturing (Figure 1, small red-orange triangles with black lines) or a wastewater injection well (Figure 1, inverted triangles), showing a spatial correlation. There are also many production wells without any nearby seismicity, although many of these wells are located more than 10 km from WHAR, so we would be less likely to detect seismicity near these wells, if it exists. We also check for a temporal correlation between seismicity in each cluster and the start date of wastewater injection at disposal wells, as well as the duration of stimulation stages at all production wells within a 2 km radius of the cluster, considering the absolute event location uncertainty.

2.7.1 Wastewater injection wells

Table S7 lists all wastewater disposal wells within the map area in Figure 1 (inverted triangles labeled by well number) active during the study period 2010-06-01 to 2010-09-01. Injection Wells 1 and 5 (colored by depth in Figure 1), which started injecting during the study period on 2010-07-07 and 2010-08-16 respectively, are the two injection wells located closest to the Guy-Greenbrier Fault.

The magnitude-time plots for located (blue) and assigned (black) events in each cluster (Figures 8, 13-15) show the start date of injection at Wells 1 and 5 as black dashed lines.

2.7.2 Stimulated production wells

Table S8 identifies all 53 production wells within the map area in Figure 1 (small triangles with black lines, colored by their true vertical depth) stimulated during the study period 2010-06-01 to 2010-09-01. The triangle indicates the surface location of the well, while the line shows the horizontal well path from heel to toe. We first queried the public Arkansas Oil and Gas Commission well database [*AOGC*, 2017a] for all production wells in the three counties spanning our map area (Faulkner, Cleburne, and Van Buren), then retained permit numbers for only the 53 wells inside the map boundaries in Figure 1 that were stimulated during the study period. We then searched the Arkansas Oil and Gas Commission Document Imaging Wells File Cabinet [*AOGC*, 2017b] by permit number for detailed production well data, including precise horizontal well trajectories and information about perforation and stages of hydraulic fracturing stimulation. The quality of stimulation data available varies widely depending on the company that collected and submitted the data. Some wells have detailed logs of the exact timing, injection rates, pressures, volumes, and chemical composition of each fluid injection within every stage of stimulation, while other wells have a short summary with only the start and end dates of stimulation. If timing information is available for stimulation stages, we convert the stimulation times from local Arkansas time (Central Daylight Time) to UTC time by adding 5 hours, for consistency with the seismic data. The magnitude-time plots for located and assigned events in each cluster (Figures 8, 13-15) show the duration of stimulation at all production wells within 2 km of the cluster (listed for each cluster in last column of Table S5) as a purple box, spanning the time from the start of the first stage to the end of the last stage.

2.7.3 Seismicity clusters near the Guy-Greenbrier Fault

Figure 8 focuses on seismicity located on or near the Guy-Greenbrier Fault, within Clusters 1-5 (blue boxes) in Boxes 1 and 2 from Figure 1, along with nearby production wells (small triangles with horizontal well path lines) and injection wells (inverted triangles). The magnitude-time plots for located and assigned events in each cluster explore temporal correlations between injection, stimulation, and the occurrence of seismicity.

Cluster 1, the northernmost cluster in Figure 8 located just northwest of the Guy-Greenbrier Fault, is the largest cluster with 3,192 total events (667 located and 2,525 assigned, Table S5).

Most events in Cluster 1 are shallow, with depth 2-3 km. These events locate on three east-west oriented structures, perpendicular to the north-south horizontal well path orientations of the 5 nearest production wells overlapping this cluster on the map. In addition, the magnitude-time plot for Cluster 1 shows an abrupt increase in both located and assigned seismicity that closely coincides with the timing and duration of stimulation (purple boxes) at the 7 nearest production wells in July 2010 (except for well 42069, which was stimulated in June 2010, and is temporally correlated with some $M_L < 1$ events in Cluster 1), with the seismicity lasting for several weeks after the end of stimulation before decaying with time. Cluster 1 is also located near injection Well 1 (about 3 km away), and most events occur after injection began at Well 1 with a time delay of about a week, but the obvious spatial and temporal correlations with the nearby stimulated production wells lead us to conclude that Cluster 1 seismicity was likely induced by hydraulic fracturing, rather than by wastewater injection.

Cluster 1 had the highest quality data, including a large number of earthquake locations, several stimulated production wells, and comprehensive stimulation data with start and stop times for all stages at each well. This led us to a more detailed investigation of spatial and temporal correlations between seismicity on different structures within Cluster 1 and each stage of stimulation at the 5 nearest production wells. Figure 9 examines the time evolution of Cluster 1 seismicity and stimulation stages at the 5 nearest production wells (permit numbers 42146, 42389, 42262, 43344, 43343), which are both colored by time with Day 0 defined as 2010-07-16 00:00:00 UTC. The event locations in Figure 9a are slightly offset from the horizontal well paths, which we attribute to our 2 km absolute location error resulting from the sparse 3-station network used for location. Figure 9b shows the seismicity shifted ~0.7 km southeast relative to the Figure 9a locations, which now completely overlie the 5 well paths, making it easier to view the detailed correlations where seismicity and stimulation stages on particular well sections have matching colors. This is motivated by our knowledge that relative location errors are much smaller and the geometry of locations is consistent with stimulation. In addition, the shifted locations in Figure 9b, which are within the 2 km absolute location uncertainty, agree with the back-azimuth derived from P-wave polarization analysis at station ARK2 (Figure S5).

Movie S1 displays the cumulative time evolution of Cluster 1 seismicity and stimulation stages at production wells 42146, 42389, 42262, 43344, 43343, both colored by the number of days since 2010-07-16 00:00:00 UTC (defined as Day 0); shifted event locations from Figure 9b are plotted. Figure 10a-d shows seismicity and stages during four different time intervals from Movie S1. The first stimulation stage started at the toe of the easternmost well 43343, and stages alternated between well 43343 and the adjacent well 43344 moving south during the first 3 days (red) in a zipperfrac pattern [*Vermylen and Zoback*, 2011], with seismicity closely following (Figure 10a). Then on day

3 (orange), while stimulation continued on wells 43343 and 43344, stimulation started at the toe of well 42389 to the west, and stages alternated between well 42389 and the adjacent westernmost well 42146 moving south toward the heel, again in a zipperfrac pattern, with seismicity also migrating in the same direction (Figure 10b, c). On day 10 (cyan), stimulation started at the toe of the center well 42262, again moving north to south, and the seismicity predictably follows the stages (Figure 10d). Seismicity persisted at the southeastern corner of Cluster 1 (Figure 10c,d) even after stimulations near the heel of wells 43343 and 43344 finished. Figure 10e displays a magnitude-time plot of the 16-day stimulation time period examined in Figure 9, Figure 10a-d, and Movie S1, with stimulation stages from each well plotted in a different color. The seismicity rate is higher during or immediately following the stimulation stages, which have a short duration of a few hours each, while seismicity tapers off during longer breaks between stimulation (during days 2-3, 9-10, and 14-16). Figure 9, Figure 10, and Movie S1 demonstrate a compelling spatial and temporal correlation of seismicity in Cluster 1 with individual stages of hydraulic fracturing stimulation.

Cluster 2 is located about 1 km north of the Guy-Greenbrier Fault, just south of Cluster 1 (Figure 8). It has a large number of events (1,078 total) located at a depth of 3-4 km. Most events in Cluster 2 are located on a 0.75 km-long, east-west oriented structure similar to those in Cluster 1, which is nearly orthogonal to the north-south well paths of the 8 nearby production wells. In the magnitude-time plot for Cluster 2, some earthquakes happen following stimulation at well 42069 in June 2010 (Table S8). In July 2010, a few events follow the start of injection at Well 1 (located just 1 km away), but the seismicity rate does not experience a large increase until the end of July, following stimulation at the remaining 7 nearby production wells. Compared to Cluster 1, there is a longer time delay between the onset of stimulation and the rapid increase in seismicity; most events in Cluster 2 actually occur after stimulation has ended. Such time delays, longer than a week, have been observed in other cases of hydraulic fracturing induced seismicity [*Schultz et al.*, 2015a, 2016]. The location, orientation, and timing of Cluster 2 seismicity suggest that these events were probably induced by hydraulic fracturing, rather than by wastewater injection. However, we cannot completely exclude the possibility that Cluster 2 was induced by wastewater injection, due to its depth, timing, and proximity to Well 1.

Most seismicity in Cluster 3, the closest earthquake cluster to injection Well 1, is oriented along the strike of the Guy-Greenbrier Fault (Figure 8), although there is a small east-west oriented sub-cluster of events at the northern boundary of Cluster 3 (Cluster 3C from Table S5). Cluster 4, located farther southwest away from the production wells, contains fewer events, also located on the Guy-Greenbrier Fault. The magnitude-time plots show that seismic activity in Clusters 3 and 4

significantly increases following injection at Well 1, after a short 3-day time delay, but is not affected much by stimulation later in July. These events have depth 3-4 km and have lower magnitude (mostly $M_L < 1$) than events in Clusters 1 and 2. The abrupt increase in seismicity starting on 2010-08-29 was reported in *Ogwari et al.* [2016] as the beginning of the Guy-Greenbrier sequence, but we see a lower level of microseismicity initiate and persist within a few days of injection. We conclude that Cluster 3 and 4 events have a stronger spatial and temporal correlation with, and thus are more likely to be induced by, wastewater injection at Well 1, rather than stimulation.

The presence of distinct east-west trending structures formed by Clusters 1, 2, and 3C motivated us to explore the source mechanism of these events. We select 300 events from Cluster 1, 159 events from Cluster 2, and 22 events from Cluster 3C (Figure 11a) and plot their first motions (black "u": up, red "d": down) on a composite focal mechanism projected onto the lower hemisphere (Figure 11b). Since we have sparse station coverage, we assume that all 3 clusters have the same mechanism. Cluster 1 events are shifted ~0.7 km southeast as in Figure 9b for the first motion calculation. If we assume a double-couple source mechanism, we can manually fit two nodal planes to the first motion data, one trending ~N75°E and the other oriented ~N15°W (Figure 11b, black lines). If the ~N75°E nodal plane is the fault plane, which is a reasonable assumption given the east-west orientation of seismicity, the first motions indicate right-lateral strike-slip motion along this fault. However, given the regional ~N60°E maximum horizontal stress orientation [*Hurd and Zoback*, 2012], we would expect left-lateral strike-slip motion along east-west oriented faults. Local heterogeneity in the stress orientation is unlikely because the regional ~N60°E stress orientation is consistent with right-lateral strike-slip motion on the nearby favorably oriented ~N30°E Guy-Greenbrier Fault [*Horton*, 2012]. This contradiction between the expected left-lateral slip from the stress orientation, and the observed right-lateral motion on the focal mechanism, rules out the possibility that these events in Clusters 1, 2, and 3C are left-lateral strike-slip earthquakes activated on preexisting east-west faults favorably oriented in the regional stress field [*Maxwell*, 2013]. Instead, we relax the double-couple assumption, and suggest that these events have a combination of shear and tensile faulting. Although the sparse data are inconclusive, the restricted region of dilatational first motions near the center of the focal sphere (Figure 11b, red "d") could be explained by a non-double-couple mechanism with a volumetric component resulting from opening of small east-west-oriented fractures [*Sileny et al.*, 2009; *Fischer and Guest*, 2011; *Vavrycuk*, 2011], which is an intended goal of hydraulic fracturing to facilitate flow of hydrocarbons. The east-west seismicity is oriented perpendicular to the well paths, which supports this idea, although there are several events in Cluster 1 with $M_L > 2$ (Figure 8), which is higher than the expected $-3 < M < 0$ magnitude range of microseismicity from opening hydraulic

fractures [*Warpinski et al.*, 2012; *Rubinstein and Mahani*, 2015]. We note that our interpretation is limited by the lack of first motion data at enough stations, and it is possible that these 3 clusters actually have different mechanisms, contrary to our assumption.

Cluster 5 is located farther to the southwest on the Guy-Greenbrier Fault (Figure 8), in Box 2 (Figure 1). These events align with the strike of the Guy-Greenbrier Fault and have depths around 5 km. They are deeper than events in Clusters 1-4 farther northeast along the fault and were reported by *Ogwari et al.* [2016] as the first four events on the southern section of the fault. Cluster 5 was definitely not induced by injection at nearby Well 5, because most events occurred before the start of injection. There are two stimulated production wells near Cluster 5, and most of the events occur after stimulation at well 43114, so it is possible that Cluster 5 was induced by stimulation; however, the along-strike orientation, deeper depth, and lower seismicity rate (compared to Clusters 1 and 2, which were likely induced by hydraulic fracturing) suggest an alternative explanation that we favor: Cluster 5 could have been triggered by diffusion of pore pressure from injection at Well 1, with a longer time delay between the start of injection and the first event in August 2010.

Figure 12 summarizes all seismicity in Clusters 1-5 with epicenters restricted to within 0.5 km of the Guy-Greenbrier Fault (Profile A-A', Figure 1). Figure 12a shows the depth of these events as a function of along-strike distance (also shown in Figure 1), along with the location and depth (magenta sections) of wastewater injection Wells 1 and 5. It also displays the depths of the target Fayetteville Shale formation, the sedimentary Paleozoic Boone Formation/Ozark Aquifer into where injection occurs, and the crystalline Precambrian basement below [*Ogwari et al.*, 2016]. Since seismicity along the Guy-Greenbrier Fault is located within the triangular area outlined by the three seismic stations (Figure 1), these event depths should be reliable. Figure 12b shows the timing of events, as well as the onset of injection at Wells 1 and 5, along the strike of the Guy-Greenbrier Fault; it is obvious that Cluster 5 events occur before injection started at Well 5. We separate out the events in Clusters 1 and 2 (blue boxes labeled C1, C2) because they were likely induced by hydraulic fracturing stimulation. *Ogwari et al.* [2016] and *Mousavi et al.* [2017] report high b-values in these areas later on in September and October 2010, which is also consistent with hydraulic fracturing induced seismicity. The remaining events along the fault, belonging to Clusters 3, 4, 5, were probably induced by wastewater injection at Well 1. We estimate an apparent hydraulic diffusivity of $D \approx 1$ m^2/s, assuming a homogeneous and isotropic medium [*Shapiro et al.*, 2002; *Shapiro and Dinske*, 2009]:

$$r = \sqrt{4\pi D t}. \quad (1)$$

Pore pressure diffuses outward from injection at Well 1, and reaches injection Well 5, located r = 5.5 km away, t = 28 days after injection started at Well 1. In comparison, *Ogwari and Horton* [2016] used a detailed numerical model and observed seismicity to estimate hydraulic diffusivity along the Guy-Greenbrier Fault during October and November 2010, when seismicity dramatically increased. They found $D \approx 0.2\text{-}0.3$ m^2/s in the northern and central sections of the fault (near Clusters 3 and 4), while in the southern section (near Cluster 5), their diffusivity was $D \approx 1.1$ m^2/s above 5 km depth and $D \approx 0.02$ m^2/s below 5 km. *Mousavi et al.* [2017] estimated a lower hydraulic diffusivity of $D \approx 0.01$ m^2/s in the northern section of the Guy-Greenbrier Fault, from fitting Equation 1 to a detailed catalog of seismicity from *Ogwari et al.* [2016] during the time period from 2010-07-07 to 2010-10-20, following injection at Well 1. The results in Figure 12 foreshadow the migration of seismicity from northeast to southwest and from the shallower Paleozoic sedimentary formation into deeper Precambrian basement seen soon afterward in September-October 2010 [*Ogwari et al.*, 2016; *Ogwari and Horton*, 2016].

2.7.4 Seismicity clusters off the Guy-Greenbrier Fault

Figure 13 takes a closer look at seismicity located 5-10 km southeast of the Guy-Greenbrier Fault, within Clusters 6-10 (blue boxes) in Box 3 from Figure 1, and nearby production wells that were stimulated during 2010-06-01 to 2010-09-01. Magnitude-time plots for located and assigned events in each cluster examine temporal correlations between injection at Wells 1 and 5, stimulation, and the occurrence of seismicity. Since these events are located outside the 3-station network, their depth estimates are unreliable. We ended up with greater (4-8 km) depth estimates than those from ANSS catalog events in this area (Figure S1), although the depths mostly agree within the large uncertainties in the catalog depths (Figure S4b). Cluster 6 contains events in an east-west orientation, and there is an obvious temporal correlation between seismicity and the duration of stimulation at the nearest production well 43043. The first detected and associated (not located) event starts about 3 hours after the onset of the first stage of stimulation, and the seismicity rate remains high until the end of stimulation, after which the seismicity rate decays rapidly. The stimulation and seismicity began a day before the start of injection at Well 1, which is too distant (6 km away) to have an immediate effect on the seismicity in Cluster 6. We therefore conclude that Cluster 6 was likely induced by hydraulic fracturing. Similarly, Cluster 7 was also likely induced by hydraulic fracturing given the strong spatial and temporal correlation between events in this cluster and stimulation at the nearest production well 43153. For this cluster the seismicity rate remains high after stimulation has ceased. The depth and orientation of events in Cluster 7 are not as accurate, given the greater distance away

from the seismic stations. Cluster 8 is spatially and temporally correlated with hydraulic fracturing stimulation at nearby well 43258 in June 2010. The seismicity rate is initially high, then decreases after the end of stimulation, but remains at a low level for over a month; however, there is no seismicity in August 2010, following stimulation at another nearby well 43154. Cluster 8 may be the same events, detected and located using a temporary seismic network by *Horton* [2012], that were reported to lie on an east-west elongated trend near the Morrilton Fault east of injection well 3.

