Final Technical Report

for the USGS Earthquake Hazards Program Grant

G22AP00243

June 01, 2022 – May 31, 2023

"Probabilistic identification of seismicity triggered by oil and gas activities and its effects on seismic hazard"

Dr. Iason Grigoratos, Prof. Stefan Wiemer

iason.grigoratos@sed.ethz.ch, stefan.wiemer@sed.ethz.ch

Swiss Seismological Service (SED), ETH Zurich

Sonneggstrasse 5, Zurich, Switzerland, +41 44 633 38 57

August 2023

Table of Contents

Abstract

The present study employed an established semi-empirical earthquake recurrence relation to model the observed seismicity in Central and Eastern United States (CEUS) given the time-history of oil and gas activities. The model is founded on established physics-based principles that directly into account the external driving forces, i.e. the oil and gas activities, via the injection rates. Although the model is generic to be applicable in any region and covers both HF and wastewater disposal (SWD) cases, it matches well the earthquake time-history across various scales and regions. Besides hindcasting past seismicity rates, our main goal is to employ the model to distinguish the causal factors of induced seismicity across CEUS, including areas with overlapping oil and gas activities, something that has always been a challenging issue that is lacking a homogeneous response in the existing literature. The release of national maps identifying areas with induced seismicity, where the long-term USGS hazard models might deviate significantly from the near-term rates, would benefit the public, various stakeholders, insurance companies, and state-regulators.

To that end, we unified about 30 earthquake catalogs and collected HF data for all the CEUS states (west of W 117). Next, we examined cases of seismicity induced by HF in Oklahoma, the Raton basin, the Eagle Ford, the Fort Worth basin, and the Permian basin. We also collected data and analyzed the following regions in terms of a relationship between seismicity and SWD: Oklahoma, Kansas, the Raton basin, the Eagle Ford, the Fort Worth basin, the Midland basin and the Delaware basin.

We concluded that wastewater disposal is a main driver behind the elevated seismicity in the Delaware basin (both shallow and deep), the Midland basin (deep), southern New Mexico (deep), central Oklahoma (deep), and south and central Kansas (deep), with HF triggering significant levels of seismicity in the Delaware basin, the Eagle Ford, and in south and southwestern Oklahoma. Our conclusions were in very good agreement with previous studies. The present investigation identified for the first time clusters triggered by SWD in the Eagle Ford (both in the center of the basin and along the Mexico-TX border), as well as clusters triggered by HF in central and northern Oklahoma. Furthermore, we were able to link all the clusters of Culberson county and the enigmatic Irving-Dallas sequence to deep SWD. We were also able to link at a statistically significant level six events with magnitude 5 and above since 2010 within our basins to wastewater disposal operations. Finally, we highlight three events that require further research in terms of their triggering process: the 2015 Venus M_w 4.0 earthquake in the Fort Worth basin, the 2011 Trinidad M_w 5.3 earthquake in the Raton basin and the 2011 Prague M_w 5.7 earthquake in central Oklahoma.

1. Introduction

As a starting point, we should clarify that this report summarizes the tasks only of the 1st year of the original 2-year proposal, given that only that portion of the project has received funding thus far. The tasks outlined for the 2nd year, have been resubmitted as a new stand-alone proposal for FY2024.

In the past decade, several US states, have experienced unprecedented seismicity rates, following an increase in the volumes of fluids that are being injected underground via hydraulic fracturing stimulations (HF) and the disposal of the co-produced waste-fluids. These earthquakes mostly affect infrastructure that was designed with little to no consideration for seismic loading, rendering them particularly vulnerable to earthquake motions. The magnitudes of these earthquakes are not always small, with at least 6 earthquakes of magnitude larger than 5 that have been linked to oil and gas activities in Oklahoma and Texas. This has caused unrest in parts of the exposed public, forcing the industry and the regulators to either stop or relocate certain energy-production activities. That said, the spatial distribution of seismicity has been highly variable. Many major plays, such as the Bakken in North Dakota, have experienced little to no induced seismicity despite thousands of wells being in operation (Weingarten et al., 2015; Verdon et al., 2016), whereas others in Oklahoma and Canada have experienced so much seismicity that regulators had to intervene (Goebel et al., 2019; Schultz et al., 2020).

When it comes to HF for oil and gas exploration globally, varying levels of seismicity have been detected in British Columbia (e.g., Kao et al., 2018; Roth et al., 2020), Alberta (e.g., Schultz et al., 2015; Bao and Eaton, 2016; Schultz and Wang, 2020), Oklahoma (Holland, 2013), Texas (Fasola et al., 2019; Grigoratos et al., 2022), Arkansas (Yoon et al., 2017), Ohio (Brudzinski and Kozlowska, 2019), China (e.g., Lei et al., 2019; Yang et al., 2020), and the United Kingdom (Kettlety et al., 2021). The vast majority of those earthquake sequences get initiated during the HF stimulations and last a few days (Schultz et al., 2020). The earthquakes tend to occur within 2 km of a stimulating HF well (Brudzinski and Kozlowska, 2019; Roth et al., 2020; Schultz and Wang, 2020; Schultz et al., 2020; Kettlety et al., 2021), with only one high- spatial-resolution data set indicating distances up to 4 km (Yu et al., 2019). That said, location uncertainties in regional catalogs often bias toward larger distances.

Wastewater (or saltwater) disposal (SWD), on the other hand, has caused widespread seismicity in Oklahoma (Grigoratos et al., 2020a), Kansas (Rubinstein et al., 2018; Ansari and Bidgoli, 2020), North Texas (Quinones et al., 2019), and along the Raton basin (Rubinstein et al., 2014), with isolated cases linked to specific wells in Ohio (Kim, 2013; Brudzinski and Kozlowska, 2019), East Texas (Frohlich et al., 2014), Arkansas (Horton 2012; Park et al., 2020), and China (Wang, Lei, et al., 2020). At a regional scale, the earthquakes usually occur at or below the depth of the target formation, with basement-fault reactivation in the presence of fluid pathways often leading to larger magnitudes (e.g., Chang and Yoon, 2020).

In this study, we unified about 30 earthquake catalogs and collected HF data for all the Central and Eastern US states (west of W 117; CEUS). Next, we examined cases of seismicity induced by HF in Oklahoma, the Raton basin, the Eagle Ford, the Fort Worth basin, and the Permian basin. We also collected data and analyzed the following regions in terms of a relationship between seismicity and SWD: Oklahoma, Kansas, the Raton basin, the Eagle Ford, the Fort Worth basin, and the Permian basin. Identifying the causal factors behind elevated seismicity rates in areas with multiple ongoing oil and gas activities has always been a challenging issue that is lacking a homogeneous response. Our study addresses this issue in a probabilistic way and can thus inform the induced seismicity zonation maps that the USGS uses to exclude earthquakes from their tectonic hazard maps (Petersen et al. 2012). It also highlights areas where intense oil and gas activities do not cause seismicity or areas where more than one causal factor is triggering seismicity. Such regions are of particular interest for future studies.

2. Data

2.1 Earthquake catalogs

We compiled a new unified earthquake catalog for the CEUS with magnitudes above 1 over the time period of January 2000 to February 2023, merging multiple different sources (Table 1). Duplicate events among catalogs were removed (Grigoratos et al., 2023), with hierarchy rules applied when more than one location or magnitude solution was available. The margins used during the duplicate-search are listed in the Supplemental Material.

Table 1. Sources of parametric earthquake data.

Source	Time-period		Magnitude Scale	Region/State	M ≥
ComCat (ANSS)	2000	Feb 2023	varied	CEUS	2.0
Mueller (2018)	1900	2018	proxy M_w	CEUS	1.0
SLU (Herrmann et al., 2011)	1962	Feb 2023	M_w	CEUS	1.5
TexNet	2017	Feb 2023	M_L	TX	1.0
TexNet Research [1]	2019	2021	M_L	TX	1.0
NTXES [2]	Oct 2008	Jul 2021	M_L	north-central TX	-1.3
Frohlich et al. (2016)	1902	2015	varied	TX	3.0
Fasola & Brudzinski (2023) [3]	Mar 2014	Feb 2020	M_L	Eagle Ford (TX)	-0.5
Frohlich & Brunt (2013) [4]	Apr 2010	Sep 2011	proxy m_{bLg}	Eagle Ford (TX)	1.3
Skoumal et al. (2020a) [5]	2017	2018	varied	Delaware basin (TX)	-1.0
Gan & Frohlich (2013) [4]	May 2009	De 2010	unknown	Permian basin (TX)	1.7
Walter et al. (2018) [4]	2008	Nov 2015	unknown	Panhandle (TX)	-1.6
Nakai et al. (2017a) [4]	Feb 2008	Feb 2010	M_L	NM & CO	-0.1
temporary USGS network in NM [6]	Oct 2010	Aug 2022	M_L	south NM	0.0
New Mexico Tech (NMT)	2017	Feb 2023	M_d	NM	1.0
Pursley et al. (2013)	1998	2009	M_d	NM & CO	2.0
Rubinstein et al. (2014; their Table 2)	1973	2013	varied	Raton basin	3.8
Wang et al. (2020) [7]	Jul 2016	Jul 2020	M_L	Raton basin	-2.2
Oklahoma Geological Survey (OGS)	1983	Feb 2023	$M_L/m_{bLg}/M_w$	OK	1.0
Schoenball & Ellsworth (2017)	May 2013	Nov 2016	varied [12]	OK	0.7
Skoumal et al. (2019) [8]	May 2010	2016	M_L [12]	OK	1.8
Holland (2013)	Jan 2011	Jan 2011	M_L [12]	south-central OK	1.8
Kansas Geological Survey (KGS) [9]	1983	Feb 2023	varied	KN	2.0
Rubinstein et al. (2018) [10]	Mar 2014	Dec 2017	M_L [12]	southern KN	-0.4
Arkansas Geological Survey (AGS)	1811	May 2021	unknown	AR	0.0
Ohio DNR	1901	Aug 2022	unknown	OH	-0.1
PASEIS (Nyblade & Homman, 2016) [4, 11]	Feb 2013	Jun 2015	M_L	PA	1.1
PASEIS [11]	2016	Aug 2022	M_L	PA	0.8

[1]extension of the public TexNet catalog that includes earthquakes of limited manual review due to their smaller size

[2]extended version provided by Heather DeShon (pers. comm., August 2022). Includes events from SMU (Quinones et al., 2019), Frohlich et al. (2012), ComCat, Ogwari er al. (2018), Scales et al. (2017), and template-matched unpublished data from Mike Brudzinski.

[3]update to Fasola et al. (2019) with recomputed magnitudes. Fasola et al. (2019) and Fasola & Brudzinski (2023) did template-matching based on TexNet events for March 2014 - December 2018, and January 2019 - February 2020, respectively. Both studies assigned the parent-location to newly detected events.

[4]used EarthScope USArray data

[5]employed template matching with the parent location assigned to newly detected events, three different subcatalogs due to three different parent-lists (TexNet; ComCat; NMT) with duplicates present between them (which we removed).

[6]this catalog is not uploaded to ComCat, because there is limited quality-control applied to it. The magnitude scale seems to require significant further calibration to match the magnitudes reported by NMT and TexNet.

[7]employed a machine-learning phase picker and relocated hypocenters, very low claimed M_c (below 0).