Cluster 9, which contains only 12 events, is located far from all production and injection wells, and the magnitude-time plot does not show any temporal correlation with injection at Well 1 or 5, or with stimulation at the nearest well 43154 located 2 km away. Therefore, we interpret these events as natural background seismicity. The Enola swarms of 1982, with 30,000 earthquakes in 3 years [*Chiu et al.*, 1984], and 2001, which had 2,500 earthquakes in 2 months with a M 4.4 as the largest event [*Rabak et al.*, 2010], were natural earthquake sequences that occurred 2 km south of Box B3, near 35.18°N, 92.2° W. The Enola swarms happened long before the start of hydraulic fracturing in the Fayetteville Shale [*Horton*, 2012] or the start of wastewater injection in 2009 (Table S7).

The events in Cluster 10 could possibly be a result of hydraulic fracturing, but the quality of our results is not sufficient to be definitive in this case. Event waveforms are noisy, leading to higher location uncertainties. There are many production wells in the large area defining Cluster 10, which are stimulated during over half the 3-month time period (Table S8): the magnitude-time plot for Cluster 10 is mostly purple, so although the most of the high-seismicity time periods overlap with stimulation, this temporal correlation is not informative.

Figure 14 closely examines seismicity located 4-8 kilometers northwest of the Guy-Greenbrier Fault, within Clusters 11-13 (blue boxes) in Boxes 4 and 5 (Figure 1), and nearby stimulated production wells. Magnitude-time plots once again highlight temporal correlations between well activity and events in these clusters. Clusters 11 and 12 are just outside our seismic network, so their depth estimates are probably reliable. In contrast, Cluster 13 is much farther away from the network, so depths for these events are unreliable, and the event waveforms at WHAR are noisy. The timing of events in Cluster 11, located directly on and oriented almost orthogonal to the well path of the nearest production well 43439, overlaps closely with the duration of stimulation at well 43439, with a rapid decay of seismicity after stimulation ends, so Cluster 11 was likely induced by hydraulic fracturing. Cluster 12, with only 14 events, is located about 1 km away from production well 43433, but the temporal correlation with stimulation at this well is weak, since there is a long delay between the end of stimulation and the seismicity. It is possible that these events, with a depth

of about 4 km, were triggered by pore pressure diffusion from injection at the nearest disposal Well 1 with a long time delay; however, we have no clear evidence to suggest that Cluster 12 events are not natural background seismicity. Cluster 13 was possibly induced by stimulation at nearby production wells 43254, 43255, and 43256 in late May and early June of 2010, after which the seismicity rate is very high. After the end of stimulation in early June, seismicity in Cluster 13 lingers at a low level during the entire 3-month study period, without being affected much by stimulation at wells 43252 and 43253 in July 2010. Many of the events in Cluster 13 were in the ANSS catalog (Figure S1).

Figure 15 shows magnitude-time plots for seismicity in Clusters 14-16, in various locations several kilometers away from the Guy-Greenbrier Fault (blue boxes labeled C14, C15, C16 in Figure 1). As these events are far from the seismic network, we detected very few events in these clusters, and their locations and depths are uncertain. Cluster 14 is most likely natural seismicity, since it is located far from any injection or production wells, and the events occur over the entire 3-month study period at a low background rate. Clusters 15 and 16 are spatially and temporally correlated with the duration of stimulation at the nearest production wells, 43244 and 43219 respectively, so they were likely induced by hydraulic fracturing.

3 Discussion

3.1 Microearthquakes induced by hydraulic fracturing are common during the Guy-Greenbrier sequence

Our analysis reveals that the initial stages of the Guy-Greenbrier earthquake sequence contain a complicated mixture of microseismicity with $M_L \leq 2.9$. The vast majority of these earthquakes are spatially and temporally related to hydraulic fracturing stimulation operations, which suggests that $M_L > 1$ seismicity induced by hydraulic fracturing is more common than widely appreciated. We identify about 3/4 (56/75) of ANSS catalog events as induced by hydraulic fracturing. Depending on the well, there is significant variation in the duration of seismicity after the end of stimulation. About 1/3 (17/53) of the production wells in this area stimulated during 2010-06-01 to 2010-09-01 are associated with seismicity (Table 2). We identify a smaller number of events, located on the Guy-Greenbrier Fault, which were likely induced by wastewater injection at Well 1 starting on 2010-07-07, and their migration southwest to greater depths anticipates the behavior of intense seismic activity to come in September and October 2010 [*Ogwari et al.*, 2016]. A small fraction of events that were uncorrelated with hydraulic fracturing or wastewater disposal may be natural background seismicity, which is known to occur in this area. Table 2 summarizes our preferred

interpretation for whether each cluster of earthquakes was natural, induced by injection, or induced by hydraulic fracturing.

We suggest that much of the microseismicity later in the entire Guy-Greenbrier earthquake sequence from September 2010 to October 2011, originally attributed to deep wastewater injection [*Horton*, 2012; *Huang and Beroza*, 2015], may instead be a result of hydraulic fracturing stimulation. Since the Guy-Greenbrier area has a history of natural seismicity from the Enola swarms in 1982 [*Chiu et al.*, 1984] and 2001 [*Rabak et al.*, 2010], one might consider the background seismicity during June 2010, before the start of injection at Well 1, to be natural tectonic seismicity; however, our study found that most of these "background" events were spatially and temporally associated with hydraulic fracturing operations. We also found that many events during July and August 2010, after the start of injection at Well 1, resulted from hydraulic fracturing rather than wastewater disposal, with some of these events located very close to or on the Guy-Greenbrier Fault. *Ogwari et al.* [2016] and *Mousavi et al.* [2017] reported a cluster of events west of the Guy-Greenbrier Fault that was likely induced by hydraulic fracturing stimulation from 2010-09-29 to 2010-10-04; such events probably also exist later in the sequence.

The combination of sensitive detection and precise location of microseismicity, and a detailed public database of disposal wells and production wells with stimulation information, has allowed us to separate events induced by hydraulic fracturing from events induced by wastewater injection (Figure 12). The presence of these multiple influences on seismicity poses significant challenges for seismic hazard mitigation, where different actions would be required for injection-induced seismicity versus hydraulic fracturing-induced seismicity [*Walters et al.*, 2015]. All events during our study period had $M_L \leq 2.9$, so they were too small to cause damage, but they do change the stresses locally. *Ogwari and Horton* [2016] found that pore-pressure changes less than 0.06 MPa can initiate seismicity on the critically stressed Guy-Greenbrier fault. We speculate that the presence of events, such as those in Cluster 1, which are larger than expected for events caused by opening new fractures [*Warpinski et al.*, 2012], could be a useful indicator that care should be taken with plans for nearby large-scale wastewater injection.

3.2 Benefits of high-sensitivity, high-resolution seismic monitoring

We demonstrate that it is possible to extract detailed information on the location and timing of microseismicity with a sparse 3-station seismic network recording continuously, with stations spaced 5-10 km apart, using high-resolution seismological techniques for event detection and location.

Methods that use waveform similarity to detect earthquakes in continuous seismic data, such as template matching, the Repeating Signal Detector [*Skoumal et al.*, 2016], and FAST [*Yoon et al.*, 2015], can significantly improve the magnitude of completeness, allowing a statistical analysis of seismicity rate changes over time and their relationship to fluid injection [*Huang and Beroza*, 2015], and reveal unknown sources of low-magnitude seismicity, such as the clusters we found to be induced by hydraulic fracturing. Many other studies have used template matching to identify seismicity induced by hydraulic fracturing [*Holland*, 2013; *Friberg et al.*, 2014; *Skoumal et al.*, 2015a,b; *Schultz et al.*, 2015a,b, 2016], as the magnitudes are often lower than for injection-induced seismicity [*Rutqvist et al.*, 2013]. In addition to improved detection, we can obtain high-precision event locations using double-difference relocation with cross-correlation derived travel times for similar pairs of events [*Waldhauser and Ellsworth*, 2000], and subsequently obtain additional temporal resolution with waveform cross-correlation. These types of detection, location, and correlation methods are well suited for induced seismicity, where many events occur in close proximity as clusters, and thus have similar waveforms when recorded at the same station. Limitations of our study include location uncertainties due to the minimal 3-station network. We have poor depth constraints for events outside the seismic network, and our absolute location uncertainty was 2 km even after improving the velocity model with the quarry data. Nevertheless, the fortunate combination of a 3-station, 3-component continuous seismic network located near the seismicity and a high-quality public well database with detailed records of injection and hydraulic fracturing stimulation allowed us to discern the relationship between microseismicity, wastewater disposal, and hydraulic fracturing in this area.

In regions where seismic networks are sparse, our study suggests a cost-effective strategy for seismic monitoring. A large number of stations is always helpful, but waveform-based detection and location methods are essential for making the most out of a limited data set. It is preferable for seismic stations to start recording continuously before the beginning of injection or stimulation operations so that background seismicity can be measured. For example, we can envision running single-station FAST at each station in a widely spaced permanent network to identify the existing background rate of low-magnitude events. If the seismicity rate or the maximum magnitude of an earthquake exceeds an acceptable threshold, or otherwise seems anomalous, a temporary network with additional stations could be deployed to enable more detailed characterization of the earthquakes, shifting limited resources where they are needed the most. Early awareness of changes in seismicity can inform timely and informed decision making for operators and regulators about whether to continue or alter injection and hydraulic fracturing activities, possibly as implemented in traffic light systems for seismic risk management [*Walters et al.*, 2015].

4 Conclusions

In Guy-Greenbrier, Arkansas, an area of unconventional natural gas production in the Fayetteville Shale, wastewater injection beginning in July 2010 was widely suspected to have induced a year-long earthquake sequence that culminated in a M_w 4.7 earthquake [*Horton*, 2012; *Huang and Beroza*, 2015; *Ogwari et al.*, 2016]. We characterized seismicity at a very fine scale during the 3-month time period 2010-06-01 to 2010-09-01, which includes background seismicity, initiation of the Guy-Greenbrier earthquake sequence, and the early seismic response to wastewater injection, with $M_L \leq 2.9$ for all events. We used sensitive event detection methods: FAST [*Yoon et al.*, 2015], and template matching [*Huang and Beroza*, 2015], to detect 14,604 similar-waveform low-magnitude earthquakes in continuous seismic data at a single station. We followed this with precise relative double-difference location of nearby earthquakes [*Waldhauser and Ellsworth*, 2000] at three stations with an improved quarry-blast constrained velocity model, then harnessed waveform similarity to refine the temporal resolution of located event clusters. Most events during these 3 months were spatially and temporally correlated with hydraulic fracturing stimulation operations at a small number of nearby production wells, while we attribute a smaller number of events, located on and starting to outline the yet-to-be-discovered Guy-Greenbrier Fault, to wastewater injection at Well 1 starting in July 2010. Many stimulated production wells have no nearby detected seismicity. Although this area has hosted swarms of natural seismicity in the past [*Chiu et al.*, 1984; *Rabak et al.*, 2010], we infer that only a small fraction of events during these 3 months are natural in origin. The simultaneous presence of seismicity induced by both hydraulic fracturing and wastewater injection, which we speculate is also true later in the earthquake sequence, presents a challenge for seismic hazard mitigation and operational decision-making with traffic light systems [*Walters et al.*, 2015]. We demonstrate that given continuous seismic data and a detailed public well database with injection and stimulation information, it is possible to obtain high-resolution seismological observations even with a sparse 3-station network. We advocate continuous seismic monitoring for anomalous earthquake activity before starting injection or hydraulic fracturing.

Acknowledgments

The facilities of IRIS Data Services, and specifically the IRIS Data Management Center, were used for access to continuous seismic data used in this study (seismic networks AG and 7F). IRIS Data Services are funded through the Seismological Facilities for the Advancement of Geoscience and EarthScope (SAGE) Proposal of the National Science Foundation under Cooperative Agreement EAR-1261681. We obtained injection and production well data from the Arkansas Oil and Gas Commission Database

[*AOGC*, 2017a,b]. ANSS Comprehensive Earthquake Catalog (ComCat) data was downloaded from the U.S. Geological Survey website: https://earthquake.usgs.gov/data/comcat/. We ran FAST on high-performance computing clusters provided by the Stanford Center for Computational Earth and Environmental Science (CEES). We used Seismic Analysis Code (SAC) [*Helffrich et al.*, 2013] to manually pick P and S arrivals, ObsPy (with NumPy and Matplotlib) for seismological data processing and visualization [*Beyreuther et al.*, 2010], and Generic Mapping Tools (GMT) to generate maps [*Wessel et al.*, 2013]. Paul Ogwari showed us how to access stimulation data for Arkansas production wells [*AOGC*, 2017b] and shared valuable insights on the Guy-Greenbrier sequence. Mostafa Mousavi provided a detailed, thorough review of the initial manuscript. We thank Paul Friberg and Paul Ogwari for helpful reviews of the submitted manuscript. This paper also benefited from discussions with Mark Zoback, Rob Skoumal, Rongmao Zhou and Diane Frazier from BHP Billiton, and industrial affiliate members in the Stanford Center for Induced and Triggered Seismicity (SCITS). We thank Karianne Bergen for assistance with hierarchical clustering and Martin Schoenball for hypoDD help. C.Y. was funded by a Chevron Stanford Graduate Fellowship and by National Science Foundation grant EAR-1551462. The authors acknowledge financial support from SCITS.

References

Arkansas Oil and Gas Commission (AOGC) (2017a). *Arkansas Well Data Search*, http://www.aogc.state.ar.us/welldata/wells/default.aspx (last accessed March 2017).

Arkansas Oil and Gas Commission (AOGC) (2017b). *Arkansas DocuWare Document Imaging Wells File Cabinet*, http://aogc2.state.ar.us:8080/DocuWare/PlatformRO/WebClient?orgId=1 (last accessed March 2017).

Atkinson, G. M., Eaton, D. W., Ghofrani, H., Walker, D., Cheadle, B., Schultz, R., Shcherbakov, R., Tiampo, K., Gu, J., Harrington, R. M., Liu, Y., van der Baan, M., and Kao, H. (2016). Hydraulic Fracturing and Seismicity in the Western Canada Sedimentary Basin. *Seismol. Res. Lett.* 87, 3, doi:10.1785/0220150263.

Bao, X. and Eaton, D. W. (2016). Fault activation by hydraulic fracturing in western Canada. *Science*, doi:10.1126/science.aag2583.

Beyreuther, M., Barsch, R., Kischer, L., Megies, T., Behr, Y., and Wassermann, J. (2010). ObsPy: A Python toolbox for seismology. *Seismol. Res. Lett.* 81, 3, 530–533, doi:10.1785/gssrl.81.3.530.

Bisrat, S., DeShon, H. R., and Rowe, C. (2012). Microseismic Swarm Activity in the New Madrid Seismic Zone. *Bull. Seismol. Soc. Am.* 102, 3, 1167–1178, doi:10.1785/0120100315.

Bormann, P. (Editor) (2012). Seismic Sources and Source Parameters, in *New Manual of Seismological Observatory Practice (NMSOP-2)* (Chapter 3), IASPEI, GFZ German Research Centre for Geosciences, Potsdam, ed. 2, http://nmsop.gfz-potsdam.de (last accessed July 2017), doi:10.2312/GFZ.NMSOP-2.

British Columbia Oil and Gas Commission (BCOGC) (2012). *Investigation of Observed Seismicity in the Horn River Basin*, https://www.bcogc.ca/node/8046/download (last accessed June 2017).

British Columbia Oil and Gas Commission (BCOGC) (2014). *Investigation of Observed Seismicity in the Montney Trend*, https://www.bcogc.ca/node/12291/download (last accessed June 2017).

Cattaneo, M., Augliera, P., Spallarossa, D., and Lanza, V. (1999). A Waveform Similarity Approach to Investigate Seismicity Patterns. *Natural Hazards* 19, 123–138.

Chiu, J. M., Johnston, A. C., Metzger, A. G., Haar, L., and Fletcher, J. (1984). Analysis of analog and digital records of the 1982 Arkansas earthquake swarm. *Bull. Seismol. Soc. Am.* 74, 5, 1721–1742.

Clarke, H., Eisner, L., Styles, P., and Turner, P. (2014). Felt seismicity associated with shale gas hydraulic fracturing: The first documented example in Europe. *Geophys. Res. Lett.* 41, doi:10.1002/2014GL062047.

Davies, R., Foulger, G., Bindley, A., and Styles, P. (2013). Induced seismicity and hydraulic fracturing for the recovery of hydrocarbons. *Mar. Pet. Geol.* 45, 171–185, doi:10.1016/j.marpetgeo.2013.03.016.

Deichmann, N., and Garcia-Fernandez, M. (1992). Rupture geometry from high-precision relative hypocentre locations of microearthquake clusters. *Geophys. J. Int.* 110, 501–517.

Ellsworth, W. L. (2013). Injection-induced earthquakes. *Science* 341, doi:10.1126/science.1225942.

Farahbod, A. M., Kao, H., Cassidy, J. F., and Walker, D. (2015). How did hydraulic-fracturing operations in the Horn River Basin change seismicity patterns in northeastern British Columbia, Canada? *The Leading Edge* 34, 658–663, doi:10.1190/tle34060658.1.

Fischer, T., and Guest, A. (2011). Shear and tensile earthquakes caused by fluid injection. *Geophys. Res. Lett.* 38, L05307, doi:10.1029/2010GL045447.

Friberg, P. A., Besana-Ostman, G. M., and Dricker, I. (2014). Characterization of an earthquake sequence triggered by hydraulic fracturing in Harrison County, Ohio. *Seismol. Res. Lett.* 85, 6, 1–13, doi:10.1785/0220140127.

Green, D. N., and Neuberg, J. (2006). Waveform classification of volcanic low-frequency earthquake swarms and its implication at Soufriere Hills Volcano, Montserrat. *J. Volcanol. Geotherm. Res.* 153, 51–63, doi:10.1016/j.jvolgeores.2005.08.003.

Harris, D. (2006). Subspace detectors: Theory, *Lawrence Livermore Natl. Lab. Rep. UCRL-TR-222758*, Lawrence Livermore National Laboratory, Livermore, California.

Harris, D. B. and Dodge, D. A. (2011). An Autonomous System for Grouping Events in a Developing Aftershock Sequence. *Bull. Seismol. Soc. Am.* 101, 2, 763–774, doi:10.1785/0120100103.

Havskov, J., and Ottemoller, L. (2010). Location, in *Routine Data Processing in Earthquake Seismology* (Chapter 5), Springer, New York, ed. 1, 101–150, https://link.springer.com/book/10.1007%2F978-90-481-8697-6.

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

Helffrich, G., Wookey, J., and Bastow, I. (2013). *The Seismic Analysis Code: A Primer and User's Guide*, Cambridge University Press, United Kingdom, ed. 1.

Holland, A. A. (2013). Earthquakes Triggered by Hydraulic Fracturing in South-Central Oklahoma. *Bull. Seismol. Soc. Am.* 103, 3, 1784–1792, doi:10.1785/0120120109.

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, 250–260, doi:10.1785/gssrl.83.2.250.

Huang, Y., Beroza, G. C., and Ellsworth, W. L. (2016). Stress drop estimates of potentially induced earthquakes in the Guy-Greenbrier sequence. *J. Geophys. Res. Solid Earth* 121, 1–11, doi:10.1002/2015GL065170.

Huang, Y., and Beroza, G. C. (2015). Temporal variation in the magnitude-frequency distribution during the Guy-Greenbrier earthquake sequence. *Geophys. Res. Lett.* 42, 6639–6646, doi:10.1002/2016JB013067.

Hurd, O., and Zoback, M. D. (2012). Intraplate earthquakes, regional stress and fault mechanics in the Central and Eastern U.S. and Southeastern Canada. *Tectonophysics* 581, 182–192, doi:10.1016/j.tecto.2012.04.002.

Kafka, A. L. (1990). *Rg* as a Depth Discriminant for Earthquakes and Explosions: A Case Study in New England. *Bull. Seismol. Soc. Am.* 80, 2, 373–394.

Kissling, E., Ellsworth, W. L., Eberhart-Phillips, D., and Kradolfer, U. (1994). Initial reference models in local earthquake tomography. *J. Geophys. Res.* 99, 19,635–19,646.

Leskovec, J., Rajaraman, A., and Ullman, J .D. (2014). Clustering, in *Mining of Massive Datasets* (Chapter 7), Cambridge University Press, New York, ed. 2, 73–130, http://www.mmds.org.

Maeda, N. (1985). A method for reading and checking phase times in auto- processing system of seismic wave data. *Zisin=Jishin* 38, 365–379.

Massa, M., Eva, E., Spallarossa, D., and Eva, C. (2006). Detection of earthquake clusters on the basis of waveform similarity: An application in the monferrato region (Piedmont, Italy). *J. Seismol.* 10, 1–22, doi:10.1007/s10950-006-2840-4.

Maxwell, S. C., Jones, M. B., Parker, R. L., Leaney, W. S., Mack, M., Dorval, D., D'Amico, D., Logel, J., Anderson, E., and Hammermaster, K. (2010). Fault activation during hydraulic fracturing. *AAPG Search Discov* 90172, doi:10.1190/1.3255145.

Maxwell, S. (2013). Unintentional Seismicity Induced by Hydraulic Fracturing. *CSEG Rec.* 38, 40–49.

Mousavi, S. M., Ogwari, P. O., Horton, S. P., and Langston, C .A. (2017). Spatio-temporal evolution of frequency-magnitude distribution and seismogenic index during initiation of induced seismicity at Guy-Greenbrier, Arkansas. *Phys. Earth Planet. Inter.* 267, 53–66, doi:10.1016/j.pepi.2017.04.005.

Ogwari, P. O., Horton, S. P., and Ausbrooks, S. (2016). Characteristics of Induced/ Triggered Earthquakes during the Startup Phase of the Guy-Greenbrier Earthquake Sequence in North-Central Arkansas. *Seismol. Res. Lett.* 87, 3, doi:10.1785/0220150252.

Ogwari, P. O., and Horton, S. P. (2016). Numerical model of pore-pressure diffusion associated with the initiation of the 2010-2011 Guy-Greenbrier, Arkansas earthquakes. *Geofluids* 16, 954–970, doi:10.1111/gfl.12198.

Petersen, T. (2007). Swarms of repeating long-period earthquakes at Shishaldin Volcano, Alaska, 2001-2004. *J. Volcanol. Geotherm. Res.* 166, 177–192, doi:10.1016/j.jvolgeores.2007.07.014.

Prejean, S. G., and Ellsworth, W. L. (2001). Observations of Earthquake Source Parameters at 2 km Depth in the Long Valley Caldera, Eastern California. *Bull. Seismol. Soc. Am.* 91, 2, 165–177.

Rabak, I., Langston, C., Bodin, P., Horton, S., Withers, M., and Powell, C. (2010). The Enola, Arkansas, Intraplate Swarm of 2001. *Seismol. Res. Lett.* 81, 3, 549–559, doi:10.1785/gssrl.81.3.549.

Raleigh, C. B., Healy, J. H., and Bredehoeft, J. D. (1976). An experiment in earthquake control at Rangely, Colorado. *Science* 191, 4233, 1230–1237, doi:10.1126/science.191.4233.1230.

Rowe, C. A., Thurber, C. H., and White, R. A. (2004). Dome growth behavior at Soufriere Hills Volcano, Montserrat, revealed by relocation of volcanic event swarms, 1995-1996. *J. Volcanol. Geotherm. Res.* 134, 199–221, doi:10.1016/j.jvolgeores.2004.01.008.

Rubinstein, J. L. and Mahani, A. B. (2015). Myths and Facts on Wastewater Injection, Hydraulic Fracturing, Enhanced Oil Recovery, and Induced Seismicity. *Seismol. Res. Lett.* 86, doi:10.1785/0220150067.

Rutqvist, J., Rinaldi, A. P., Cappa, F., and Moridis, G. J. (2013). Modeling of fault reactivation and induced seismicity during hydraulic fracturing of shale-gas reservoirs. *J. Pet. Sci. Eng.* 107, 31–44, doi:10.1016/j.petrol.2013.04.023.

Schaff, D. P., Bokelmann, G. H. R., Ellsworth, W. L., Zanzerkia, E., Waldhauser, F., and Beroza, G. C. (2004). Optimizing Correlation Techniques for Improved Earthquake Location. *Bull. Seismol. Soc. Am.* 94, 2, 705–721.

Schultz, R., Stern, V., Novakovic, M., Atkinson, G., and Gu, Y. J. (2015a). Hydraulic fracturing and the Crooked Lake Sequences: Insights gleaned from regional seismic networks. *Geophys. Res. Lett.* 42, 2750–2758, doi:10.1002/ 2015GL063455.

Schultz, R., Mei, S., Pana, D., Gu, Y. J., Kim, A., and Eaton, D. (2015b). The Cardston Earthquake Swarm and Hydraulic Fracturing of the Exshaw Formation (Alberta Bakken Play). *Bull. Seismol. Soc. Am.* 105, 6, doi:10.1785/0120150131.

Schultz, R., Wang, R., Gu, Y. J., Haug, K., and Atkinson, G. (2016). A seismological overview of the induced earthquakes in the Duvernay play near Fox Creek, Alberta. *J. Geophys. Res. Solid Earth* 122, doi:10.1002/2016JB013570.

Shapiro, S. A., Rothert, E., Rath, V., and Rindschwentner, J. (2002). Characterization of fluid transport properties of reservoirs using induced microseismicity. *Geophysics* 67, 1, 212–220, doi:10.1190/1.1451597.

Shapiro, S. A., and Dinske, C. (2009). Fluid-induced seismicity: Pressure diffusion and hydraulic fracturing. *Geophysical Prospecting* 57, 301–310, doi:10.1111/j.1365-2478.2008.00770.x.

Sileny, J., Hill, D. P., Eisner, L., and Cornet, F. H. (2009). Non–double-couple mechanisms of microearthquakes induced by hydraulic fracturing. *J. Geophys. Res.* 114, B08307, doi:10.1029/2008JB005987.

Skoumal, R. J., Brudzinski, M. R., and Currie, B. S. (2015a). Earthquakes Induced by Hydraulic Fracturing in Poland Township, Ohio. *Bull. Seismol. Soc. Am.* 105, 1, 189–197, doi:10.1785/0120140168.

Skoumal, R. J., Brudzinski, M. R., and Currie, B. S. (2015b). Distinguishing induced seismicity from natural seismicity in Ohio: Demonstrating the utility of waveform template matching. *J. Geophys. Res. Solid Earth* 120, 6284–6296, doi:10.1002/2015JB012265.

Skoumal, R. J., Brudzinski, M. R., and Currie, B. S. (2016). An efficient repeating signal detector to investigate earthquake swarms. *J. Geophys. Res. Solid Earth* 121, 5880–5897, doi:10.1002/2016JB012981.

Thelen, W. A., Allstadt, K., De Angelis, S., Malone, S. D., Moran, S. C., and Vidale, J. (2013). Shallow repeating seismic events under an alpine glacier at Mount Rainier, Washington, USA. *J. Glaciol.* 59, 214, 345–356, doi:10.3189/2013JoG12J111.

Vavrycuk, V. (2011). Tensile earthquakes: Theory, modeling, and inversion. *J. Geophys. Res.* 116, B12320, doi:10.1029/2011JB008770.

Vermylen, J. P., and Zoback, M. D. (2011). Hydraulic fracturing, microseismic magnitudes, and stress evolution in the Barnett Shale, Texas, USA. Paper SPE-140507-MS presented at SPE Hydraulic Fracturing Conference. Society of Petroleum Engineers, The Woodlands, Tex.

Waldhauser, F., and Ellsworth, W. L. (2000). A Double-Difference Earthquake Location Algorithm: Method and Application to the Northern Hayward Fault, California. *Bull. Seismol. Soc. Am.* 90, 6, 1353–1368.

Walters, R. J., Zoback, M. D., Baker, J. W., and Beroza, G. C. (2015). Characterizing and Responding to Seismic Risk Associated with Earthquakes Potentially Triggered by Fluid Disposal and Hydraulic Fracturing. *Seismol. Res. Lett.* 86, 4, doi:10.1785/0220150048.

Wang, R., Gu, Y. J., Schultz, R., Kim, A., and Atkinson, G. (2016). Source analysis of a potential hydraulic-fracturing-induced earthquake near Fox Creek, Alberta. *Geophys. Res. Lett.* 43, 564–573, doi:10.1002/2015GL066917.

Warpinski, N. R, Du, J., and Zimmer, U. (2012). Measurements of Hydraulic-Fracture-Induced Seismicity in Gas Shales. *PE Hydraul. Fract. Technol. Conf., SPE 151597,* The Woodlands, Texas, 6-8 February 2012, doi:10.2118/151597-PA.

Wessel, P., Smith, W. H. F., Scharroo, R., Luis, J. F., and Wobbe, F. (2013). Generic Mapping Tools: Improved version released. *Eos Trans. AGU* 94, 409–410, doi:10.1002/2013EO450001.

Yoon, C. E., O'Reilly, O., Bergen, K. J., and Beroza, G. C. (2015). Earthquake detection through computationally efficient similarity search. *Science Advances* 1, e1501057, doi:10.1126/sciadv.1501057.

Table 1. New 1D velocity model, constrained by quarry blast location, used to locate all earthquakes in this study.

Depth (km)	P wave (km/s)	S wave (km/s)	V_p/V_s
0.0	4.06	2.46	1.650
1.22	5.57	3.22	1.730
2.89	6.12	3.27	1.872
6.23	6.23	3.58	1.740
13.0	6.24	3.71	1.682

Table 2. Summary of 16 seismicity clusters from 2010-06-01 to 2010-09-01, names of wells associated with each cluster, and our preferred interpretation of whether they are natural, induced by hydraulic fracturing stimulation, or induced by wastewater injection at Well 1.

Cluster Number	Interpretation	Associated Well Names
1	Hydraulic fracturing	42146, 42389, 42262, 43344, 43343
2	Hydraulic fracturing	42069, 43375, 43376
3	Wastewater injection	Injection Well 1
4	Wastewater injection	Injection Well 1
5	Wastewater injection	Injection Well 1
6	Hydraulic fracturing	43043
7	Hydraulic fracturing	43153
8	Hydraulic fracturing	43258
9	Natural	–
10	Hydraulic fracturing	Several wells, but not definitive
11	Hydraulic fracturing	43439
12	Natural	–
13	Hydraulic fracturing	43254, 43255, 43256
14	Natural	–
15	Hydraulic fracturing	43244
16	Hydraulic fracturing	43219

Figure 1. Map of Guy-Greenbrier area in central Arkansas (red box, inset at lower left) with earthquake locations, seismic stations, wastewater injection wells, and production wells with hydraulic fracturing stimulation during the time period 2010-06-01 to 2010-09-01. Profile A-A' shows only seismicity located within 0.5 km of the Guy-Greenbrier Fault. Profile B-B', perpendicular to the Guy-Greenbrier Fault, shows all seismicity: circled events are located on the Guy-Greenbrier Fault, while other seismicity is off the Guy-Greenbrier Fault. The profiles include locations of ANSS catalog events after this time period (small gray dots) to delineate the location of the Guy-Greenbrier Fault. Later figures zoom in on areas enclosed in red boxes B1-B5 (B1 and B2 in Figure 8, B3 in Figure 13, B4 and B5 in Figure 14). Blue boxes C14-C16 indicate isolated clusters of seismicity discussed in Section 2.7. Fault traces are from *Horton* [2012].

Figure 2. Schematic illustration of combining FAST similarity matrix output from multiple components at a single station WHAR as sparse matrix addition. Each square represents a pair of fingerprints (which can be mapped back to waveforms) from two different times in the continuous data. Gray squares with high similarity indicate times when similar waveforms occur for each component. Black squares indicate times when waveforms are similar on all 3 components.

Figure 3. Comparison of earthquakes detected by FAST and template matching during the time period 2010-06-01 to 2010-09-01. Both methods detect the same 12,368 events (blue). FAST detects an additional 658 events that template matching did not find (cyan), while template matching detected 1,578 events that FAST fails to detect (magenta). We detect a total of 14,604 events using either FAST, template matching, or both methods.

Figure 4. (a) Three quarry blasts with similar waveforms recorded between 2010-06-01 and 2010-09-01. (b) Google Earth satellite imagery from before (2009-07-23) and after (2010-09-15) the 3 quarry blasts. We infer that the blasting occurred at the circled notch (red), which is present in the post-blast image but absent from the pre-blast image. This notch location is used as ground truth for the 3 quarry blasts, allowing us to refine the velocity model.

Figure 5. 1D velocity model comparison. We use the updated Central Arkansas 1D velocity model from Table 3 in *Ogwari et al.* [2016] (dashed) as a starting model, then calculate a refined velocity model constrained by the quarry blast location (solid). The refined model has slightly lower P and S wave velocities at shallow depth and is used to locate all earthquakes in this study.

Figure 6. Magnitude-frequency distribution of all 14,604 detected events (Figure 3): 1,740 located events (blue), 6,508 assigned events (black), 6,356 unassigned events (red). Although we were unable to locate the assigned events, we can categorize them as belonging to Clusters 1-16 through cross-correlation of event waveforms at station WHAR (Section 2.6). The predominantly low-magnitude unassigned events are the remaining detected events from Figure 3 that are too noisy to either locate or associate with existing clusters. The largest 3 events were not located because they occurred before 2010-06-11, when data was available only at station WHAR.

Figure 7. Representative stack waveform (top) and normalized waveforms aligned with cross-correlation (bottom) of all 3,192 earthquakes belonging to Cluster 1, recorded on each component of station WHAR (east: left; north: center; vertical: right). We located the 667 largest events (shaded orange) in Cluster 1. Although we were unable to locate the 2,525 lower-magnitude events due to a lack of high-quality picks at stations ARK1 and ARK2, their waveforms at station WHAR are similar to the located event waveforms, so they can be used to improve the temporal resolution of events in Cluster 1.

878 **Figure 8.** Zoomed map view of seismicity in Clusters 1-5 (blue boxes), Boxes 1-2 (red boxes in Figure 1 -
879 see legend), and their spatial and temporal relationship to nearby stimulated production wells (small triangles
880 colored by depth, labeled by permit number from Table S8, listed for each cluster in last column of Table S5)
881 and wastewater injection wells (inverted triangles colored by depth). Earthquakes on the map are circles colored
882 by depth and sized by relative magnitude. These events are located on or near the labeled Guy-Greenbrier Fault,
883 with ~N30°E strike [*Horton*, 2012; *Ogwari et al.*, 2016]. Thick black arrows indicate the ~N60°E orientation
884 of maximum horizontal compressive stress in this region [*Hurd and Zoback*, 2012].