[8]employed template matching with relocations for newly detected events; added 0.38 M_L units to each event following Kavoura et al. (2020)

[9]no focal depths

[10]extended by one year to December 2017

[11]mining blasts were removed, see Nyblade & Homman (2016) and Homman & Nyblade (2020)

[12]we used the proxy M_w values already computed by Grigoratos et al. (2020a)

The hierarchy (in descending order) for the origin (time and location) solutions was: TexNet, Holland (2013), Walter et al. (2018), Gan & Frohlich (2013), Nakai et al. (2017a), Rubinstein et al. (2018), KGS, Schoenball & Ellsworth (2017), NTXES, OGS, USGS NM network, PASEIS (after 2016), Ohio DNR, Pursley et al. (2013), Frohlich et al. (2016), Fasola & Brudzinski (2023), ComCat, Rubinstein et al. (2014), PASEIS (2013-2015), AGS, NMT, TexNet Research, Skoumal et al. (2020a), Wang et al. (2020), Skoumal et al. (2019), SLU, Mueller (2018). The order was mostly based on expert opinion and on basic principles outlined in Grigoratos et al. (2023), such as assuming that local networks have better station coverage and more accurate velocity models. Non-located template-matched events were given low priority. The events from Schoenball & Ellsworth (2017) were limited to latitudes above 35 and below 37, because the poor station-coverage outside these limits led to very large epicentral differences when compared to other sources. Solutions of Schoenball & Ellsworth (2017) with focal depths above 15km were also discarded for the same reason. Furthermore, Rubinstein et al. (2018) has several events located within Kansas that probably occurred in northern OK (according to other sources). Finally, we did not compile the earthquakes detected by the Lajitas array in south-central Texas (Frohlich et al., 2020), because they had very large epicentral differences (up to 40 km), when compared to their TexNet duplicates (Grigoratos et al., 2020a).

The hierarchy (in descending order) for the magnitude solutions was: SLU, ComCat, TexNet, Holland (2013), OGS, Rubinstein et al. (2018), KGS, USGS NM network, Rubinstein et al. (2014), Fasola & Brudzinski (2023), Ohio DNR, Pursley et al. (2013), NMT, PASEIS, AGS, NTXES, Frohlich et al. (2016), Walter et al. (2018), Gan & Frohlich (2013), Nakai et al. (2017a), TexNet Research, Wang et al. (2020), Skoumal et al. (2020a), Skoumal et al. (2019), Schoenball & Ellsworth (2017), Mueller (2018). We gave priority to original M_w estimates (SLU, ComCat), and to state-wide networks. Low priority was given to template-matched events (due to their relative magnitudes), and to Mueller (2018) given the issues identified in Grigoratos et al. (2020a). In hindsight, we should have given lower priority to the magnitudes of the USGS NM network, given the now known issues with the calibration of its magnitude scale. That said, in any case, the magnitudes play a very minimal role in our analyses.

The unified catalog was split into various subcatalogs to account for spatio-temporal variations in the magnitude of completeness (M_c). Subcatalogs were created for the Delaware basin, the Midland basin, the Eagle Ford (EF), the Fort Worth Basin (FWB), the NM-CO border, and for OK-KN. The magnitude of completeness was estimated for each subcatalog for different time-periods based on the slope of the magnitude-frequency-distribution. The methods we used in this study are not sensitive to the M_c and thus including more events by slightly underestimating the M_c is encouraged. For this reason, a more quantitative approach for the estimation of M_c was not needed. If we were to compute seismic hazard estimates that require accurate b-values, the situation would have been different. A lack of need for accurate b-values is also why we did not homogenize the magnitudes into M_w, and we rather allowed all the various magnitude scales to co-exist. In very broad terms, most magnitudes below 3.5 are local magnitudes, while most magnitudes above 3.5 are moment magnitudes.

Next, we address the issue of declustering, i.e. the removal of foreshocks and aftershocks. Earthquake sequences induced by wastewater disposal and hydraulic fracturing have a common triggering mechanism (one or more wells), and therefore their spatial and temporal distribution would be inherently non-random, even if there were no earthquake-to-earthquake interactions. Of course, such interactions do occur, especially when the seismicity is linked to wastewater disposal; hydraulic fracturing tends to lead to relatively fewer, yet non-trivial, earthquake-to-earthquake interactions. Furthermore, the earthquake recurrence models we use in this study (section 3.2) can only model mainshocks. For all these reasons, we decided to decluster the various subcatalogs. Acknowledging that any declustering method that is agnostic to the main driver of the seismicity rate changes, i.e. the injection rates, cannot fully capture the complex spatio-temporal clustering nature of induced seismicity, we opted for the declustering algorithm of Aden-Antoniow et al. (2022), which utilizes the nearest-neighbor clustering algorithm of Zaliapin et al. (2008). This algorithm is robust for small samples and stable with Mc, yet sensitive to the d- and w-parameters that are defined by the user (we opted for 1.5 and 0.5 respectively, based on sensitivity tests). We applied one further modification by post-processing the clusters to always keep the largest event in each one as the mainshock (instead of the earliest one).

2.2 Injection data

2.2.1 SWD

Sources of wastewater disposal data were the B3 database (for TX, NM and OK), the Oklahoma Corporation Commission (OCC; for OK), the Colorado Oil & Gas Conservation Commission (for the Raton basin only), the KGS (for KN), the Environmental Protection Agency (EPA; for a few counties in OK), Kyle Murray (pers. comm., May 2019; for OK) and the studies by Weingarten et al. (2015; for all states), Barbour et al. (2017; for Pawnee county in OK) and Norbeck and Rubinsten (2018; for OK and KN). The B3 wastewater disposal data are assumed complete

till 2021, but also cover 2022, because B3 extrapolates trends from nearby wells in a proprietary way to fill in any gaps. The OCC data are assumed complete for 2021, with only daily data for the Arbuckle wells (which we convert to monthly) being available for 2022 and early 2023. The KGS disposal are available only through 2021. The Colorado Oil & Gas Conservation Commission disposal data are complete through at least 2022. The EPA data (relevant only for a few counties mostly within OK) are assumed largely incomplete after 2016 (given our sources).

Our analyses required monthly injection rates, the coordinates of the well and is some basins the target-formation. Only wastewater disposal wells were included; Enhanced Oil Recovery (EOR) wells were not taken into account, given that EOR operations aim to stabilize, rather than increase, the pore pressure. The (true vertical) depth was also collected, although not used explicitly, because all our analyses were in 2D. We reviewed some of OCC's data before 2009 against the 1012A forms (when available online) to confirm the units of the reported values and utilized similar corrections made by Murray (pers. comm., May 2019) for the data between 2009 and 2017 (see Grigoratos et al., 2020a). Only for OK, sometimes well metadata and rates for a given year were available from multiple spreadsheets. Priority was given (in descending order) to: manually reviewed 1012A data (by Murray or Grigoratos et al., 2020a), OCC's daily data, OCC's monthly data, OCC's annual data, and Weingarten et al. (2015). Finally, for all wells, we identified and corrected (in an automated way) keystroke errors leading to the omission or repetition of the last digit in the value of a monthly volume.

Regarding the SWD data within the Permian basin, we divided them into two groups, based on whether they were targeting the formations above or below the HF interval (Katie Smye, pers. comm., May 2023). More details about the regional stratigraphy are available in Hennings et al. (2021) and Smye et al. (2021). Notably, the target-formations either directly reported within B3 or inferred from the B3 depths did not produce realistic spatial distributions and are likely highly inaccurate. Finally, we did not employ the Delaware basin pore-pressure models used in Grigoratos et al. (2022), because they were only available for that basin and we wanted to analyze the results consistently across space. That is why we used as input the disposal rates, following of Grigoratos et al. (2020a).

2.2.2 HF

Sources of HF data were the FracFocus Chemical Disclosure Registry (for all states), the IHS Markit databases (for TX), and the OCC *FracNotices* (for OK). Reporting to FracFocus started in 2011 and became mandatory after January 2012, whereas the IHS Markit database dates back multiple decades and (within TX) is sourced from the Texas Railway Road Commission (RRC). The OCC requires a Hydraulic Fracture Notice Form (*FracNotice*) to be filed 48 hours prior to any hydraulic fracturing operations within OK. *FracNotices* were first required to an OCC district office in 2012. Electronic filing became available starting in July of 2016, was required by December of 2016, and is now the mandatory form of submittal. The operators provide the well surface and bottom-hole location, anticipated number of stages, average fluid volume per stage, scheduled start and end dates of operation, and initiation date of well flowback. In summary, for OK, we used HF data through December 2022 (leveraging the *FracNotices*), while for all other states through September 2022 (~5 month reporting lags are common within FracFocus).

Our analyses required the start- and end-date of each HF stimulation, the coordinates of the stimulating well, and the average daily injected fluid volume. The latter was obtained by dividing the reported total volumes by the duration of the stimulation. The (true vertical) depth was also collected, although not used, because all our analyses are in 2D. When a frac-job was reported by both FracFocus and IHS, we used dates and volumes from FracFocus and location and depths from IHS. Within OK, data from *FracNotices* were given priority over FracFocus. When aggregating the HF volumes into blocks, we used the midpoint between the well's "heel" and "toe."

Our pre-processing of the HF data was based on the work of Grigoratos et al. (2022). FracFocus contains one fracture job per entry, with start/end- dates, "Total Base Water Volume" (Volw), "Total Base Non-Water Volume" (Volo), true vertical depth, and surface coordinates for each well. Records from IHS, on the other hand, have multiple entries per fracture job, each with a single date and multiple fluid-types and volumes. We aggregated the IHS entries into fracture jobs merging multi-row entries based on their 14-digit API and reorganized the volumes into the two classes adopted by FracFocus (Volw, Volo). The biggest issue with IHS is the lack of a clear start/end-date for each stimulation. The single date provided is assumed to be the start-date, even though for a significant percentage of the wells this was not the case. Lacking any information on the end-date from IHS, we had to use the injected volume as a proxy for duration. To do that, we used the regression derived by Grigoratos et al. (2022). No meaningful correlation was observed for Volo.

$$\begin{aligned} duration &= 3, && \text{if } Volw < 45000 \text{ bbls} \\ duration &= 19 \cdot log_{10} Volw - 85.4, && \text{otherwise} \end{aligned} \tag{1}$$

Usually, non-water based volumes (Volo) represent <1% of the total amount. However, according to IHS some TX basins had larger (yet single-digit) percentages, for unknown reasons. That is why we decided to use as total volume the sum of Volo and Volw.

3. Methods

3.1 Hypotheses testing for causal factors of induced seismicity

The goal is to investigate what percentage of the observed seismicity within CEUS oil plays is triggered by HF and SWD respectively. Past approaches at addressing this issue usually examined a single causal factor and can be grouped into four types of analyses: spatio-temporal association filters (SAFs), time-series cross-correlations, earthquake clustering analyses and hindcasting correlation ratios. The pros and cons of each of these approaches are outlined in Grigoratos et al. (2022). Here we will focus on the latter approach.

Hindcasting correlation ratios involves hindcasting the observed seismicity with a set of competing statistical models to derive likelihood ratios, which are then converted to p-values via reshuffling tests (McClure et al., 2017). This approach can be applied to identify any type of induced seismicity, as long as there is a reasonable earthquake reccurence model linking the observed seismicity to the suspected human activity. The more physics-based these reccurence models are, the better constrained the results. Recent updates to the method (Grigoratos et al., 2020b) make it independent to the spatial grid needed to aggregate the injection and seismicity data into blocks and account for the presence of background tectonic seismicity. We believe that this is the only approach capable of capturing potential links between distributed seismicity and large-scale injection into permeable formations. This is partly because it can be easily paired with a pore pressure diffusion model, capturing far-field pressure changes observed up to 25km away from disposal wells (Brown et al., 2017; Ansari et al., 2019).

Since several CEUS states, like Texas (Savvaidis et al., 2020; Skoumal et al., 2020a) and Oklahoma (Grigoratos et al., 2020b; Skoumal et al., 2018), are likely affected both by HF and wastewater disposal, with overlapping presence of both types of wells, there is a clear need for a robust hypothesis testing protocol. We believe that the aforementioned hindcasting correlation ratios are best suited for this task. Grigoratos et al. (2020b) have successfully applied this method to identify seismicity due to wastewater disposal in Oklahoma, avoiding spurious correlations with HF. Further updates to the methodology were made to render it applicable to daily hydraulic fracturing (HF) operations and pressure-driven modeling of wastewater disposal (Grigoratos et al., 2022).

Here, we briefly summarize this framework, originally introduced by McClure et al. (2017) to investigate solely wastewater disposal. First, injection and seismicity data are aggregated into spatial blocks of several km at fixed time intervals. Two block-specific hypotheses are constructed: (i) a null hypothesis model that assumes no relationship between injection and seismicity, and (ii) an alternative model that does assume a relationship. The null hypothesis, with total likelihood L_0, assumes that all earthquakes are of tectonic origin and is modelled using the classic Gutenberg–Richter (G-R) relationship relationship (Gutenberg & Richter, 1956):

$$\lambda = 10^{a_{tec} - bm} \tag{2}$$

The alternative hypothesis, with total likelihood L_1, does assume a relationship between injection and seismicity and is modelled using equations 3 and 4 (section 3.2). Once both L_1 and L_0 are computed for each block, we define their ratio as R. Although R values greater than 1.0 (i.e., $L_1 > L_0$) indicate that the alternative hypothesis is more likely than the null hypothesis, this criterion is not sufficient to statistically reject the null hypothesis. To do the latter, we need a reference statistical distribution for the ratio R in which the null hypothesis is true (R_{null}). This reference distribution enables the comparison of the two hypotheses in a rigorous way, even if they differ in model-complexity. R_{null} was empirically generated by reshuffling the injection data (with subsequent recalculation of L_1 and L_0). Implementation details are available in Grigoratos et al. (2020b; 2022). In a way, R_{null} reflects how likely it is for the injection-driven model (equations 3 and 4) to find purely coincidental correlation between the observed seismicity and random injection data. By comparing the generated R_{null} values in each block with the single R value obtained from the real injection and seismicity time-series (R_{obs}), we can statistically determine how confident we are that the improved correlation of the alternative hypothesis is not coincidental. The metric used to quantify this confidence level is the statistical p-value, defined as $p = (\eta + 1)/\kappa$, where η is the number of synthetic datasets with R_{null} greater than the R_{obs} value from the real data, and κ is the total number of synthetic datasets (in our case 100). A p-value of 0.05, for example, translates to a confidence level C of 95%, given that $C = 1 - p$.

In the present study, we need to examine two external causal factors (HF and SWD). From a statistical standpoint, we cannot simply construct a single test to directly compare the two alternative models, because then we would have no null hypothesis to reject. Therefore, we will run the p-value analysis twice, assuming a single alternative causal factor at a time (either HF or SWD), by appropriate manipulation of equation 3. In both runs, the null hypothesis assumed all earthquakes are of tectonic origin (equation 2). This leads to a pair of p-value maps, each one illustrating where significant portions of the observed seismicity have superior correlation with either HF or SWD operations, compared to the purely tectonic baseline.

3.2 Generalized Seismogenic Index model

To hindcast the spatio-temporal changes in the seismicity rates we need a physics-based earthquake recurrence model that takes into account the external driving forces, i.e. the oil and gas activities. Since most of the stress changes are caused either directly or indirectly by the injected volumes, the earthquake recurrence model should use as input the injection rates of HF or SWD. We will employ the framework of Grigoratos et al. (2020a; 2022), who expanded the Seismogenic Index model (Shapiro et al., 2010). The original Seismogenic Index model is itself a modified version of the G-R relationship and predicts that the number of induced earthquakes is proportional to the pore pressure change (dP), which can be approximated by the injected volume. The Seismogenic Index model has been successfully used to hindcast induced seismicity related to hydraulic fracturing and geothermal applications (Dinske and Shapiro, 2016). It was originally developed for fluid injection in a single borehole, with the resulting pore-pressure perturbation (and seismicity) being centered around the well in a hydraulically homogeneous and isotropic fluid-saturated medium (Shapiro 2015). The modifications later made by Grigoratos et al. (2022) made it applicable to large-scale HF operations involving a large number of wells. The modifications made by Grigoratos et al. (2020a) were intended to render it applicable to large-scale wastewater disposal, taking into account the background tectonic rate and the stressing-rate dependency of the time lag between injection and seismicity rate changes (Norbeck & Rubinstein, 2018). All injection and seismicity data are aggregated into blocks of several km, on a monthly basis for SWD and on a daily basis for HF. The two fundamental block-specific equations of the earthquake recurrence model are:

$$\lambda[t] = \left(10^{a_{tec}} + v_{lag}[t] \cdot 10^{\Sigma}\right) \cdot 10^{-bm} \tag{3}$$

$$t_{lag}[t] = \frac{\theta}{v[t]} \tag{4}$$

where λ is the rate of earthquakes with magnitude larger or equal to m, a_{tec} is the background tectonic activity rate above magnitude 0 (in log10 units), b is the log–linear slope of the G-R relation, t is time (in days or months), Σ is the block-specific Seismogenic Index (Shapiro et al., 2010) and v is the monthly disposal rate for SWD (m^3) or the daily injected volume for HF (m^3). The variable t_{lag} is a proxy for the injection-rate-dependent time-lag between injection and seismicity rate changes. For simplicity, it is assumed to be inversely proportional to v (Grigoratos et al., 2020a), with θ being the factor of (inverse) proportionality. Next, v_{lag} is derived from v and t_{lag} as follows: each monthly/daily value of v is shifted in the future based on the corresponding value of t_{lag} (rounded integer) on that day/month *t*. The two block-specific model-parameters, Σ and θ, are estimated using maximum-likelihood, assuming that each month (or day) the declustered seismicity follows a non-homogeneous Poisson process with varying mean rate $\lambda[t]$ (driven by injection). Finally, given that the time-lag for HF is usually a few days, we also tested (for comparison) a constant time-lag t_{lag}^{c}, which we will treat as a model parameter that replaces equation 4 (only for HF). Given that almost all HF-induced earthquakes happen either during or right after the stimulation (Schultz et al., 2020; Verdon & Bommer 2021b), for the HF analyses we limited both $t_{lag}[t]$ and t_{lag}^{c} to values between 0 and 14 days, to avoid overfitting in the presence of other triggering factors (e.g. SWD).

Apart from the two free-parameters Σ and θ, two other parameters need to be defined a-priori to run the calibration, the b-value and a_{tec}. We should note that the actual b-value used in equations 3 and 2 to estimate λ[t] at M equal to M_c is not critical, as long as the same value is consistently used for the calibration of Σ and θ (or t_{lag}^{c}); we opted for a typical b-value of 1. The b-value becomes critical when used to extrapolate the statistics to larger magnitudes, as discussed in Grigoratos et al. (2020a). The value of a_{tec} for each examined region was approximated (via extrapolation assuming a b-value of 1.0) based on the regional ComCat data above M 3 between 1985 and 2009. Data only through 2009 were used because a change-point in the CEUS seismicity rates was observed in 2009 (Ellsworth, 2013). Our results are not sensitive to the a_{tec} value, because it is usually orders of magnitude smaller than the implied a-value from induced seismicity; even doubling or halving of a_{tec} has negligible effects on our results.

Given that the time-lag for SWD can be months or years, prior injection history is important when modelling a specific time-period (Grigoratos et al., 2020b). Therefore, even though we perform the maximum likelihood regression for Σ, θ (and L_1, L_0) only during time-periods of low M_c, we processed all the available disposal data dating back to 2006 or 2000 when simulating (for any given value of θ) the time-dependent variable v_{lag}. In other words, the months considered in computing the total maximum likelihood of each block can be a subset of the entire $v_{lag}[t]$ time-series, allowing the user to focus on specific time-periods of interest. This workaround of Grigoratos et al. (2022) further generalizes the model, rendering it applicable even when the seismicity data have gaps in the time-domain. If, for example, the seismic network was down for a month or even years (e.g. Quinones

et al., 2019, their figure 4), we could exclude that period from the regression, while including it when constructing the v_{lag} time-series.

3.3 Diffusion of disposed volumes

Following Grigoratos et al. (2020a), we acknowledge that injection in one block will affect the pore pressure and stress field in a neighboring block, we distributed the injected volumes of each well in space and time following the Theis (1935) equation for transient, radial flow in nonleaky vertically confined aquifers of infinite areal extent. Although the Theis (1935) equation simulates fluid flow and the associated changes in pore pressure (i.e. head), we applied it to the injection volumes because the Seismogenic Index model uses injection volumes as a proxy for changes in pore pressure. Note that, here, the spatial and temporal distributing of the injection volumes is an attempt to model the effects of pore pressure diffusion, and it is distinct from the seismicity time-lag modeled by equation 4.

Given a constant rate of discharge from a well (v in units of m^3/s), the drawdown s (m) as a function of time since the beginning of pumping (t) and radial distance from the well (r) is given by (Theis, 1935):

$$s[r,t] = h_0 - h[r,t] = \frac{v}{4\pi T}\int_{\frac{r^2}{4Dt}}^{\infty}\frac{e^{-u}}{u}du \quad (5)$$

where h_0 is the initial fluid head, $h[r,t]$ is the head during pumping, T is the transmissivity (m^2/s) and D is the hydraulic diffusivity (m^2/s). In our application, v is injection rate rather than discharge rate and thus equation 6 provides the increase in head rather than the drawdown. We use equation 5 to distribute the monthly injected volumes spatiotemporally before we utilize them in equation 4. Further implementation details are provided in Grigoratos et al. (2020a).

The only parameter needed to compute the distribution factors is the diffusivity, D. Since the Theis equation describes radial flow within a single layer, we cannot model the effects of the injection layer and the basement separately. Instead, the parameter D represents the large-scale diffusivity that is often derived from seismicity migration analyses (Parotidis and Shapiro, 2004). Within Oklahoma this value ranges between 0.1 and 2 m^2/s (e.g. Goebel *et al.*, 2017). Grigoratos et al. (2020a) opted for *D* equal to 0.3 m^2/s. In this study, for each region of interest, we performed our computations for three different values of D, equal to 0.3, 1.0 and 2.0 m^2/s and commented on the sensitivity of the results.

3.3 Spatial oversampling

As far as the p-values are concerned, Grigoratos et al., (2020a) demonstrated their sensitivity to the spatial 2D grid selected to aggregate the data into blocks and how oversampling can lead to stable and grid-independent results. The model-calibration was done in a novel way across two different spatial scales that balance the need to have sufficiently large blocks to have enough data to perform the regression and the potential for small-scale spatial variations of the model parameters to be present. First, we defined a grid of base-blocks across the study area with size X by X km. Using this grid as reference, we then defined 5 different secondary grids of double sized blocks (2X by 2X km) across the study area, with their relative placement with respect to each base-block defined according to Figure 1. We first compute a p-value for each block of the 5 secondary grids. Then, the p-value of each base-block is taken as the median p-value of the 5 secondary blocks within which it resides. This procedure produces results that are essentially grid-independent. The size of the base-blocks when dealing with SWD data was 10x10km (justification within Grigoratos et al., 2020a), while for HF it was 5x5km (justification within Grigoratos et al., 2022).