885 **Figure 9.** Time evolution of seismicity in Cluster 1, and hydraulic fracturing stimulation at the 5 nearest
886 production wells (labeled by permit number from Table S8), near north end of the Guy-Greenbrier Fault.
887 Earthquakes (circles sized by relative magnitude), as well as stimulated sections of the production wells during
888 each stage of hydraulic fracturing, are colored by time with Day 0 defined as 2010-07-16 00:00:00 UTC. (a)
889 Actual seismicity locations, which exhibit an offset from the well paths. (b) Seismicity locations shifted ~0.7
890 km southeast relative to the locations from (a), which makes it easier to see the spatial and temporal correlations
891 between seismicity and stimulation stages. We display these shifted locations in Figure 10 and Movie S1. The
892 shifted locations, which are within the 2 km absolute location uncertainty, agree with the back-azimuth derived
893 from P-wave polarization analysis at station ARK2 (Figure S5).

894 **Figure 10.** Time evolution of seismicity in Cluster 1, and hydraulic fracturing stimulation at the 5 nearest
895 production wells (labeled by permit number from Table S8), near north end of the Guy-Greenbrier Fault.
896 We display the shifted seismicity locations from Figure 9b. In (a)-(d), earthquakes (circles sized by relative
897 magnitude), as well as stimulated sections of the production wells during each stage of hydraulic fracturing, are
898 colored by time with Day 0 defined as 2010-07-16 00:00:00 UTC. This figure shows seismicity and stimulated
899 stages during different time intervals after the start of stimulation: (a) 0 days to 3 days 8 hours, with early
900 stimulations at wells 43343 and 43344; (b) 3 days 8 hours to 6 days 8 hours, with later stimulations at wells
901 43343 and 43344, and early stimulations at wells 42146 and 42389; (c) 6 days 8 hours to 10 days 6 hours,
902 with later stimulations at wells 42146 and 42389; (d) 10 days 6 hours to 16 days, with stimulations at well
903 42262. Movie S1 displays cumulative Cluster 1 seismicity and stages of stimulation for the entire 16-day time
904 period. (e) Time evolution of magnitudes for located (blue) and assigned (black) events during the 16 days
905 of stimulation, with labeled time intervals for (a)-(d). We plot the stimulation duration of all stages from a
906 particular production well in a different color.

Figure 11. (a) Map view of selected east-west oriented events (colored by depth, sized by relative magnitude) in Clusters 1, 2, 3C with first motions represented on composite focal mechanism. Cluster 1 events are shifted ~0.7 km southeast as in Figure 9b. Nearby stimulated production wells (small triangles colored by depth, labeled by permit number from Table S8) and wastewater injection wells (inverted triangle colored by depth) are shown. (b) Composite focal mechanism from first motion polarity (black "u": up, red "d": down) of selected events in Clusters 1, 2, 3C, on lower hemisphere projection. Black lines show nodal planes that best fit the first motion polarity data, assuming a double-couple source. Thick black arrows indicate the ~N60°E orientation of maximum horizontal compressive stress in this region [*Hurd and Zoback*, 2012].

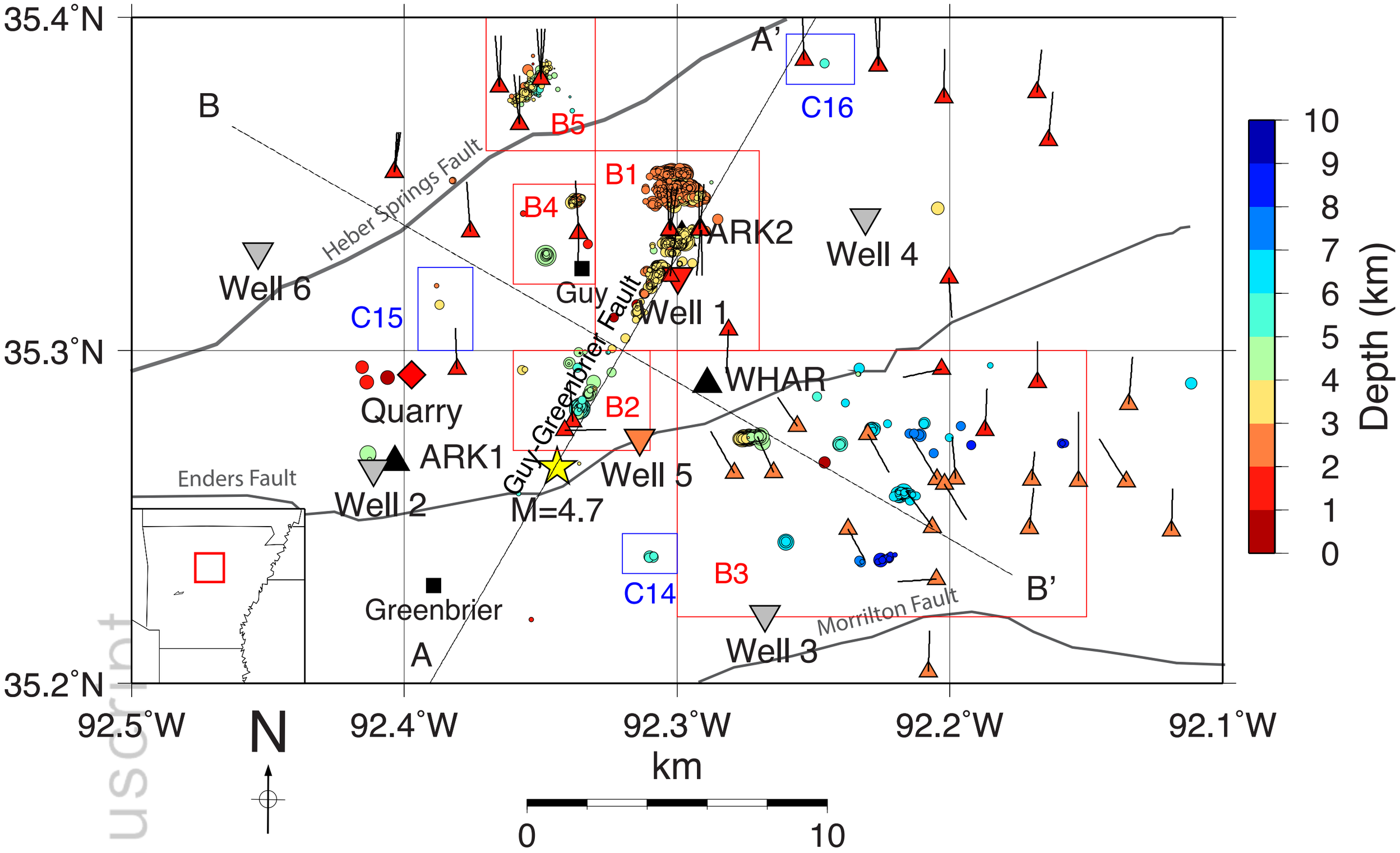
Figure 12. Summary of seismicity located within 0.5 km of the Guy-Greenbrier Fault, with wastewater injection Wells 1 and 5 (Table S7), along the southwest to northeast cross-section Profile A-A' (Figure 1). Earthquakes are circles colored by depth and sized by relative magnitude. Clusters 1 and 2 (blue boxes labeled as C1, C2) were likely induced by hydraulic fracturing stimulation. The remaining events belonging to Clusters 3, 4, 5 were probably induced by wastewater injection at Well 1. (a) Depth of events as a function of along-strike distance. Later ANSS catalog event locations (small gray dots), from 2010-09-01 to 2011-10-31, delineate the depth extent of the Guy-Greenbrier Fault. Magenta section shows the depth of wastewater injection at Wells 1 and 5. Depths for the Fayetteville Shale, Boone Formation/Ozark Aquifer, and Precambrian basement were obtained from *Ogwari et al.* [2016]. (b) Time of events and wastewater injection (arrows) as a function of along-strike distance. We estimate a hydraulic diffusivity of D $\approx$ 1 m^2/s for pore pressure diffusion from injection at Well 1.

Figure 13. Zoomed map view of seismicity in Clusters 6-10 (blue boxes), Box 3 (red box in Figure 1 - see legend), and their spatial and temporal relationship to nearby stimulated production wells (small triangles colored by depth, labeled by permit number from Table S8, listed for each cluster in last column of Table S5). Earthquakes on the map are circles colored by depth and sized by relative magnitude. These events are located off the main Guy-Greenbrier Fault, to the southeast.

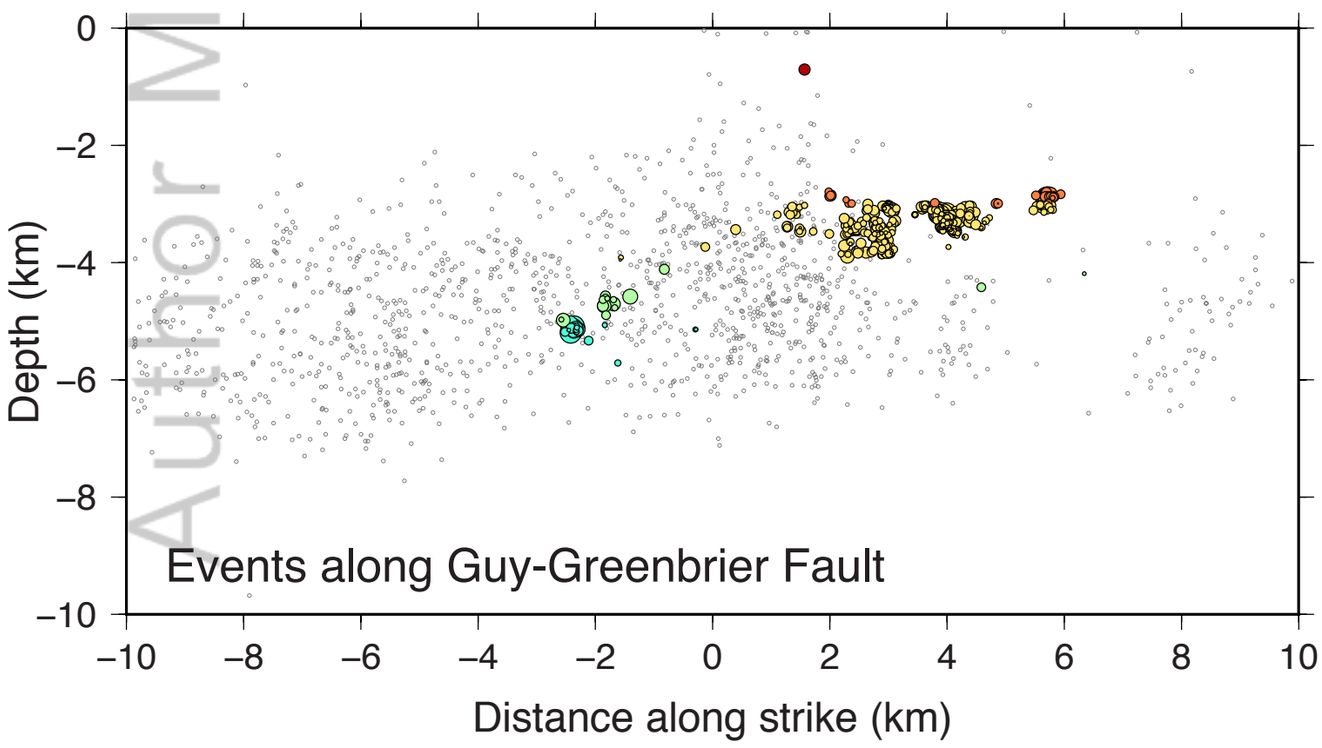
Figure 14. Zoomed map view of seismicity in Clusters 11-13 (blue boxes), Boxes 4-5 (red boxes in Figure 1 - see legend), and their spatial and temporal relationship to nearby stimulated production wells (small triangles colored by depth, labeled by permit number from Table S8, listed for each cluster in last column of Table S5). Earthquakes on the map are circles colored by depth and sized by relative magnitude. These events are located off the main Guy-Greenbrier Fault, to the northwest.

Figure 15. Seismicity in Clusters 14-16 (locations in blue boxes, Figure 1), and their temporal relationship to nearby stimulated production wells (listed for each cluster in last column of Table S5) and wastewater injection wells. These events are in isolated clusters located off the main Guy-Greenbrier Fault.

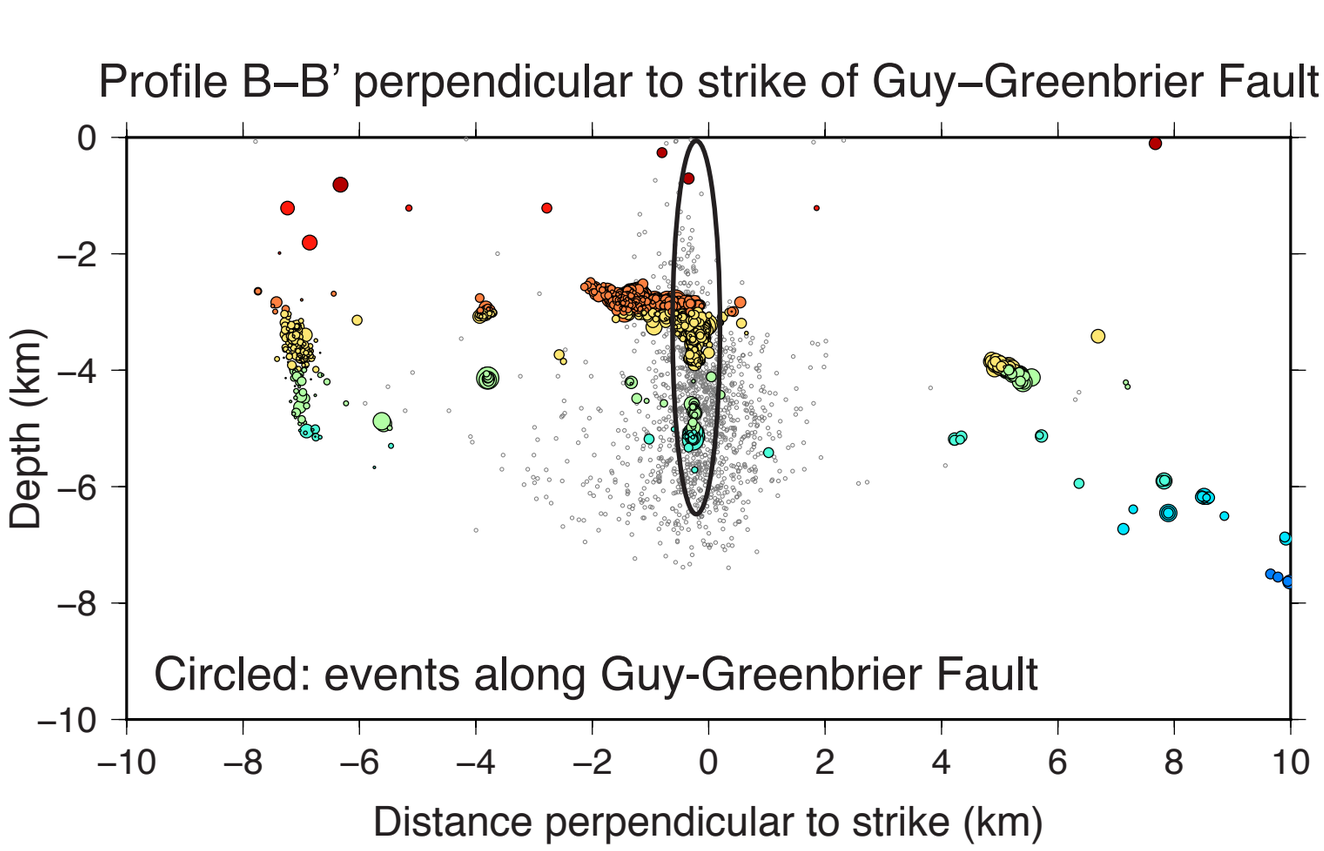
Figure 1.



Profile A–A’ along strike of Guy–Greenbrier Fault



Profile B–B’ perpendicular to strike of Guy–Greenbrier Fault



Legend

- New detected and located earthquakes, 2010-06-01 to 2010-09-01
  Sized by magnitude (-1.5 < M < 2.9), colored by depth
- Catalog events (ANSS), only on profiles: 2010-09-01 to 2011-10-31
  Indicates location of Guy-Greenbrier Fault
- Seismic stations
- Cities
- Quarry
- Largest earthquake in sequence, M 4.7, 2011-02-27
- Wastewater injection wells (colored by depth: started injecting 2010-06-01 to 2010-09-01)
- Production wells with hydraulic fracture stimulation, 2010-06-01 to 2010-09-01 (triangle: surface location and heel, colored by depth; line: horizontal extent to toe)
- Earthquake clusters (C1-C16)
  Clusters C1-C13 are within B1-B5, see later figures
- Zoom-in areas (B1-B5)

Figure 2.

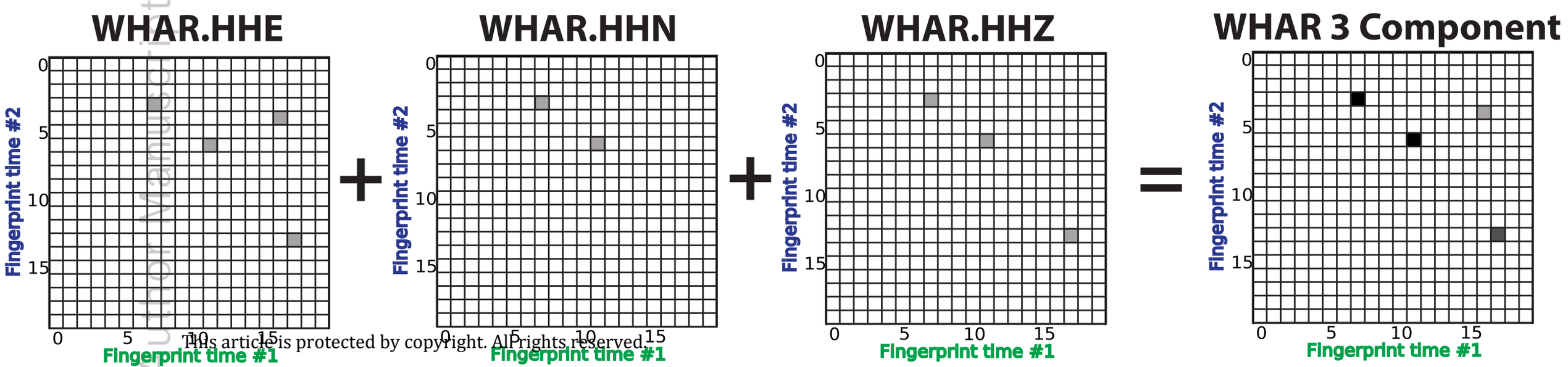
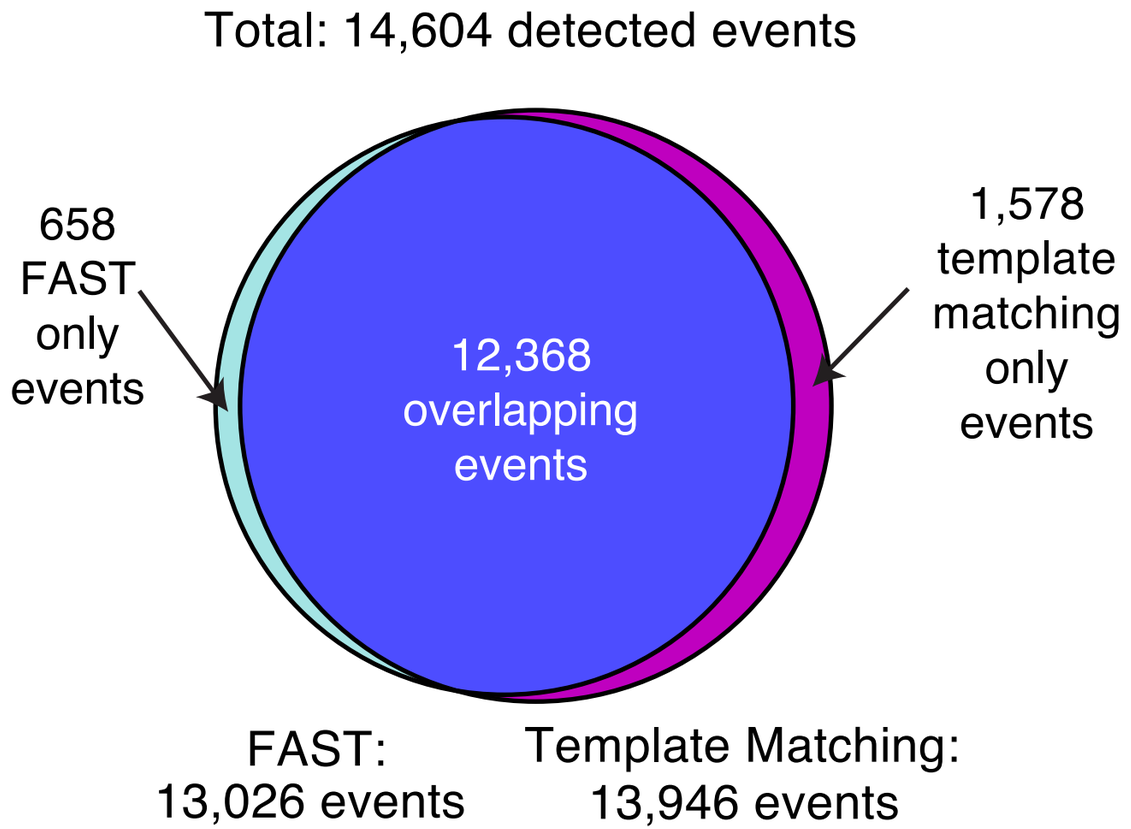


Figure 3.



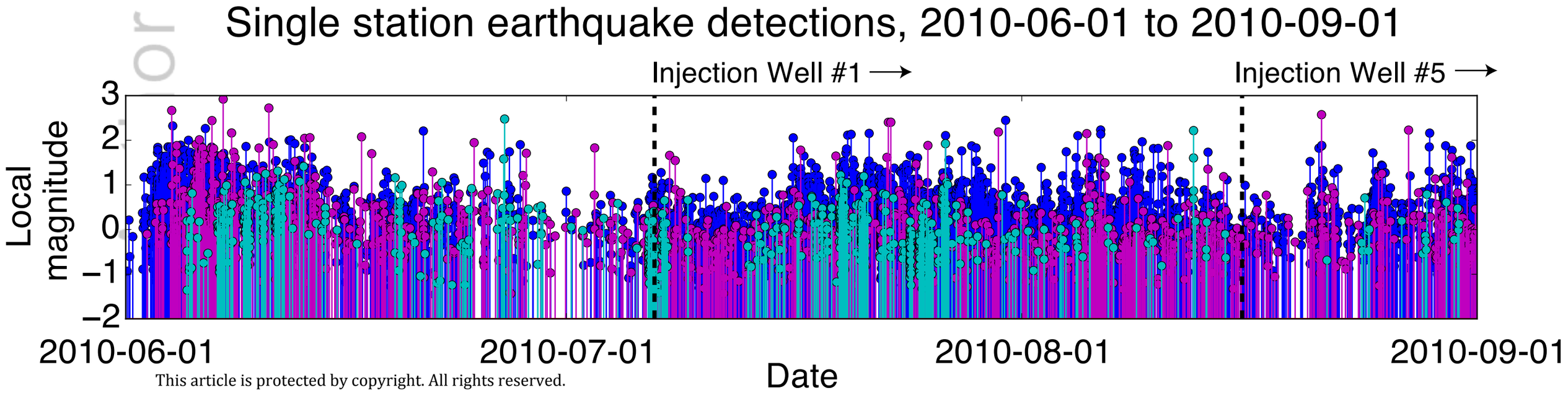
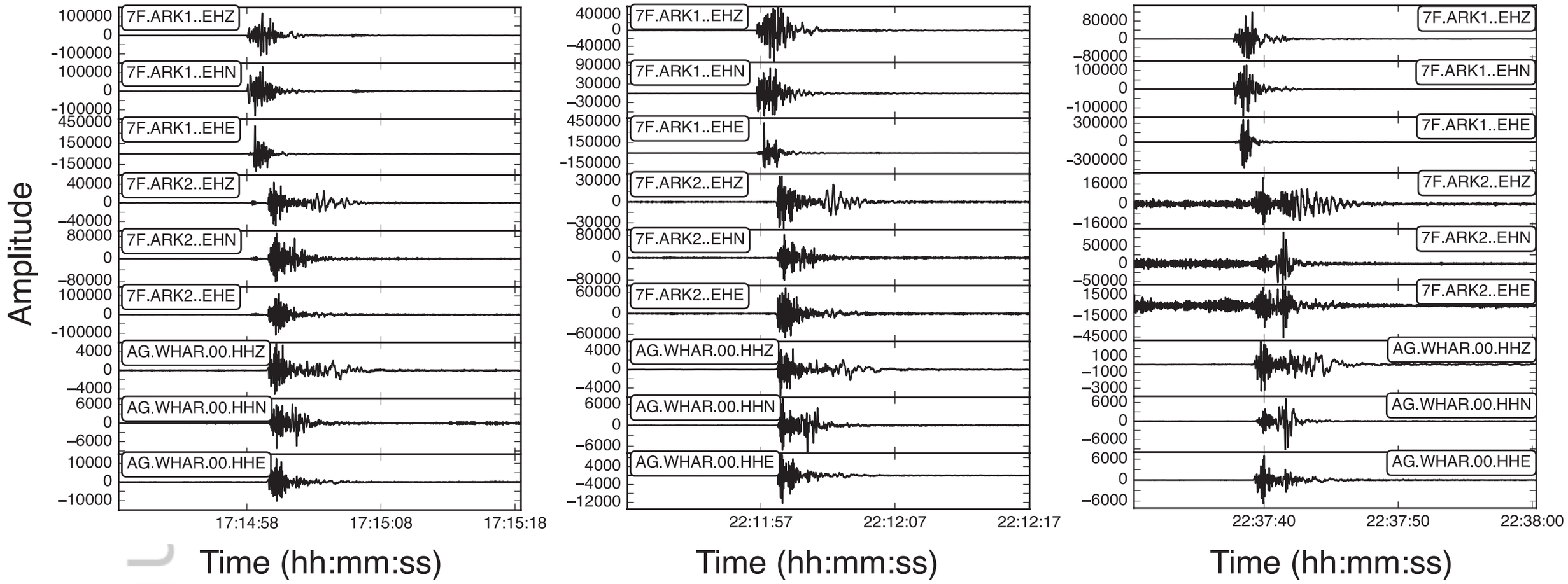


Figure 4.

a)



b)

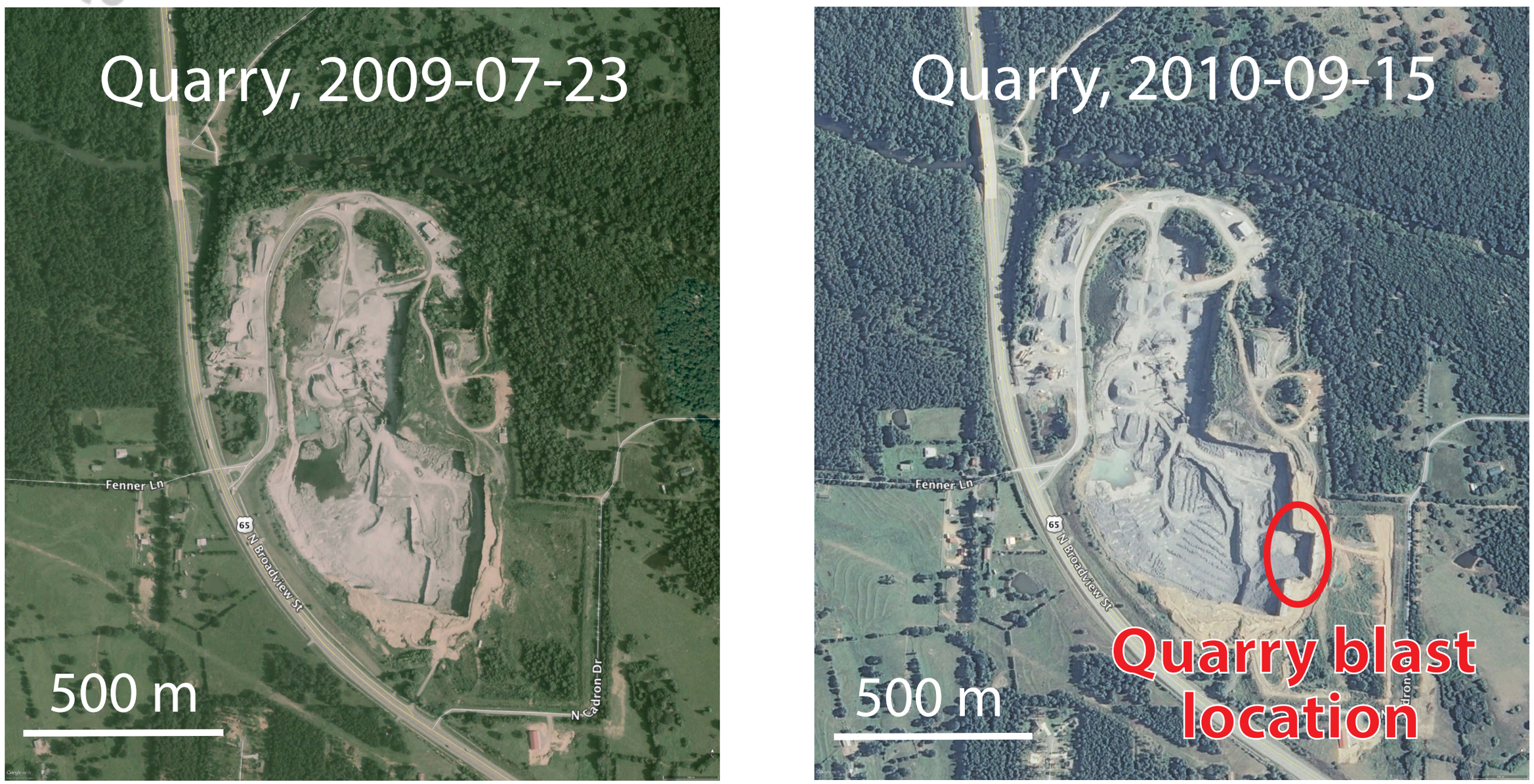


Figure 5.

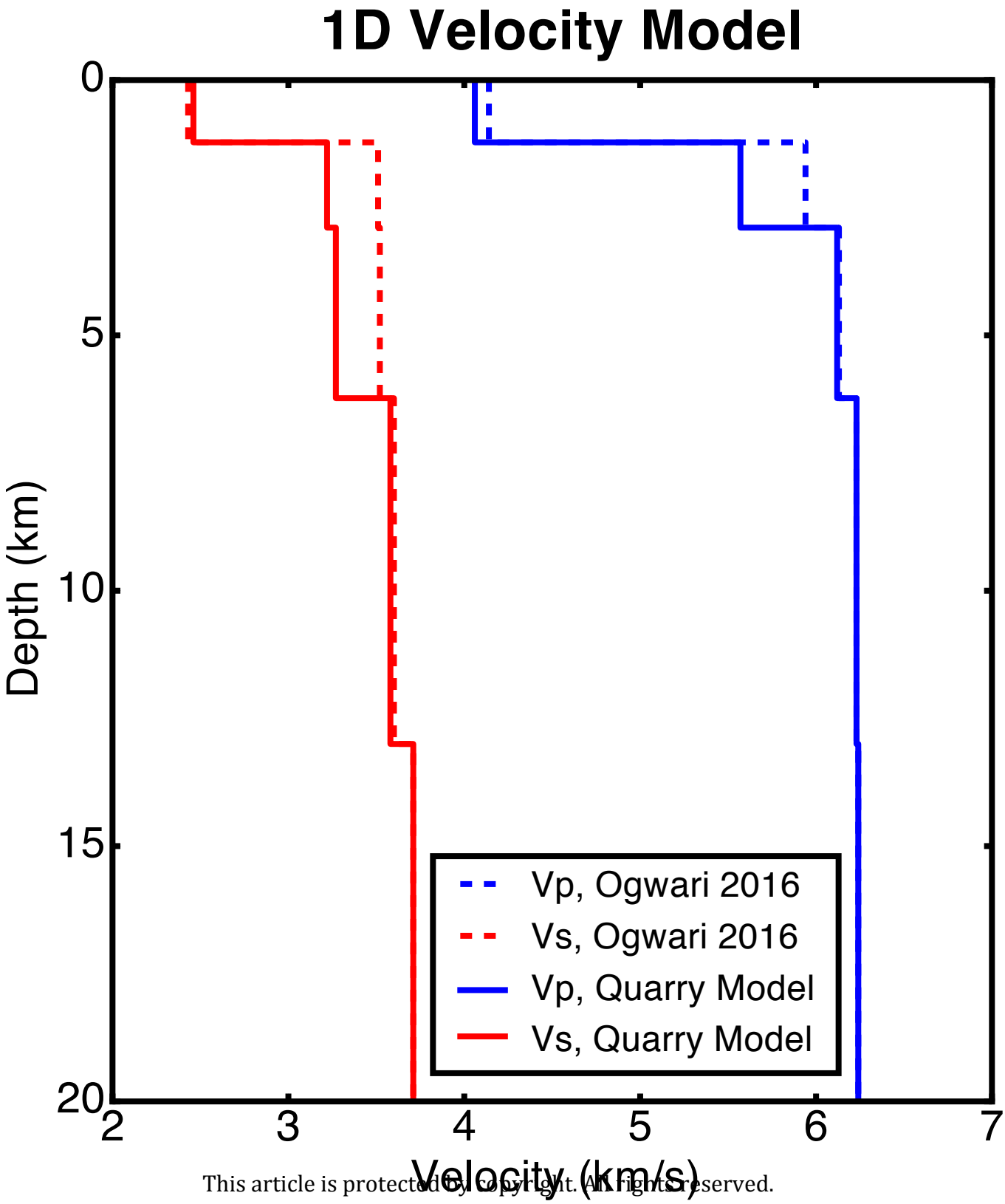


Figure 6.