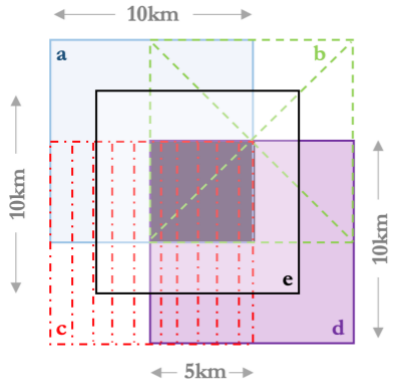

Figure 1. Schematic placement of five secondary 10x10km blocks (a-e) on top of a single 5x5km base-block (dark grey shade). Figure from Grigoratos et al. (2022).

Similar considerations regarding the set-up of the grids can be made when applying the model to forecast seismicity rates given a future injection-scenario. Grigoratos et al., (2020a) introduced a novel multi-scale regression scheme with a weighting function to address the grid-sensitivity issue, an approach that is more complex than simply taking the median estimate as we did for the p-values. Here, we adopt the same framework; details are available in Grigoratos et al., (2020a; 2022).

3.4 Hydraulic Fracturing Radar

Following Grigoratos et al. (2022), we used the Hydraulic Fracturing Radar (HFR; equation 6) as an an initial screening tool for HF within each region of interest. To that end, first we plotted a 5km base-grid with the percentage of declustered earthquakes above M_c that occurred during a HF stimulation and within 5km from it ($EQ^{\%}_{HF}$). Given the uncertainty in the end-dates of the HF stimulations coming from the IHS database, we allowed for a 3-day buffer after the assigned end-date for the temporal window. Then, we adjusted $EQ^{\%}_{HF}$ to account for the maximum total duration of stimulations within a block over the investigated study period. The reason for that adjustment is the observed asymmetry in the intensity of HF activities across blocks. Some blocks have HF stimulations every other day, while others average a few stimulation-days per year. The higher the percentage of days with ongoing stimulations, the higher the chances of coincidental correlation with non-related seismicity.

$$HFR = EQ^{\%}_{HF} \cdot (1 - \frac{t_{HF}}{t^{max}_{HF}}) \quad (6)$$

where t_{HF} is the total stimulation-days in the block and t^{max}_{HF} is the maximum total stimulation-days across all blocks (within the investigated area). This spatio-temporal association filter cannot rule out coincidental correlations, in which an earthquake triggered by SWD happened to occur during a nearby HF stimulation, possibly at different depths. On the other hand, this filter cannot capture potential delayed triggering of earthquakes several days or weeks after the HF stimulation has ended. That said, high resolution catalogs indicate that about 90% of earthquakes are triggered during HF stimulations (Schultz et al., 2018). Despite these important limitations, large HFR values can validate conclusions drawn from the p-value analyses. For example, blocks with smaller HFR values, perhaps below about 0.10, are likely not affected by HF, while low p-values should be expected in blocks with larger HFR values, above about 0.40. Potential counter-intuitive signals could be flags for manual review of the local data.

4. Results

In this chapter, we present the main product of our analysis, the median grid-independent p-values for the three types of oil and gas activity in question, that is, shallow SWD, deep SWD, and HF. The null hypothesis in all three cases was that all earthquakes are of (purely) tectonic origin. The lower the p-value, the better the correlation between the seismicity and the oil and gas activity in question. Very low p-values, that is, below 0.05, signal that the examined human activity is a key triggering factor of the seismicity in that block (Grigoratos et al., 2022). P-values between 0.05 and 0.10 might flag blocks where a subset of the clusters is induced. Blocks with low p-values in two different maps are likely affected by both types of human activity, with more sophisticated statistical analysis needed to properly quantify the exact ratio. Higher p-values above 0.20, and especially above 0.50, indicate that very little to no seismicity there was triggered by the analyzed type of human activity.

We believe that maps like these could inform either the operators or the regulators, acting as a blueprint for potential mitigation measures or well-specific investigations. We should caveat this by clarifying that the p-values we derived for each type of oil and gas activity cannot be directly compared with great decimal accuracy given that their derivation varies significantly. That is why we opted to use color-coded intervals to display the results. Even though the functional forms (equations 3–4) are very similar, the likelihood functions they are based on uses different temporal resolution (daily or monthly), varying number of temporal data points, and different physical parameters (highly pressurized flow rate vs lower pressure disposal rate). In the presence of accurate hypocentral depths, cross sections can also be illuminating (Grigoratos et al., 2022; their figure 2).

4.1 Declustering

We compiled a new unified earthquake catalog for the CEUS merging multiple different sources (Table 1). Then, the unified catalog was split into various subcatalogs to account for spatio-temporal variations in the magnitude of completeness (M_c). Subcatalogs were created for the Delaware basin, the Midland basin, the EF, the FWB, the NM-CO border, and for OK-KN. The magnitude of completeness was estimated for each subcatalog for different time-periods based on the slope of the magnitude-frequency-distribution. Next, we removed foreshocks and aftershocks using the declusterin algorithm of Aden-Antoniow et al. (2022), which utilizes the nearest-neighbor clustering algorithm of Zaliapin et al. (2008). This algorithm is robust for small samples and stable with Mc, yet sensitive to the d- and w-parameters that are defined by the user (we opted for 1.5 and 0.5 respectively, based on sensitivity tests). We applied one further modification by post-processing the clusters to always keep the largest event in each one as the mainshock (instead of the earliest one); this modification has a negligible effect on the p-values.

Figures 2 shows the results of the declustering for the entire CEUS area (west of W 117), for the time period between 2008 and 2022, for a magnitude cut off of M 2. The algorithm removed 33% of the events. In reality, we expect the M_c to be closer to 3 overall; that said neither the declustering algorithm nor our later analyses are sensitive to that parameter. For reference, the corresponding percentages (with different d- and w- parametrization however) for Southern California were 70% and for New Zealand was 43% (Aden-Antoniow et al., 2022). Therefore, we observe that the predominantly induced seismicity in CEUS has a lower percentage of aftershocks compared to typical active tectonic regions. We welcome this outcome, since it implies that the algorithm was able to identify that the oil and gas activities trigger newly formed clusters (mainshocks) and does not amalgamate the induced seismicity with the aftershock activity.

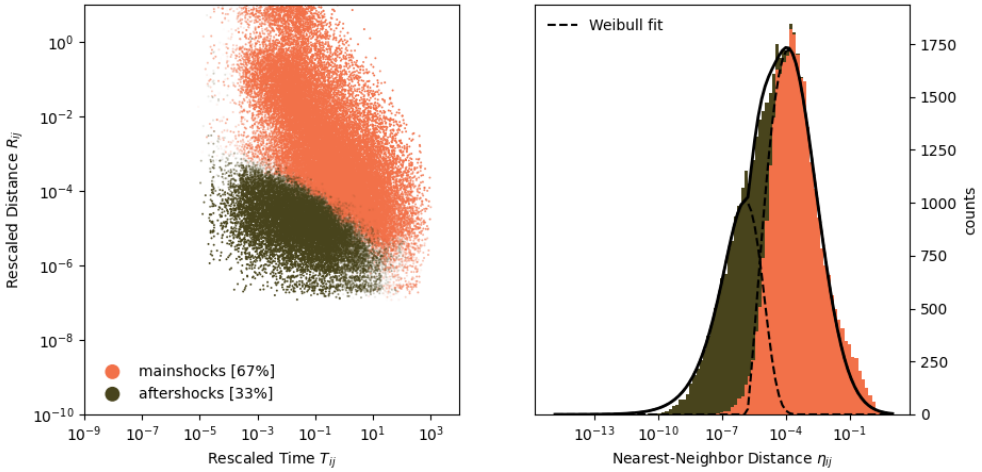

Figure 2. Declustering results for CEUS, earthquakes with M ≥ 2, between 2008 and 2022; (left) the nearest-neighbor rescaled distance R_{ij} and time T_{ij} distribution. The color indicates if the events have been classified as background events or aftershocks. (right) the two stacked nearest-neighbor distance η_{ij} distributions. The dashed black lines correspond to the fit of a Weibull function to both distributions while the black line shows the resulting sum and fit to the overall η_{ij} distribution.

Figures 3 shows the results of the declustering for OK and KN, for the time period between 2000 and 2022, for a magnitude cut off of M 2.5. Oklahoma. The algorithm removed 34% of the events, a very similar number to CEUS, as expected since OK is the dominant source of seismicity in the region since 2008. Again, in reality, we expect the M_c to be closer to 3 in early years, and reach 2.5 only after 2014 or so.

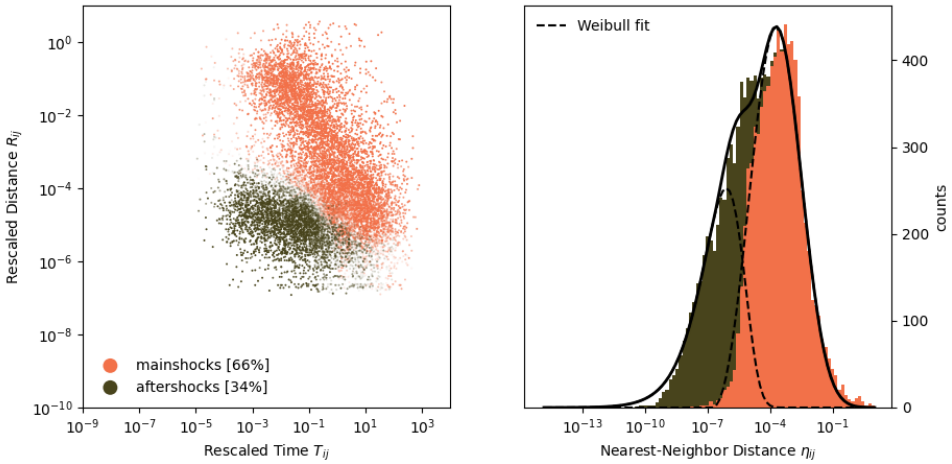

Figure 3. Declustering results for OK-KN, earthquakes with M ≥ 2.5, between 2000 and 2022. Same figure-structure as Figure 2.

Figures 4 shows the results of the declustering for the Delaware basin (west TX and south NM), for the time period between 2017 and 2022, for a magnitude cut off of M 1.5. We believe that the M_c is low in that region due to the template-matched data for 2017 and 2018, the extra events from the TexNet Research catalog and the densified network after 2020. That said, it was likely closer to 2 before 2019. The algorithm removed only 14% of the events, indicating a lack of earthquake-to-earthquake interactions and close proximity to a direct triggering source (the injecting wells). This hints at HF and perhaps shallow SWD being a key causal factor. Oklahoma, an area dominated by deep SWD had more than double the percentage of aftershocks (even for a higher magnitude cut off).