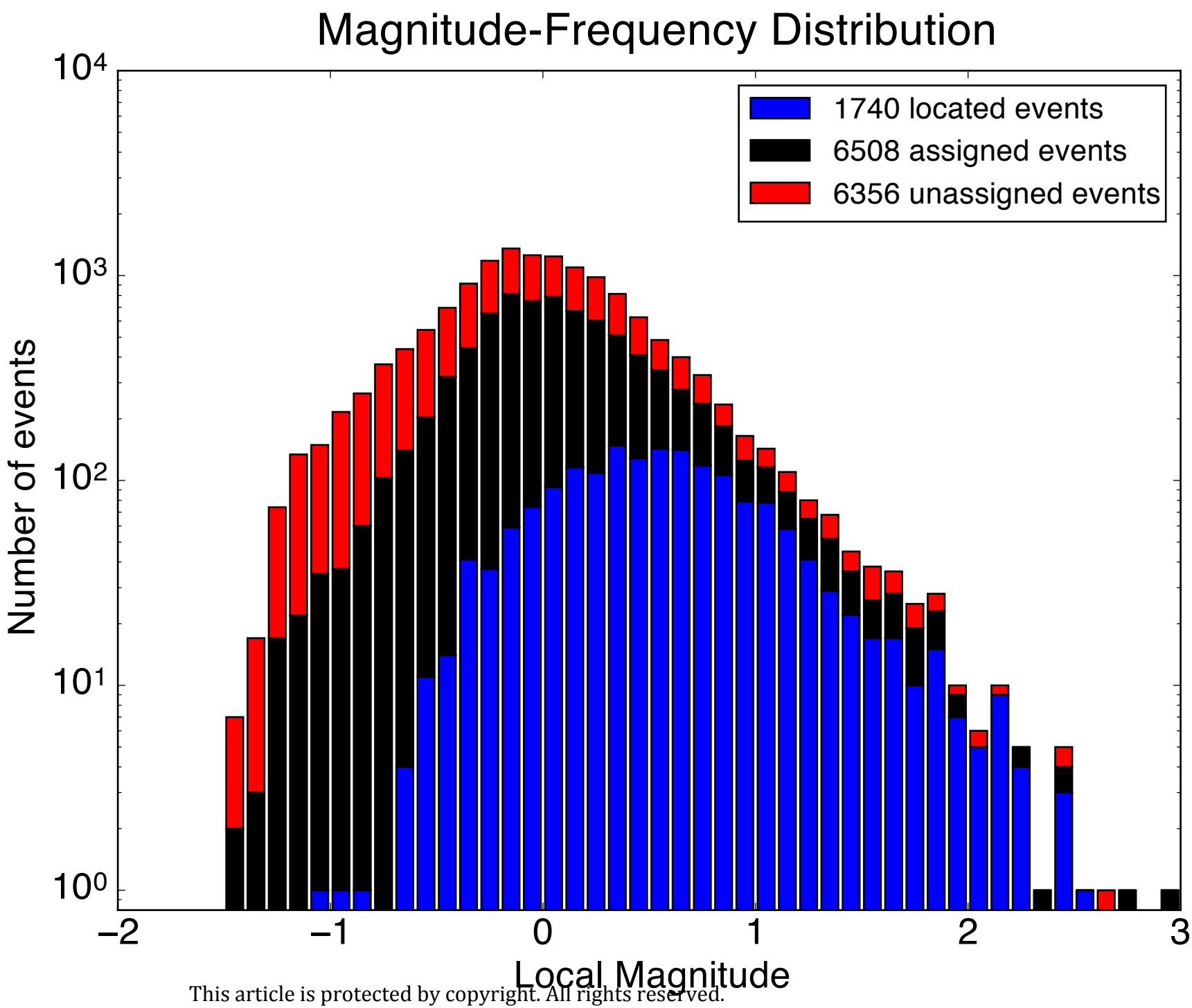


Figure 7.

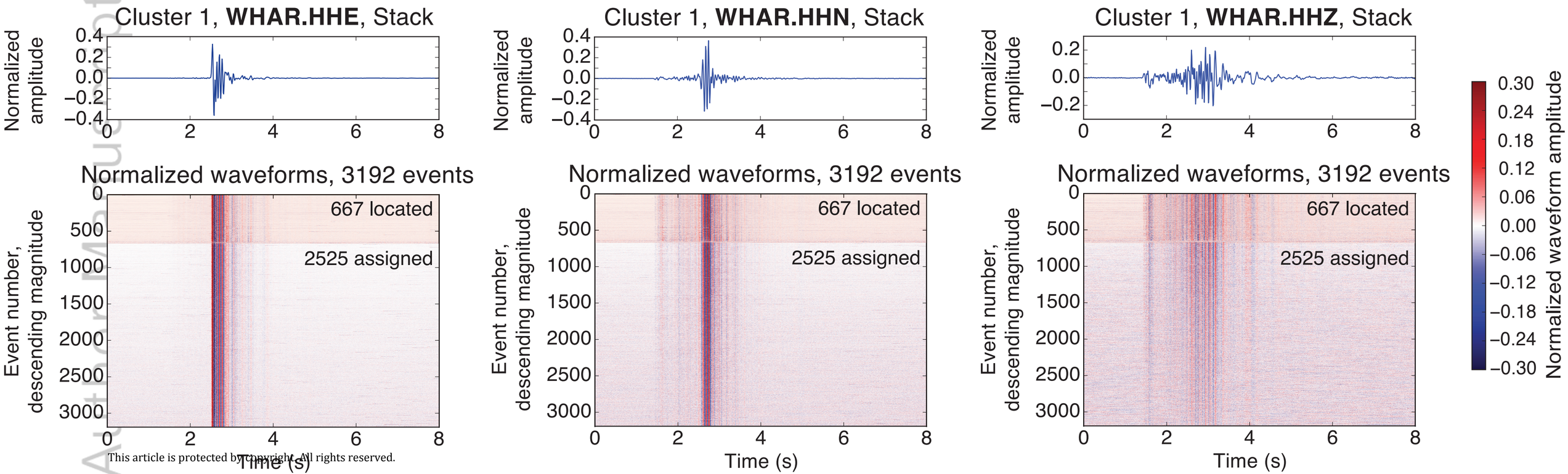
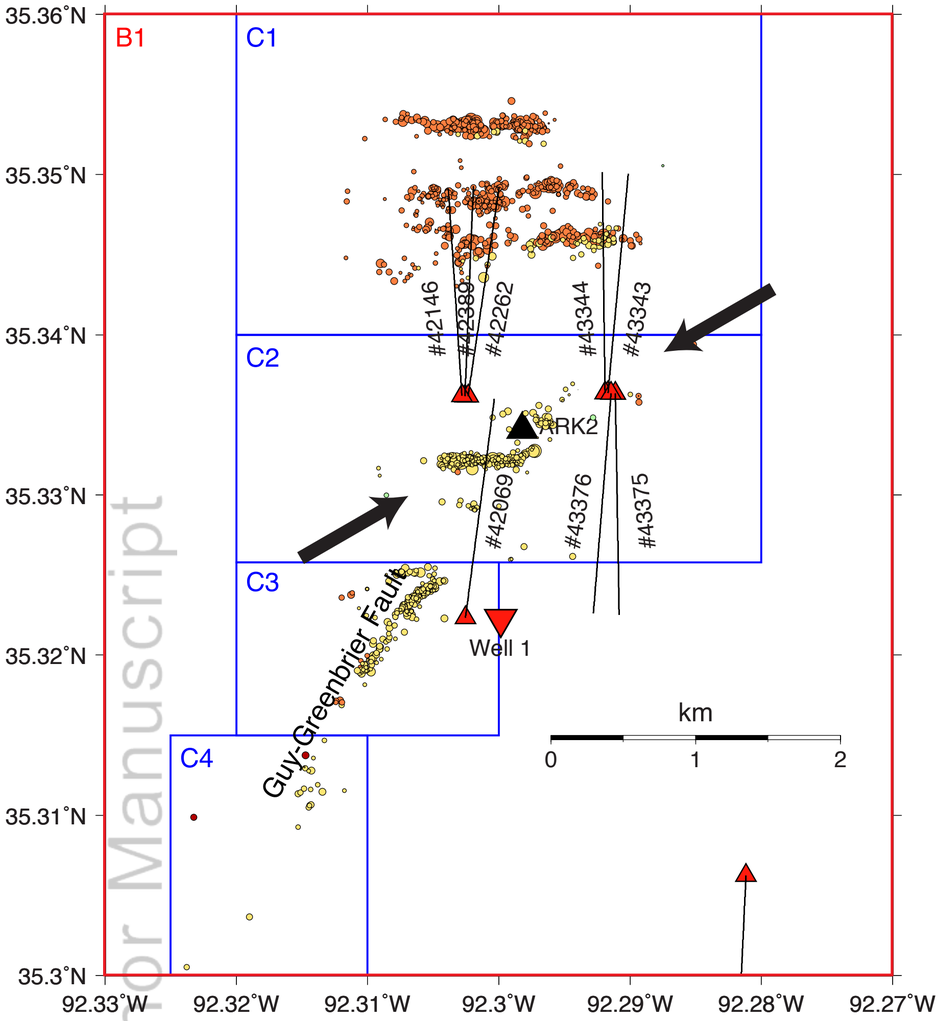
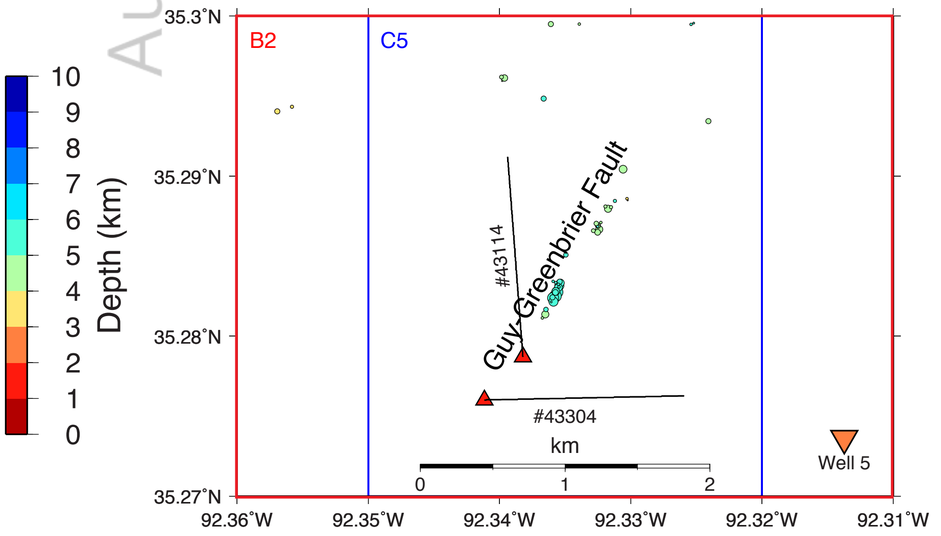


Figure 8.





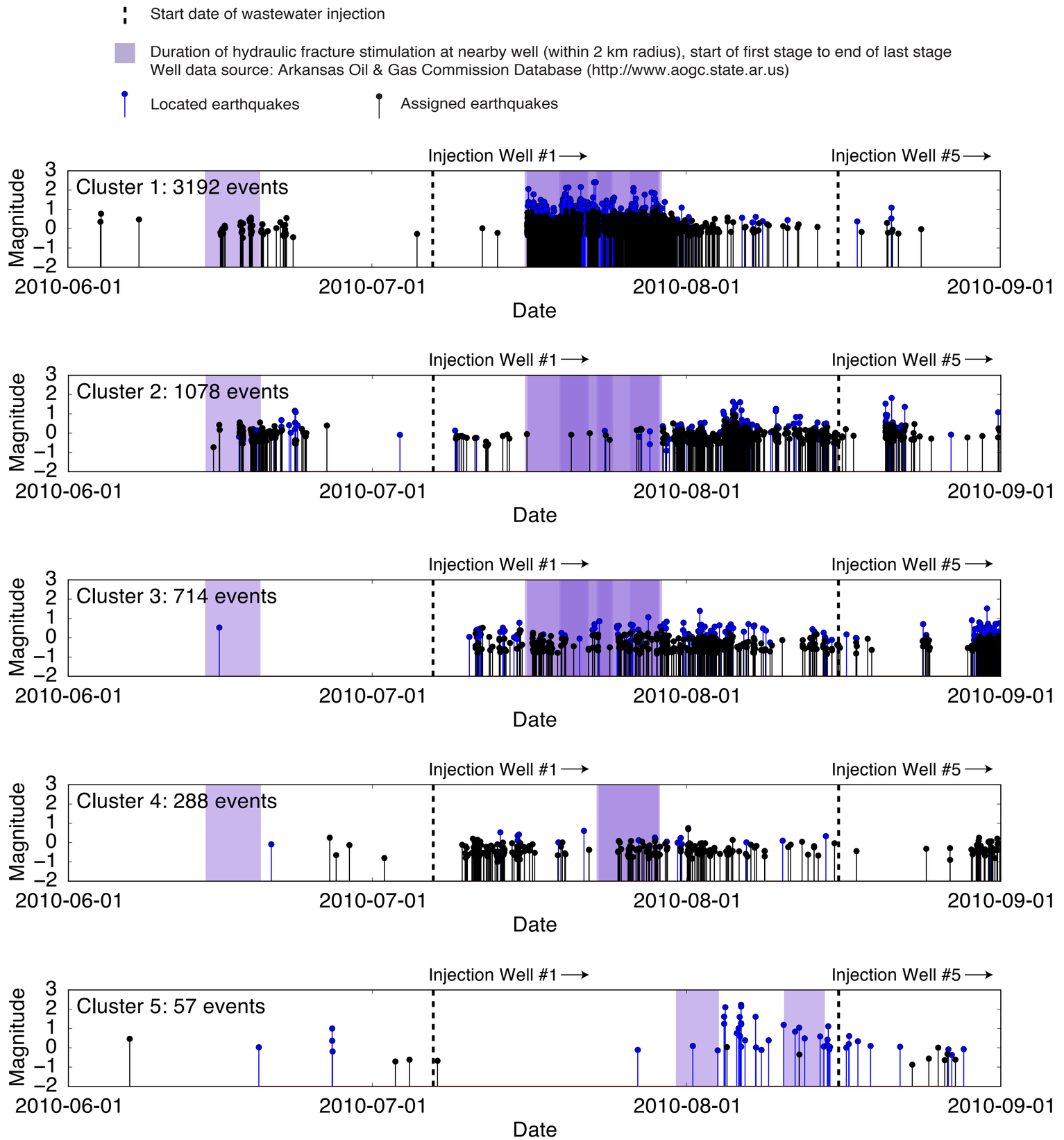


Figure 9.