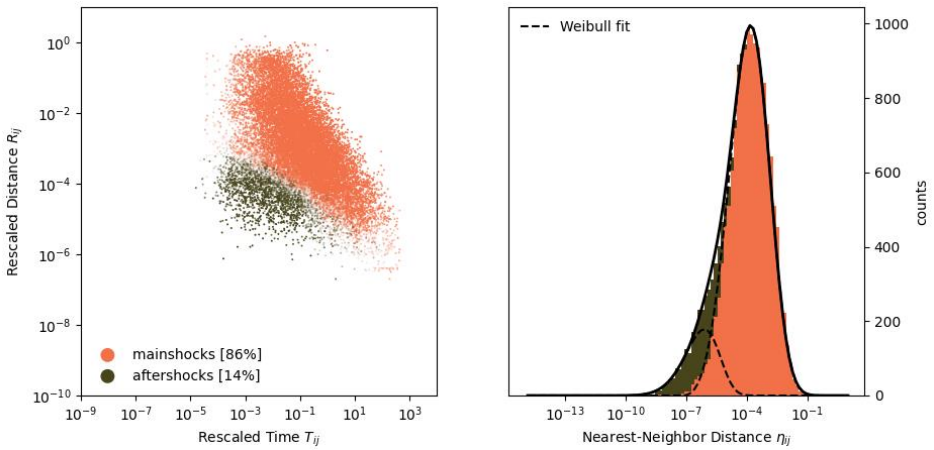

Figure 4. Declustering results for the Delaware basin, earthquakes with M ≥ 1.5, between 2017 and 2022. Same figure-structure as Figure 2.

Figures 5 shows the results of the declustering for the Midland basin (including the Midland basin), for the time period between 2017 and 2022, for a magnitude cut off of M 2. We believe that the M_c is relatively low in that region due to the extra events from the TexNet Research catalog and the densified network after 2020. That said, it was likely closer to 2.5 before 2019. The algorithm removed only 19% of the events, a similar number to the Delaware basin (likely for the same reasons). This would indicate that deep SWD is not that dominant of a causal factor in the Midland basin.

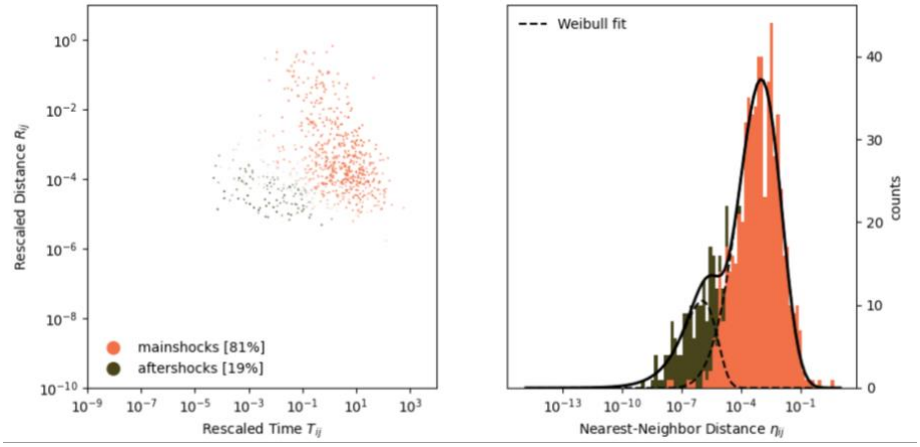

Figure 5. Declustering results for the Midland Basin, earthquakes with M ≥ 2, between 2017 and 2022. Same figure-structure as Figure 2.

We do not show declustering results for EF and the CO-NM border (although we did compute them) because the earthquake count is too low.

4.2 Eagle Ford

The Eagle Ford shale play (EF) in south-central Texas has been an extensive oil and gas production area since the 1950s and has seen extensive HF activities since around 2008 (Frohlich & Brunt, 2013). Its geologic setting is described in McKeighan et al. (2022), with recent focal mechanisms being computed by Li et al. (2021). Faulting appears dominated by normal faults.

In terms of local seismic monitoring, from November 2009 to September 2011, the EarthScope USArray (TA) deployed 25 temporary stations in and around the EF (Frohlich and Brunt, 2013). Beginning in 2016, the Texas Seismological Network (TexNet; Savvaidis et al., 2019) deployed eight broadband seismic stations in the EF. In 2018, TexNet and The University of Texas at Austin Institute for Geophysics (UTIG) deployed 23 temporary broadband seismic stations at about a 25 km interstation distance (Li et al., 2021).

The available catalogs for the EF are listed in Table 1, with TexNet being the key resource. Frohlich & Brunt (2013) produced a catalog from the TA data. Fasola et al. (2019) did template-matching based on TexNet events,

for March 2014 through December 2018. Fasola et al. (2019) and Fasola & Brudzinski (2023) did template-matching based on TexNet events for the periods of March 2014 - December 2018, and January 2019 - February 2020, respectively. Both studies assigned the parent-location to newly detected events. Fasola & Brudzinski (2023) also updated the magnitudes of Fasola et al. (2019), following Kavoura et al. (2020) and employed a repeater detector between January 2015 and February 2020 for a single-station. The events from the repeater detector have very approximate locations and thus were not included in our unified catalog. Finally, we did not utilize the relocation of Li and Savvaidis (2021) for TexNet events before 2020.

The earthquake M_c for the EF varies in time and space. Fasola et al. (2019) cited an M_c of M_L 2.3 for their catalog. Li et al. (2021) cited an M_c of M_L 2.0 for TexNet from 2017 to 2019, although this value seems optimistic to us, especially for 2017. Fasola & Brudzinski (2023) cited an M_c of M_L 0.7 for their template-matched catalog.

The EF was considered seismically dormant before the introduction of oil and gas production methods (McKeighan et al., 2022). In 1973, before the installation of a seismic network, tremors and shaking were felt in Fashing, Texas, Atascosa County (Frohlich et al., 2016). In 1993 and 2011, two earthquakes occurred in Fashing with magnitudes of 4.3 and 4.8, respectively (Frohlich et al., 2016). The latter is thought to be induced by fluid extraction (Frohlich & Brunt, 2013). In May 2018, a Mw 4.0 earthquake occurred less than 3.5 km away from two active well laterals (Fasola et al., 2019), rendering it the largest US earthquake to be linked to HF. It occurred ~10 km from the 2011 Mw 4.8 earthquake in 2011. Notably, in 2018, the rate of M ≥ 3.0 earthquakes in the EF grew to 33 times higher than background levels, with 0.3 events per year from 1980 to 2010 (Fasola et al., 2019).

Wastewater disposal wells are also present in the EF, although HF ones are more common (Figure 6). Seismicity is scattered throughout the play and is also present across the border with Mexico to the southwest. The monthly aggregate HF rates across the play peaked in 2012 and have been decreasing ever since, with a temporary upward trend between 2016 and 2019 (Figure 7). The SWD rates had been steadily increasing between 2000 and 2014, and have since been flat to downward. The declustered seismicity rates seems to have peaked in 2018 and 2019 with a secondary spike in 2022.

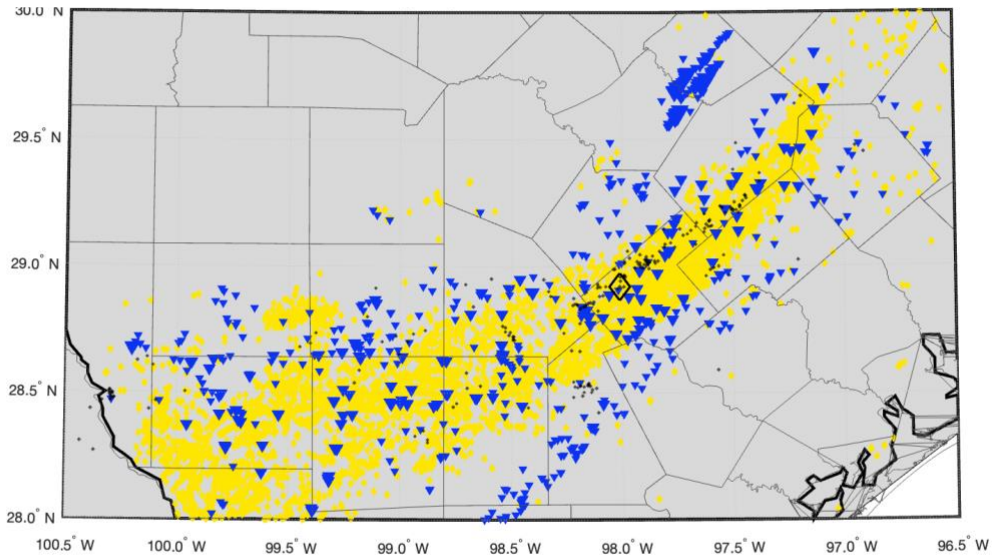

Figure 6. Map of the given investigated area showing declustered seismicity (M ≥ 2; black dots), wells (blue: SWD; yellow: HF) and county borders. The black diamond indicates the 2018 M_w 4.0 earthquake. Data only for the EF play, between 2014 and 2021.

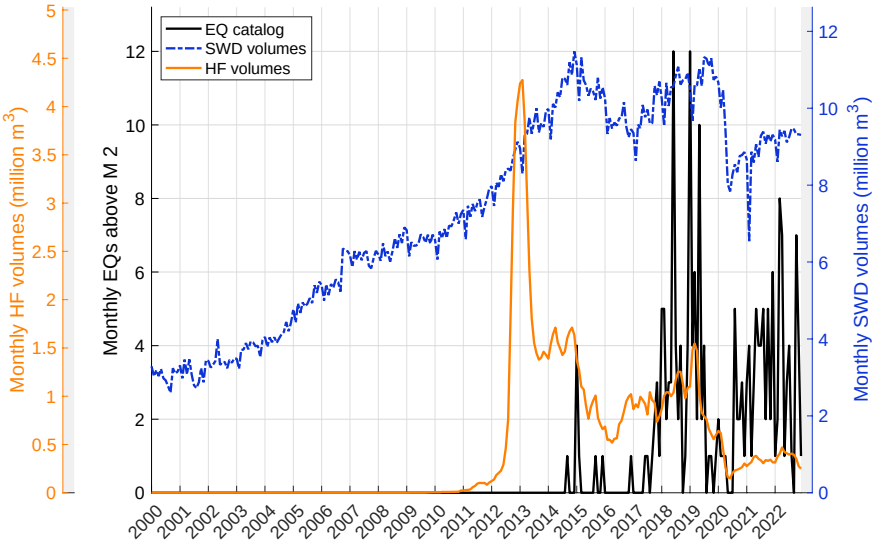

Figure 7. Time history of monthly HF volumes, SWD volumes, and declustered seismicity rates above M 2, for the given investigated area. The earthquake catalog is largely incomplete before 2017, while the HF data are incomplete before 2012.

The b-value was also estimated for illustration purposes only, based on maximum likelihood statistics (assuming a Poisson process), and by bootstrapping 1000 catalog-samples to account for measurement uncertainties behind the cataloged magnitudes (assumed to have a normally distributed standard deviation of 0.2). The magnitude binning interval was 0.1. The b-value was computed based on the regression by Weichert (1980). For the 2017 to September 2022 period, it was around 1.0, but this value is misleading due to the extremely poor fit at magnitudes larger than 3 (Figure 8).

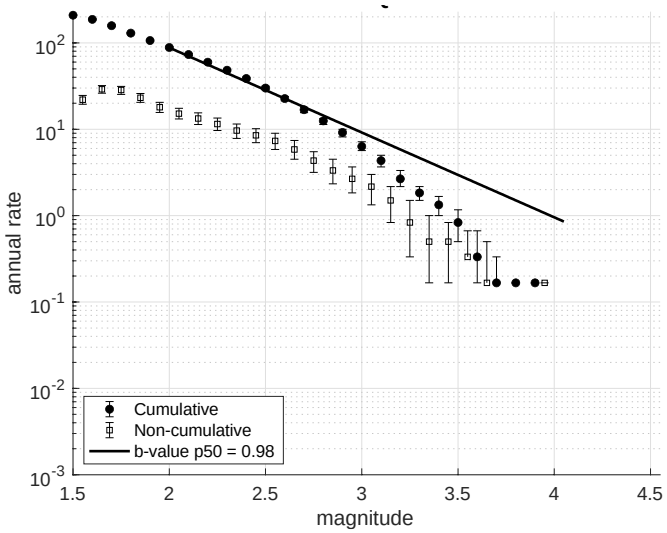

Figure 8. Bootstrapped (non-declustered) magnitude-frequency-distribution for the given investigated area, along with its fitted (median) b-value. Data between 2017 and September 2022.

4.2.1 HF

In this section we will analyze the potential link between the observed seismicity in the EF and HF. The gridded spatial distribution of total HF volumes and the earthquake epicenters is mapped in Figure 9.

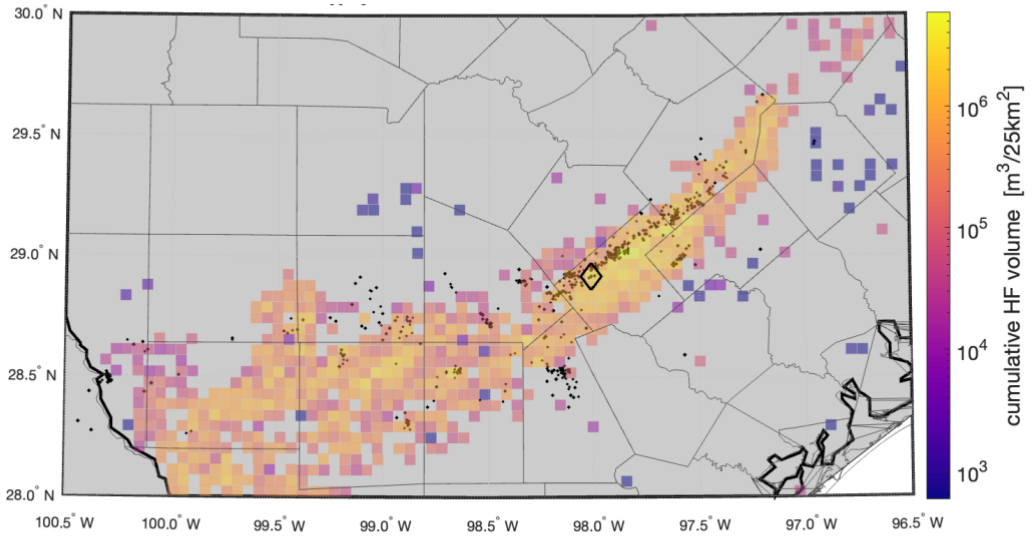

Figure 9. Map of the given investigated area showing declustered seismicity (M ≥ 2; black dots), gridded cumulative HF volumes and county borders. The black diamond indicates the 2018 M_w 4.0 earthquake. Data only for the EF play, between 2017 and September 2022.

Fasola et al. (2019) classified earthquakes from their catalog (2014–2018) within 5 to 10km and 7 days of a stimulation as induced by HF. A minimum of 5 earthquakes had to be linked to a stimulation for the classification to stand. The latter seems like an odd criterion, especially given that they did not filter for M_c. In any case, they found that 87% of their template-matched catalog, including 94 events with M_L 2.0 or above, were induced. They also evaluated the influence of operation strategy, target formation, injected volume, and number of wells per pad on the likelihood of seismicity using statistical modeling. Their key finding was that wells with multiple pads are more likely to cause seismicity. Only 8% of the 211 HF wells they examined were associated with induced seismicity, according to their set of criteria.

Fasola & Brudzinski (2023) performed the same correlation study using their catalog (2019 to February 2020). They found 96 events with $M_L \geq 2.0$ and 10 with $M_L \geq 3.0$ being flagged as induced by HF. They also confirmed that newly detected HF-induced seismicity in Live Oak county did not occur until January 2019. Notably they concluded that the total amount of fluids injected per month has a stronger influence than the total injected volume.

Figure 10 shows HFR values for the region, for the period between 2017 and September 2022, and for declustered earthquakes with magnitudes above 2. The majority of blocks have high values (above 0.25), as expected from the literature review. According to our own association-criteria (section 3.4), 55% of declustered earthquakes are associated with only 5% of the HF stimulations. Direct comparisons with the studies from the literature are challenging, because we used tighter spatio-temporal windows, a declustered catalog and accounted for the M_c. Finally, 93% of HF stimulations that eventually induced seismicity according to our criteria, started doing so during the stimulation period. This is yet another indication that the long multi-day or even multi-week time-windows often used in the literature are incorrect.

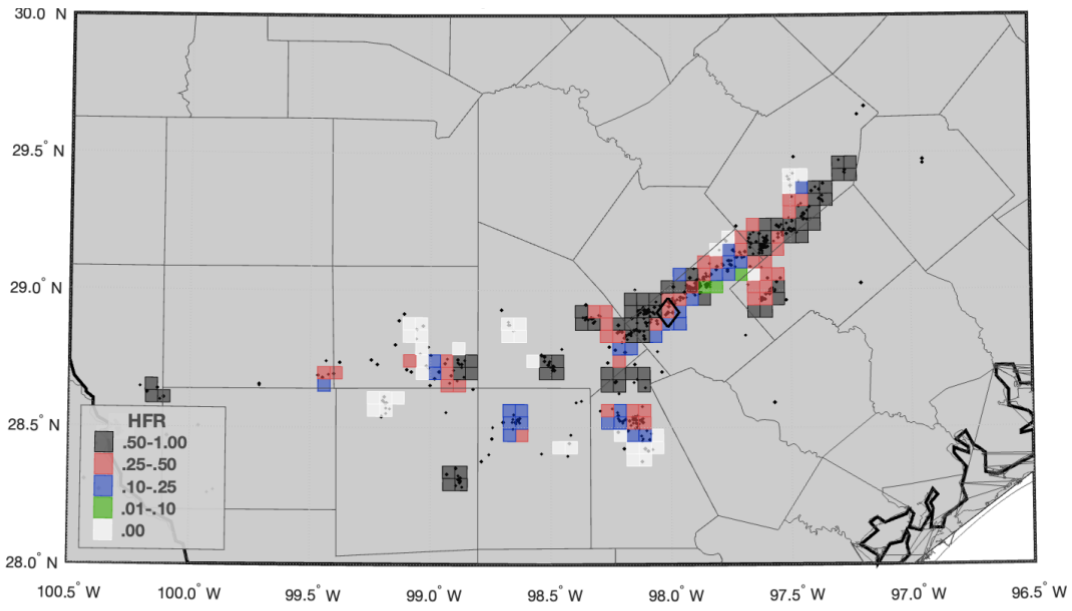

Figure 10. HFR values mapped on a 5 km grid, with all declustered earthquakes between 2017 and September 2022 above M 2 overlapped. The black diamond indicates the 2018 M_w 4.0 earthquake. Only blocks with at least 3 events are color-coded. Country borders are also mapped.

Figure 11 shows the statistical p-values for HF (p_{HF}) in the region, for the period between 2017 and September 2022. They paint the same picture as the HFR values. In particular, 66% of blocks have $p_{HF} \leq 0.05$, and 77% of blocks have $p_{HF} \leq 0.10$. Furthermore, 47% of earthquakes occurred within a block with $p_{HF} \leq 0.05$. When we combine SAF with p_{HF}, 38% of declustered earthquakes above M 2 are both linked to HF and occurred within a block with $p_{HF} \leq 0.05$. These earthquakes are almost certainly induced by HF. As far as the HF wells are concerned, only 3% of stimulations are both linked to seismicity and occurred within a block with $p_{HF} \leq 0.05$. On the other hand, 87% of stimulations are not linked to seismicity and occurred within a block with $p_{HF} \geq 0.05$. Lastly, 83% of blocks have a fitted time-lag (equation 4) of 3 days or less and the median Σ value among blocks with $p_{HF} \leq 0.05$ is -3.4.

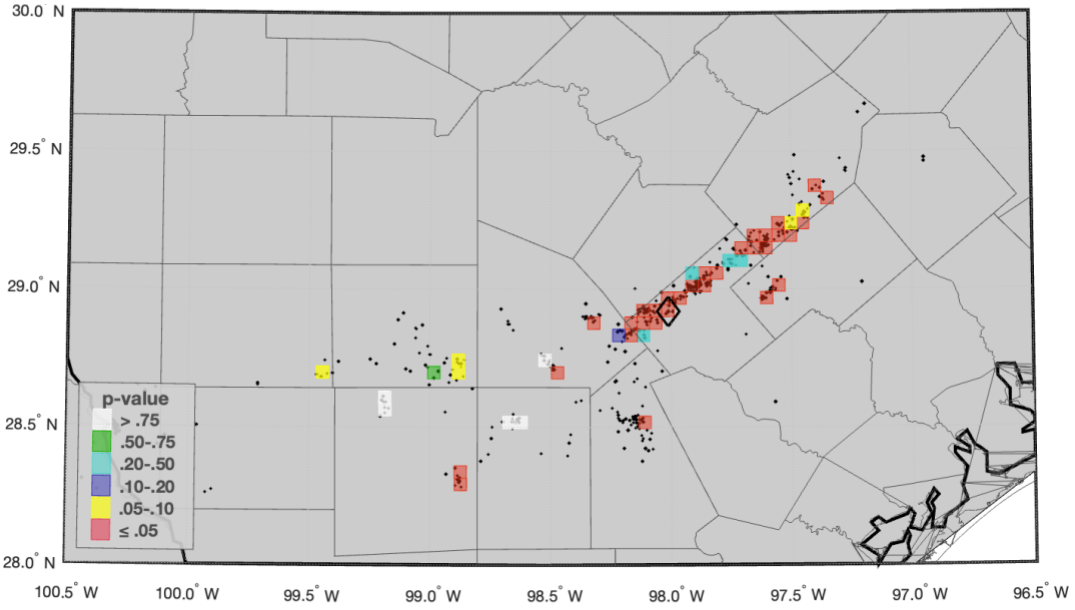

Figure 11. Statistical p-values for HF between 2017 and September 2022, with all declustered earthquakes above M 2 overlapped. Only blocks with at least 3 events are color-coded. The black diamond indicates the 2018 M_w 4.0 earthquake. Country borders are also mapped.

Figure 12 is identical to Figure 11, with only one distinction; the analysis used a non-declustered earthquake catalog. The effect on the results is negligible, highlighting the stability of our methodology.

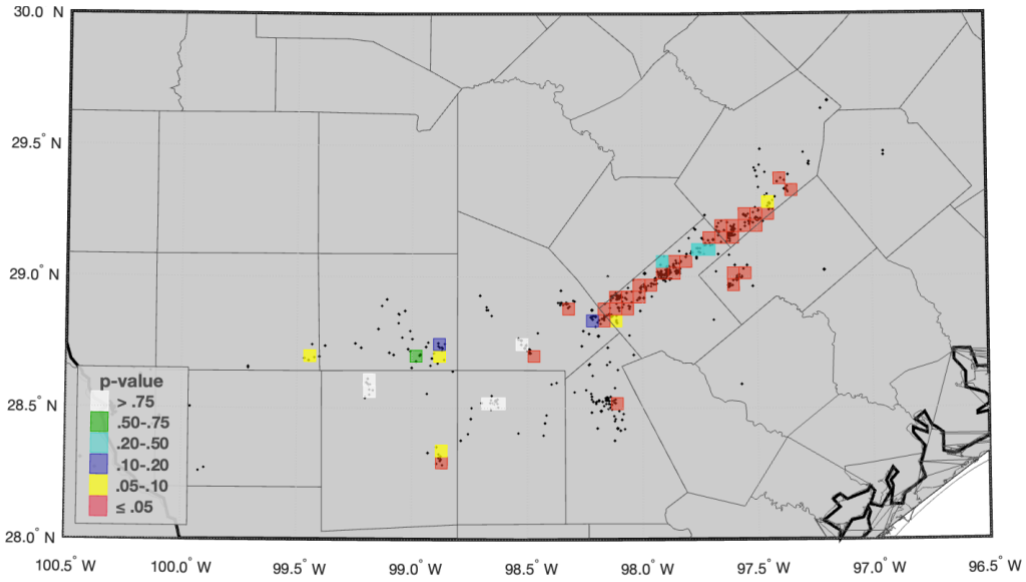

Figure 12. Same as Figure 11, but using the non-declustered catalog.