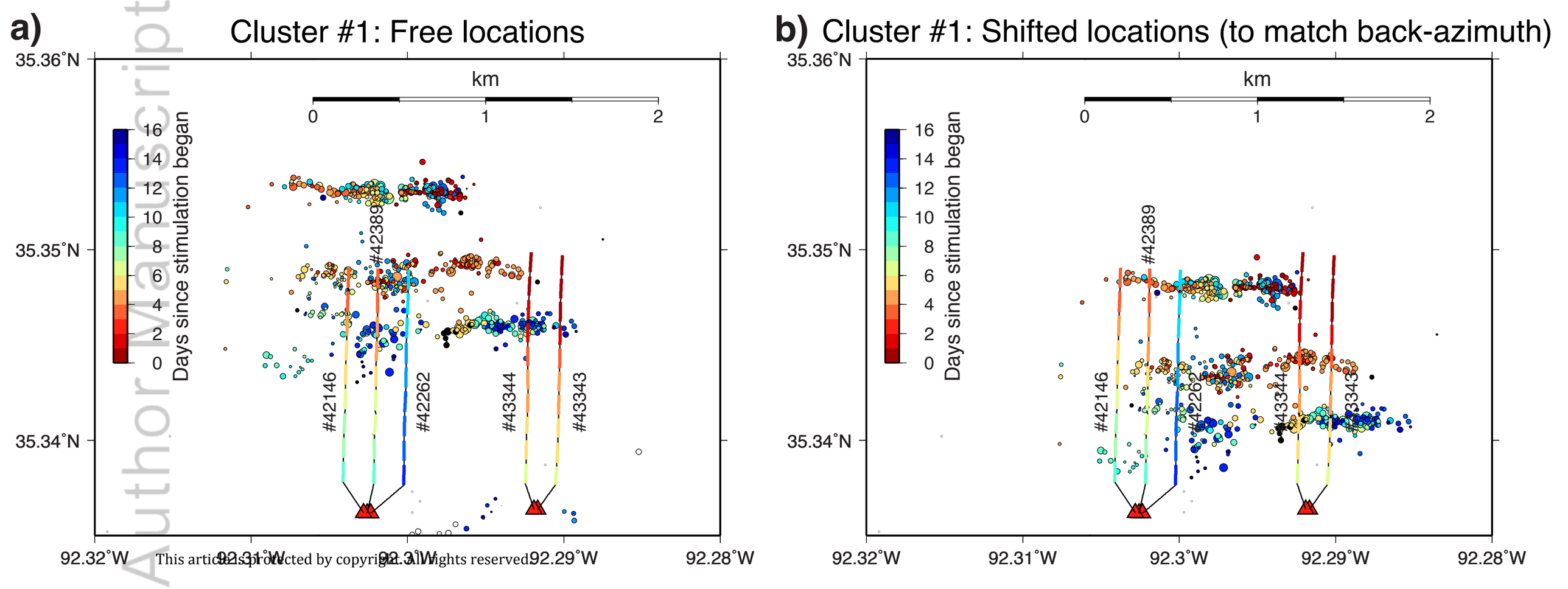


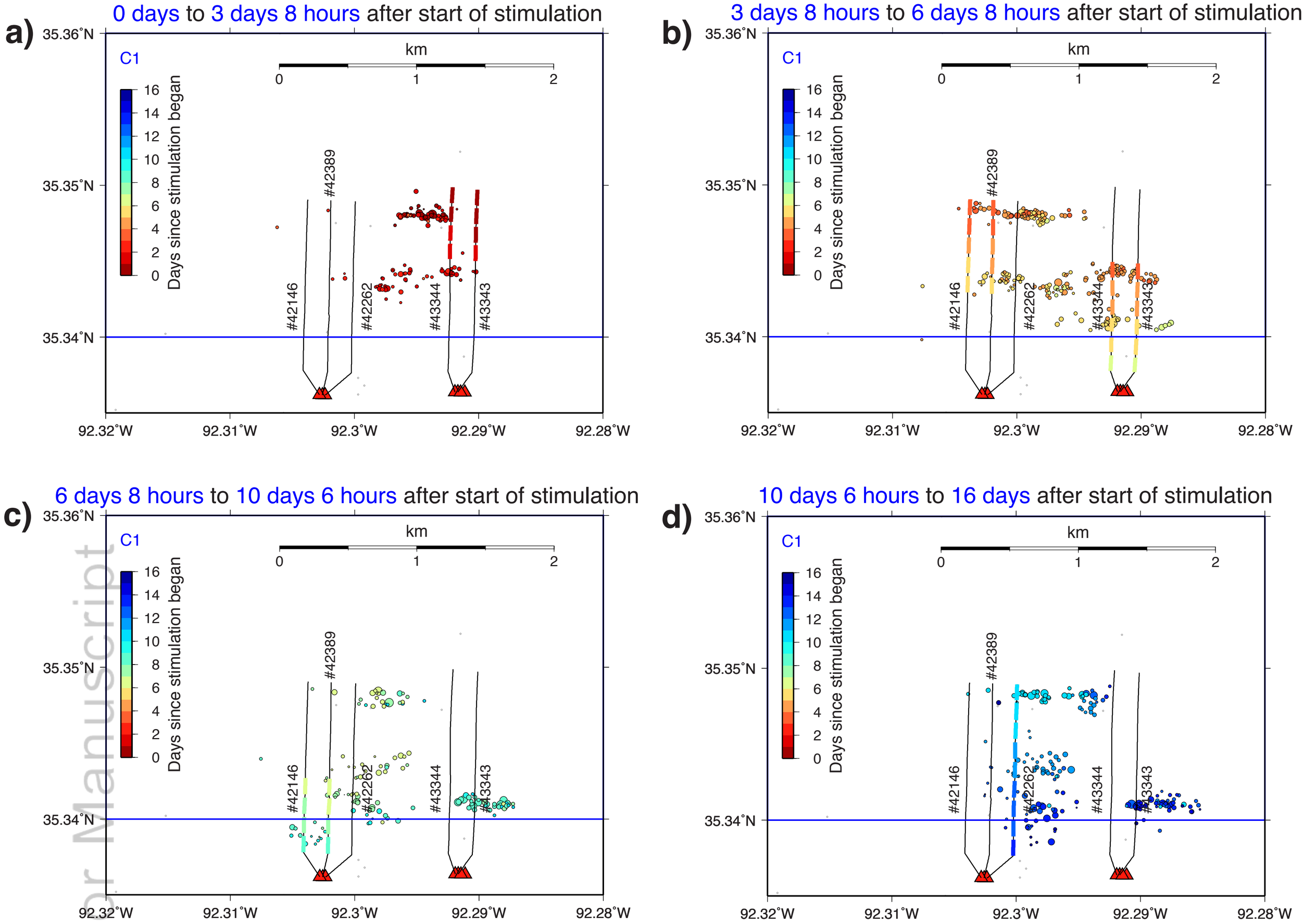
Figure 10.

a) 0 days to 3 days 8 hours after start of stimulation

b) 3 days 8 hours to 6 days 8 hours after start of stimulation

c) 6 days 8 hours to 10 days 6 hours after start of stimulation

d) 10 days 6 hours to 16 days after start of stimulation



e)

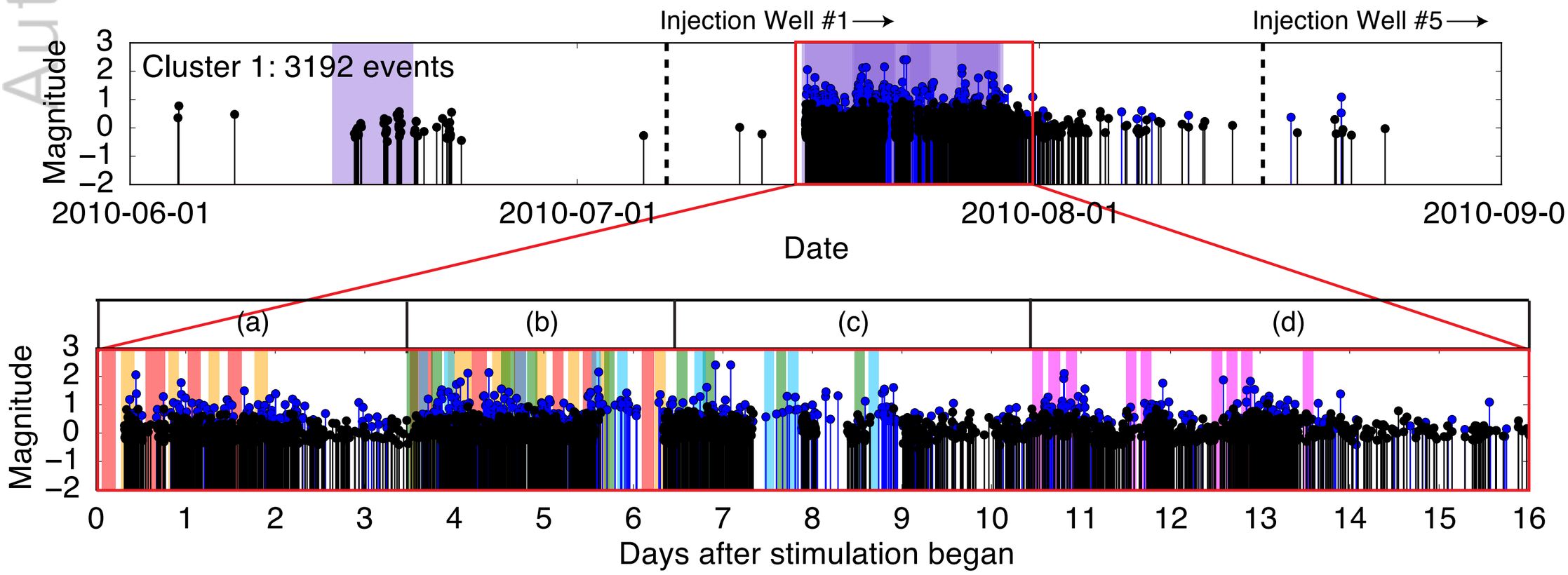
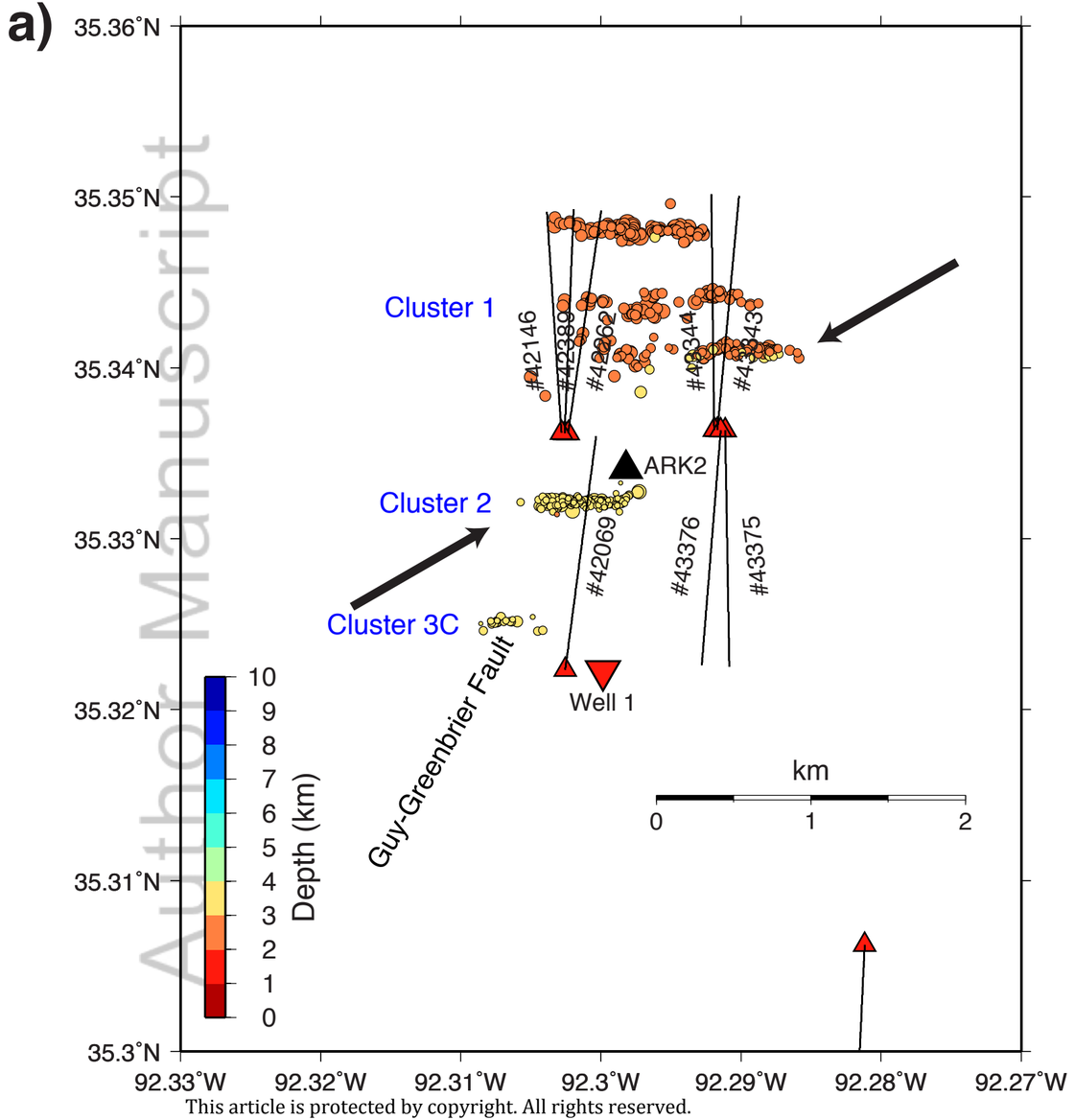


Figure 11.



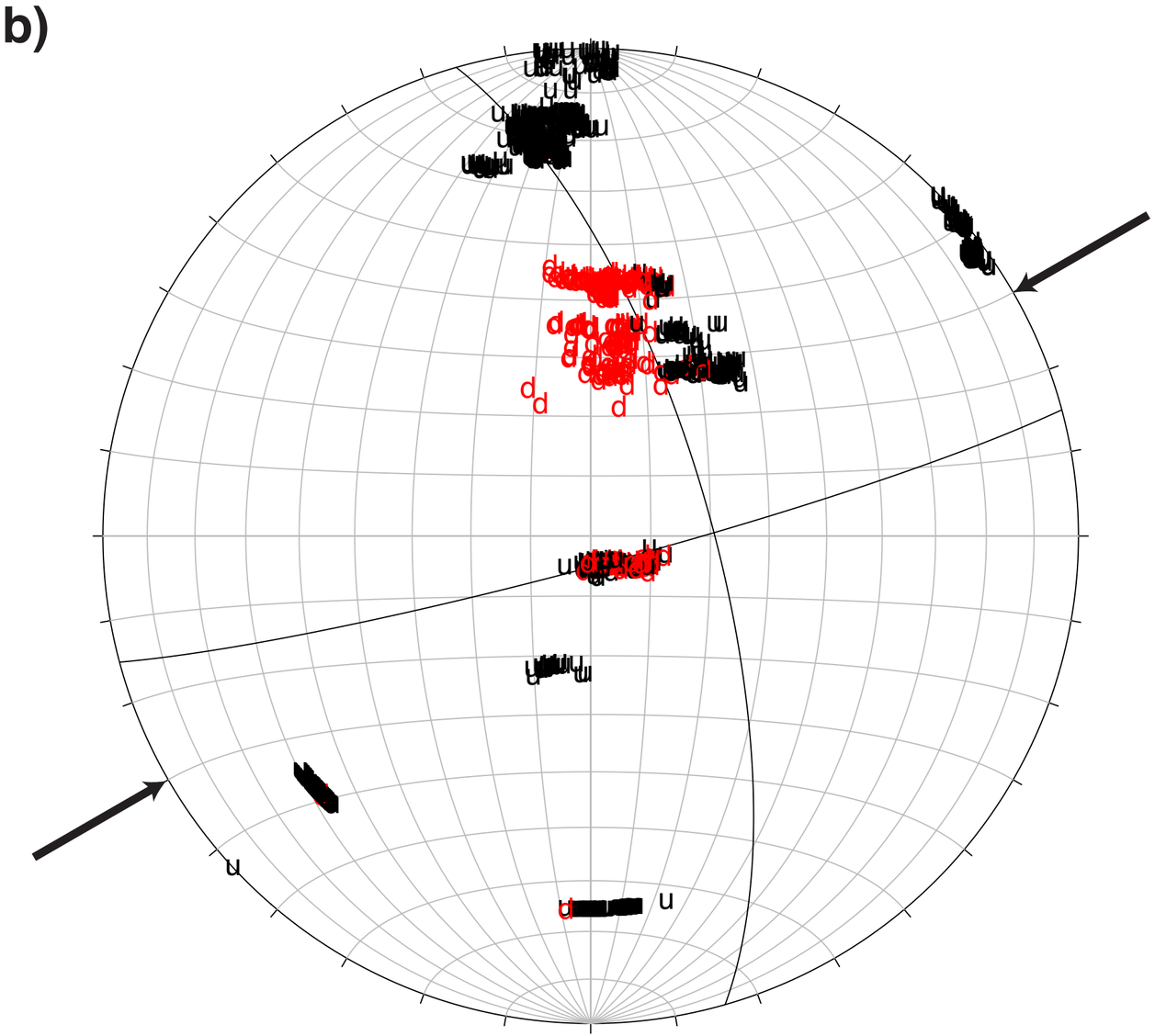


Figure 12.

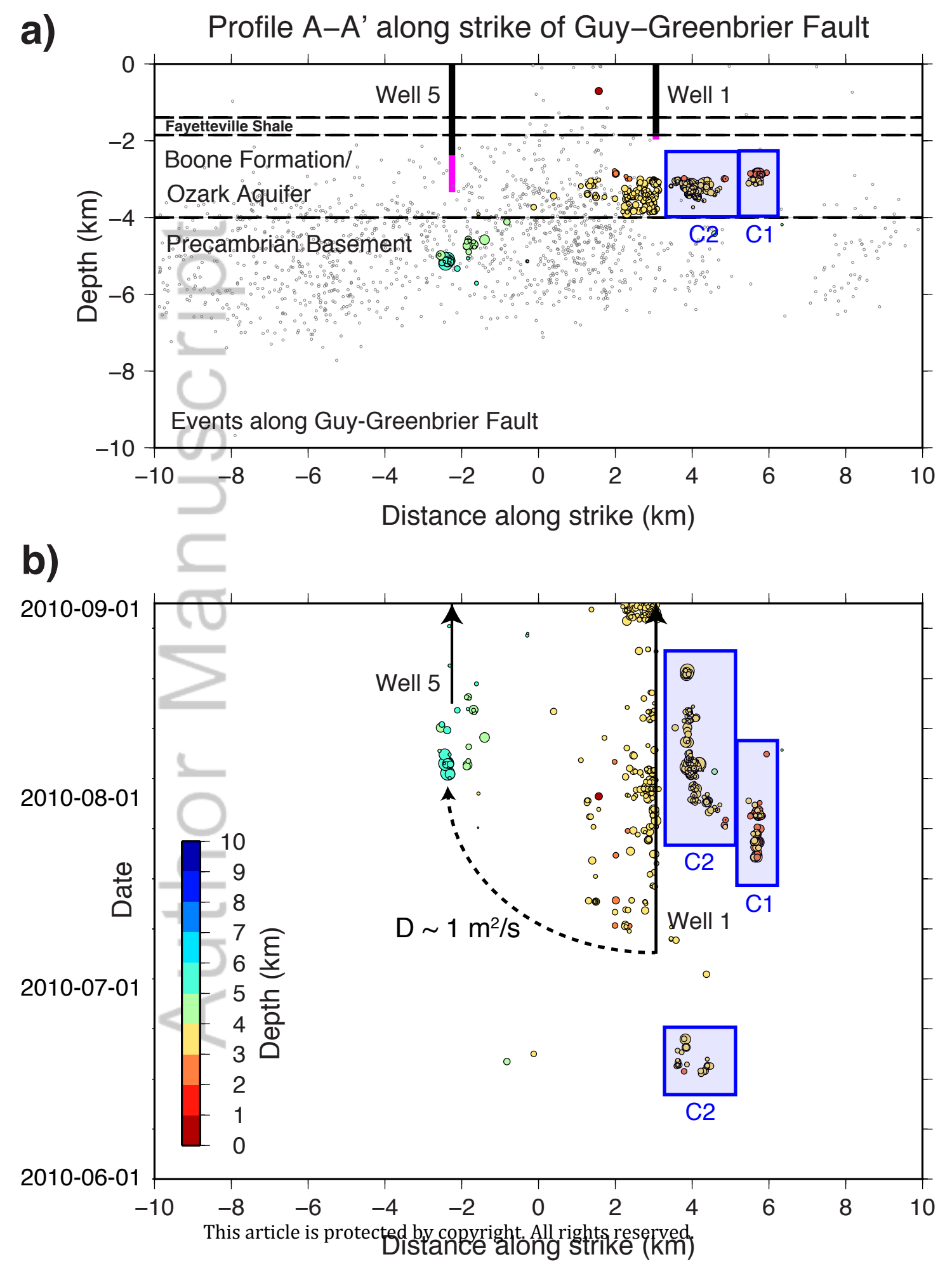


Figure 13.