Figure 13 is identical to Figure 11, with only one distinction; the analysis used the contract time-lag t_{lag}^{c} instead of the time-dependent one (section 3.2 for details). The effect on the results is negligible, highlighting the stability of our methodology.

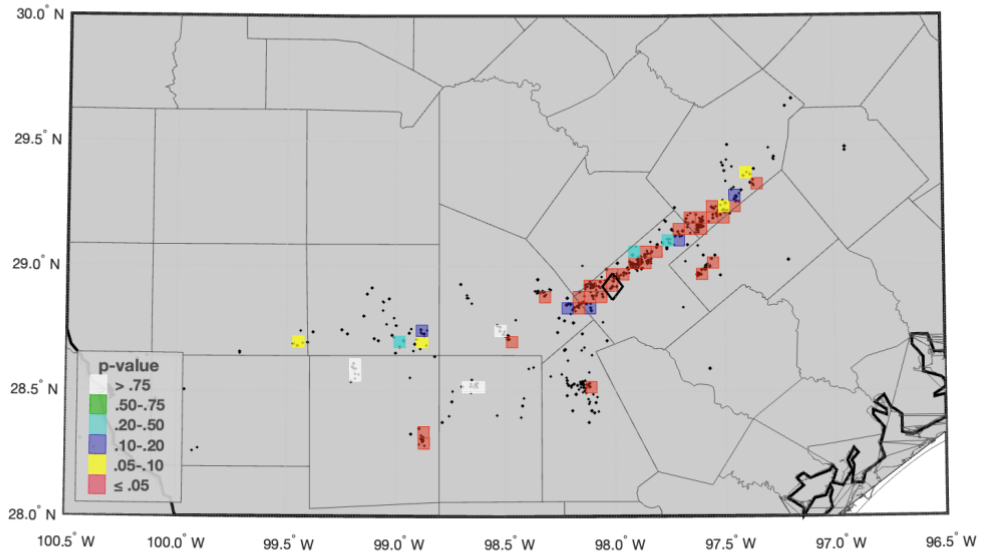

Figure 13. Same as Figure 11, but using the contract time-lag t_{lag}^{c} instead of time-dependent one (equation 4).

Figure 14 is identical to Figure 11, with only one distinction; the analysis was conducted over a much shorter time-interval; the time-period was 2017 to September 2018, instead of the default option of 2017 to September 2022. The results indicate that our method can provide early warnings, even with limited earthquake data available, before the seismicity has expanded throughout the oil and gas field. All the induced clusters flagged in Figure 14 were later also flagged in Figure 11. The sensitivity tests presented here for EF are in agreement with similar tests conducted by Grigoratos et al. (2022) or the Delaware basin.

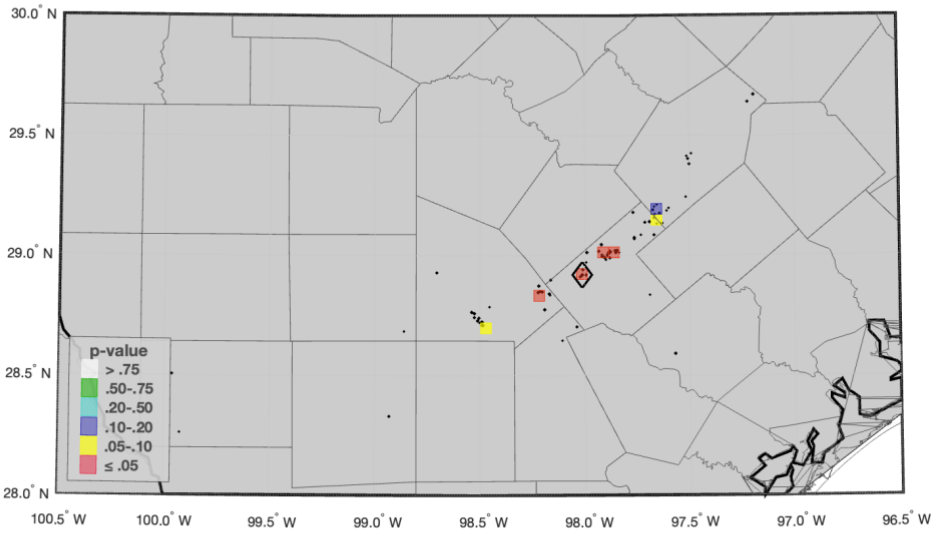

Figure 14. Same as Figure 11, but for the time-period 2017 to September 2018.

Our generalized earthquake recurrence model can also be used to hindcast or forecast seismicity rates, given an injection scenario specific to the type of oil and gas activity in question. The hindcasting performance in the EF (Figure 15) was very good and better than the simple correlation with the injection rates would imply. The simulation following the observed seismicity rates closely and even signaling the higher real M_c in 2017, by overestimating the observed seismicity rates during that year.

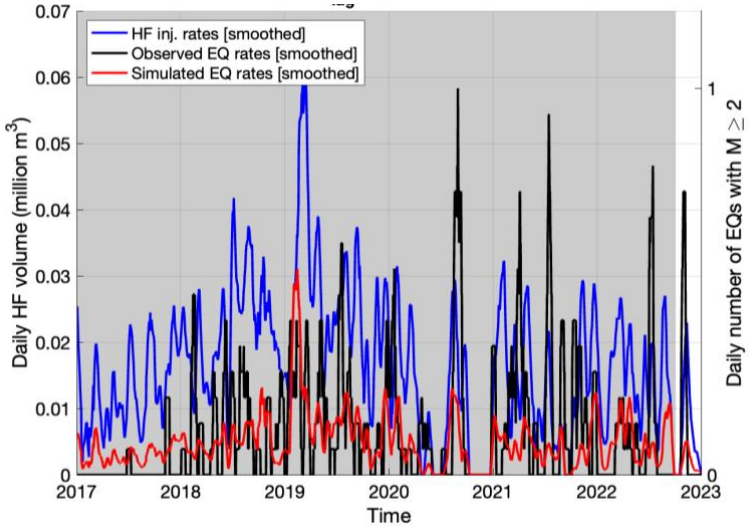

Figure 15. Two-week moving average of daily HF injection and declustered seismicity rates. The calibration period for the regression was 2017 to September 2018 (gray background). Only aggregated data from blocks with $p_{HF} \leq 0.05$ and at least three declustered earthquakes are included.

4.2.2 SWD

Contrary to HF, SWD has received little to no attention when it comes to the EF area, despite the fact that SWD volumes are clearly overlapping in space with most of the earthquake activity (Figure 16). Even in the time domain, Figure 7 shows that the SWD rates are arguably better correlated with the seismicity rates after 2014 than HF rates are.

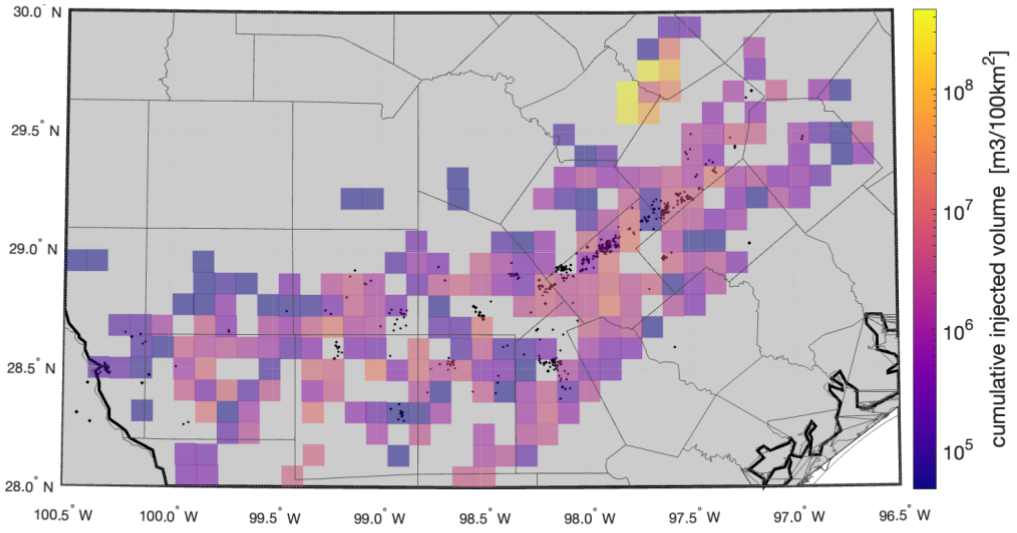

Figure 16. Map of the investigated area showing non-declustered seismicity (M ≥ 2; black dots), gridded cumulative SWD volumes and county borders. Data only for the EF play, between 2000 and 2021. The earthquake data are largely incomplete before 2014, and partly incomplete between 2014 and 2017.

Figure 17 shows the statistical p-values for SWD (p_{SWD}) in the region, for the period between 2017 and 2021, for M ≥ 2 and the large-scale diffusivity value D set to 1 m²/s. In particular, 9% of blocks have $p_{SWD} \leq 0.05$, and 17% of blocks have $p_{SWD} \leq 0.10$. Furthermore, 20% of earthquakes occurred within a block with $p_{SWD} \leq 0.05$ and 28% within a block with $p_{SWD} \leq 0.10$. Therefore, although such correlations are absent from the literature, SWD seems linked at high confidence levels with a significant number of clusters, including the 2018 M_w 4.0 sequence, which is notably also very well correlated with HF. Furthermore, SWD appears to be the sole causal factor behind the seismicity the border with Mexico to the west end of the oil and gas field. This is a novel finding, not previously reported in the literature.