Start date of wastewater injection

Duration of hydraulic fracture stimulation at nearby well (within 2 km radius), start of first stage to end of last stage
Well data source: Arkansas Oil & Gas Commission Database (http://www.aogc.state.ar.us)

Located earthquakes

Assigned earthquakes

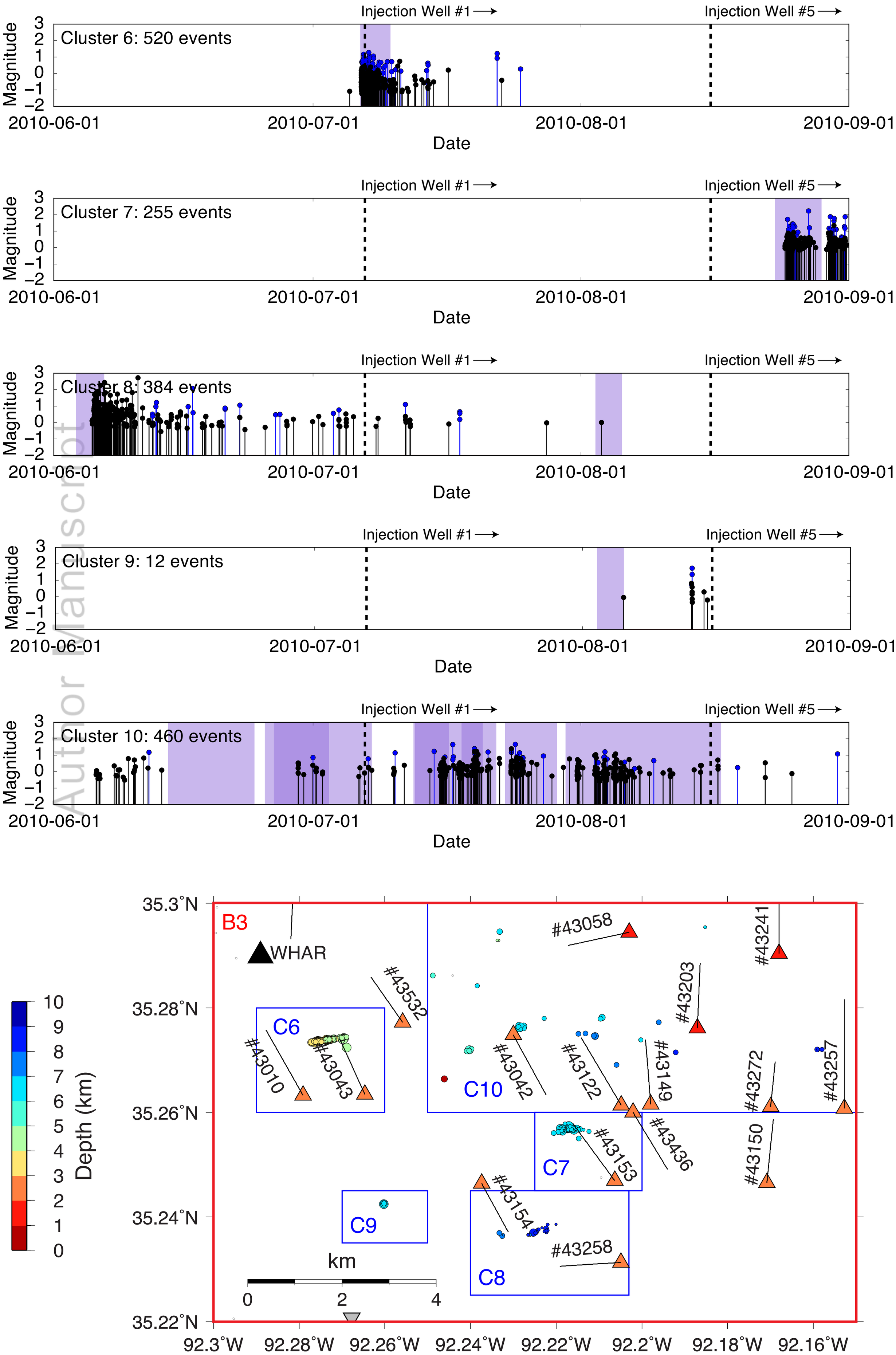


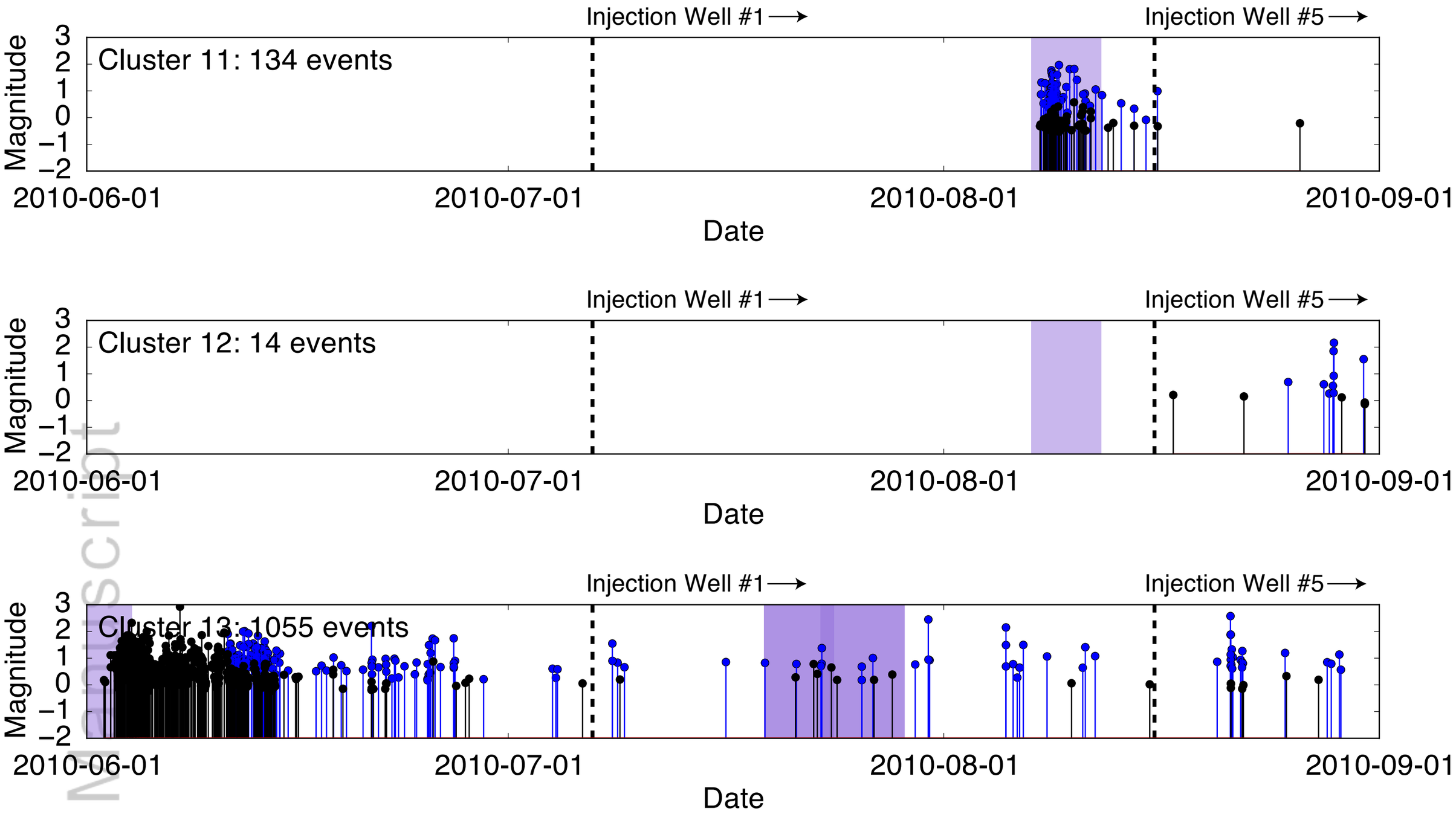
Figure 14.

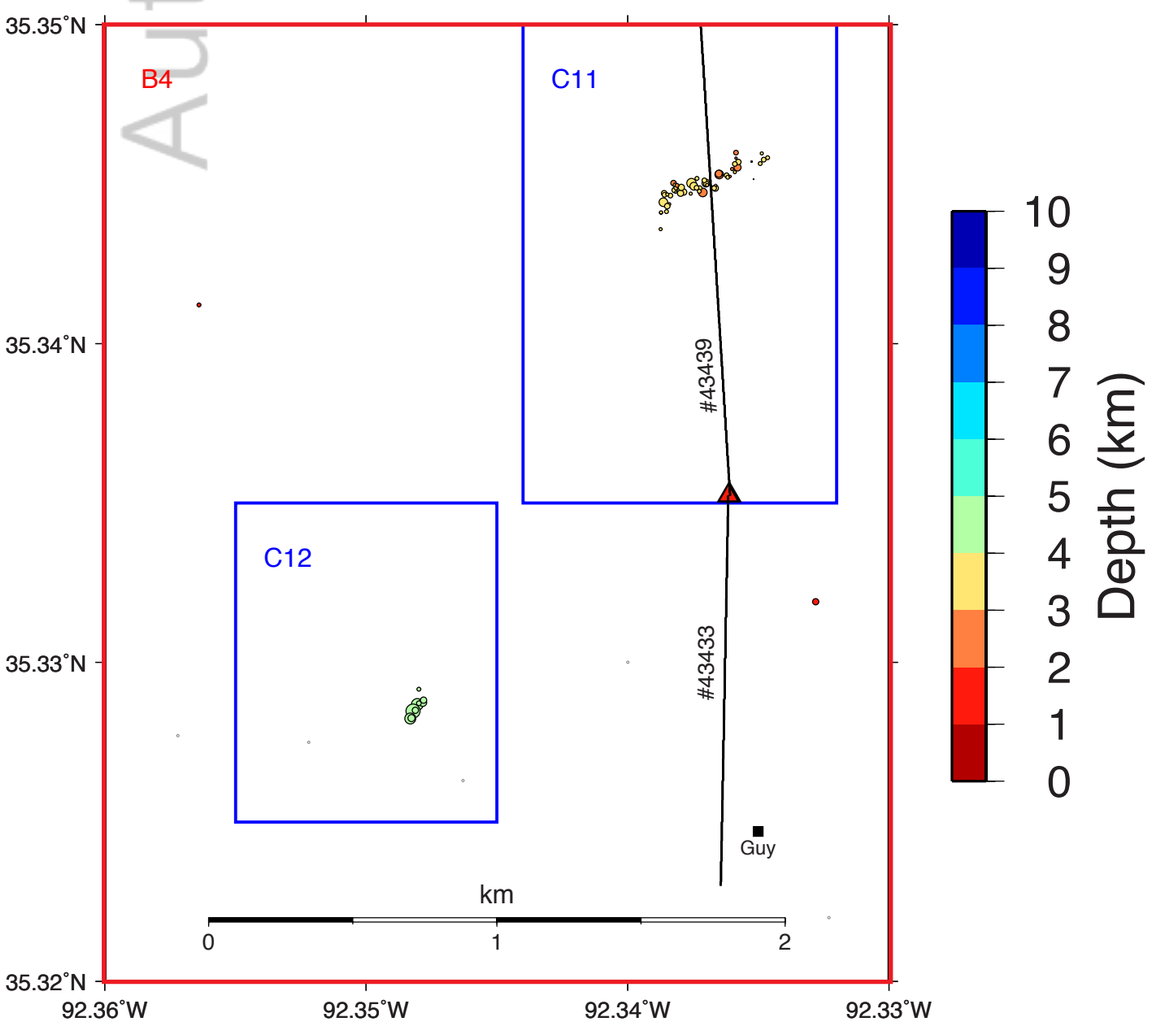
Start date of wastewater injection

Duration of hydraulic fracture stimulation at nearby well (within 2 km radius), start of first stage to end of last stage
Well data source: Arkansas Oil & Gas Commission Database (http://www.aogc.state.ar.us)

Located earthquakes

Assigned earthquakes





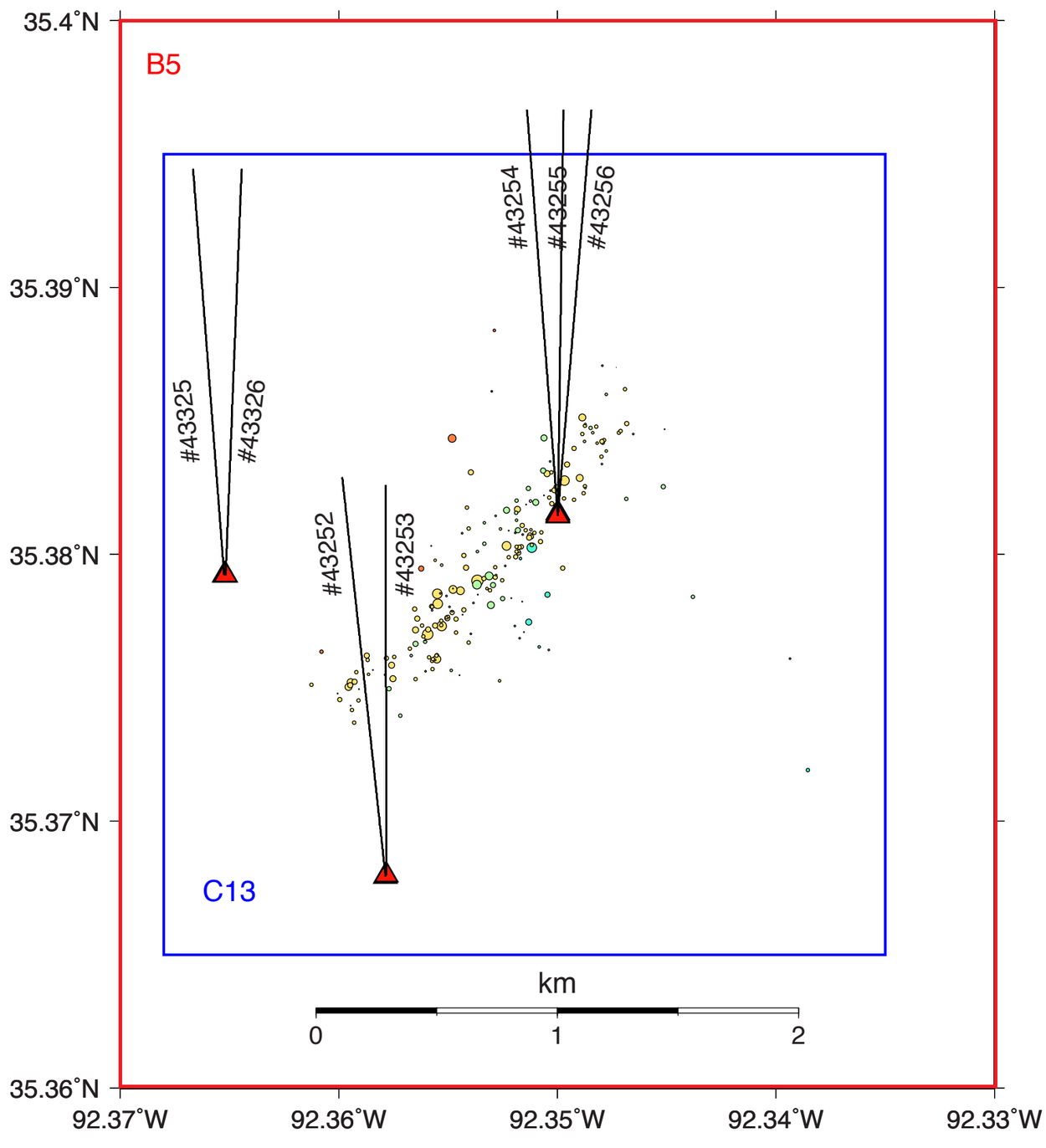


Figure 15.

Start date of wastewater injection

Duration of hydraulic fracture stimulation at nearby well (within 2 km radius), start of first stage to end of last stage
Well data source: Arkansas Oil & Gas Commission Database (http://www.aogc.state.ar.us)

Located earthquakes

Assigned earthquakes