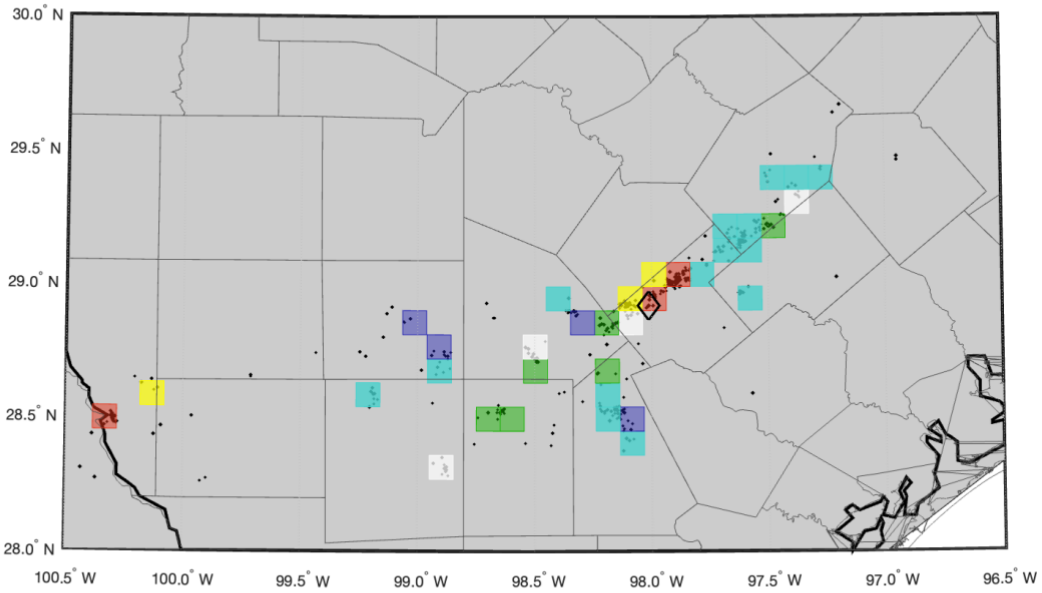

Figure 17. Statistical p-values for SWD between 2017 and 2021, for D equal to 1 m²/s, with all declustered earthquakes above M 2 overlapped. Only blocks with at least 3 events are color-coded. The black diamond indicates the 2018 M_w 4.0 earthquake. Country borders are also mapped. The color-coding scheme is the same as in Figure 11.

Figure 18 is identical to Figure 18, with only one distinction; the analysis used a non-declustered earthquake catalog. The effect on the results is minimal, highlighting the stability of our methodology. Other sensitivity tests

(not shown) also indicate the overall stability of the results with variations in the D value (e.g. equal to 0.3 or 2.0 m^2/s).

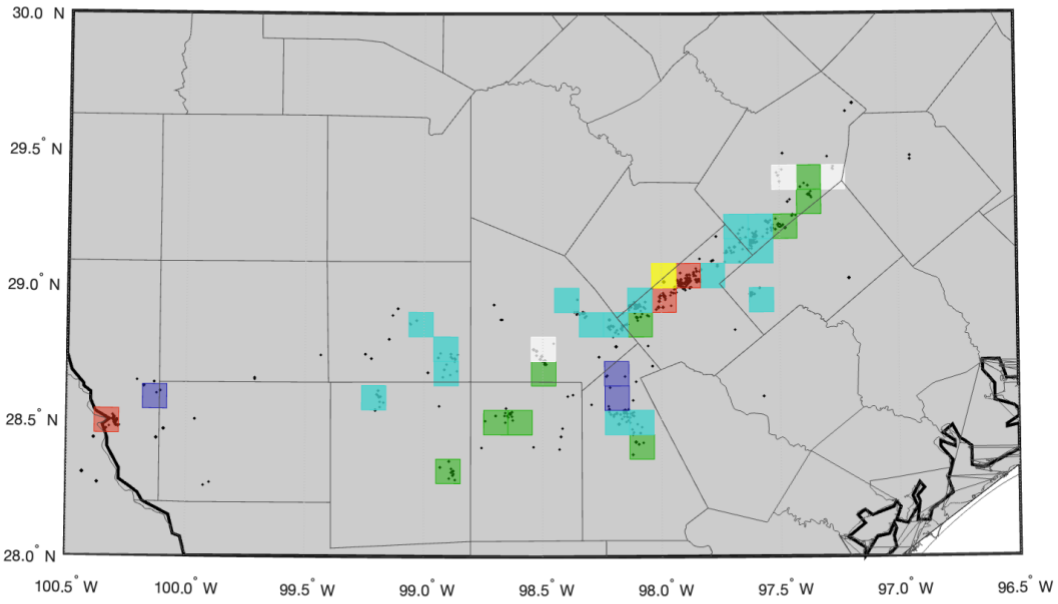

Figure 18. Same as Figure 17, but using the non-declustered catalog.

4.3 New Mexico – Colorado border

The New Mexico – Colorado border hosts two natural gas producing basins the San Juan basin (west side) and the Raton basin (east side of the border). Both basins are coalbed methane fields with production ramping up in the 1990s. Interestingly, only the Raton basin has been seismically active in recent decades, despite both hosting both HF and SWD activities (Figure 19).

SWD aiding coalbed methane production in the Raton Basin began in 1994 in Colorado expanding into New Mexico in 1999 and is primarily in the Dakota formation, a buff, conglomeratic sandstone (Johnson, 1969), with injection intervals ranging between 1250 and 2100 m below the ground surface, depending on location in the Raton Basin. There are no disposal wells in the northwestern portion of the basin because the produced formation water there meets water-quality standards for surface discharge (Rubinstein et al., 2014). Because SWD data prior to June 2006 are unavailable for New Mexico, Rubinstein et al. (2014) used the produced water volume as a proxy for the disposal volume, for prior basin-wide annual estimates. The basin-wide SWD rates increased rapidly from 2000 to 2004, then plateaued through 2011 at ~2.8 million bbls/month, and they have been gradually declining since (Glasgow et al., 2021; their Figure 1b). Reported HF injection data are available after 2011, but the exact timing of the onset of HF activities is unknown.

Historically, seismicity in the basin was rare despite an estimated M_c of 5 since 1877 (Rubinstein et al., 2014). The seismic activity took a rapid turn in 2001 however. Despite only 1 $M \geq 4$ earthquake from 1970 to 2000, 15 $M \geq 4$ earthquakes occurred from 2001 to 2020 (Glasgow et al. 2021). The peak in seismicity rates was in 2011 (Figure 20), with 20 $M \geq 3$ earthquakes in one year (Glasgow et al., 2021). In August of that same year the basin hosted its largest earthquake ever, a M_w 5.3 event. Two other important earthquake sequences have occurred since 2001, the August–September 2001 sequence (with a m_{bLg} 4.0 and a m_{bLg} 4.5 event), and the August–September 2005 sequence that included a M_w 4.6 foreshock, the M_w 5.0 mainshock and a M_w 4.0 aftershock. Notably, Glasgow et al. (2023) concluded a cascading rupture style for the 2011 mainshock, with its aftershock sequence terminating at the edge of the 2001 cluster.

There are limited and discontinuous publicly available seismic data near the Raton Basin, apart from the temporary aftershock deployments of the USGS in 2001 and 2011 (Meremonte et al., 2002; Rubinstein et al., 2014), the EarthScope USArray during which two stations bounded the basin from 2008 to 2010 (Nakai et al., 2017a; Nakai et al., 2017b), an eight-station broadband array with ~30 km inter-station spacing from July 2016 to July 2020 (Glasgow et al., 2021), and a temporary dense 96-geophone array in the southern portion of the basin from May to June 2018 (Glasgow et al., 2021). Rubinstein et al. (2014) estimated a magnitude of completeness of 3.8 for the Raton Basin from 1970 to 2011 and an M_c of around 3.0 after 2001. The latter point is supported by our own

analysis (Figure 27). Nakai et al. (2017a) cites an M_c of 1.3 for the EarthScope USArray catalog, while Glasgow et al. (2021) estimates an M_c of 0.6 for July 2016 to July 2020.

Overall, there is strong qualitative spatio-temporal evidence that the earthquakes in the Raton Basin since the year 2000 are triggered by SWD (Rubinstein et al., 2014; Glasgow et al., 2021), however more quantitative approaches (Nakai et al., 2017b) have been proven less convincing in our opinion. The 2001 sequence is effectively centered below the Wild Boar active SWD well with many epicenters within hundreds of meters of the well and most hypocenters between 1 and 3 km below the injection depth (Rubinstein et al., 2014). Wild Boar also began injecting in August 2000, one year before felt earthquakes began. At the time of the 2001 sequence, there were 4 other active SWD wells within 10 km of the earthquake sequence, with 2 wells in total injecting more than 100,000 bbls/month that year (Rubinstein et al., 2014). Meremonte et al. (2002) stated that if the 2001 earthquake sequence was natural, they would expect the seismicity to tail off and return to the lower background seismicity rate; that never happened. The 2005 M_w 5.0 event also appears in close proximity to SWD wells. Furthermore, the 2011 earthquake sequence lies within 10 km of five active SWD wells in the Raton Basin: with 2 of them injecting more than 100,000 bbls/month that year (Rubinstein et al., 2014). More recently, Glasgow et al. (2021; their Figure 11) found that greater than 90% of earthquakes occur less than 10 km from the nearest wastewater disposal well from July 2016 to July 2020, with an average earthquake to well distance of 5.8 km. On the other hand, there are some indications that the seismicity in the Raton basin might be tectonic. Glasgow et al. (2021) found that unlike some regions of induced seismicity with distinct independent background events resulting from direct forcing by fluid injection, the spatiotemporal-magnitude behavior of the seismicity in the Raton Basin is more strongly clustered, similar to tectonic regions. Their relocations also show a depth distribution that is barely reaching the top of the basement, with most of the events being located more than 4 km below the top of basement. Finally, Rubinstein et al. (2014) did some simple stress calculations to note that the production of gas or water could not be directly related to the earthquakes in the Raton Basin.

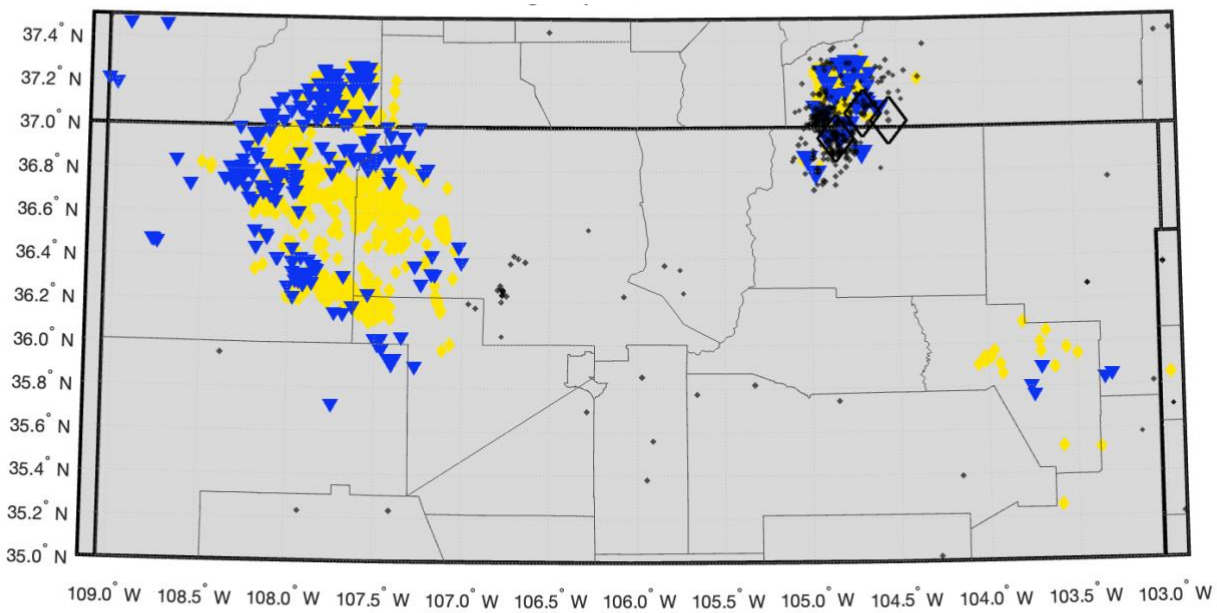

Figure 19. Map of the given investigated area showing non-declustered seismicity ($M \geq 2.5$; black dots), wells (blue: SWD; yellow: HF) and county borders. Data between 2005 and 2021. The black diamonds indicate the largest earthquakes.

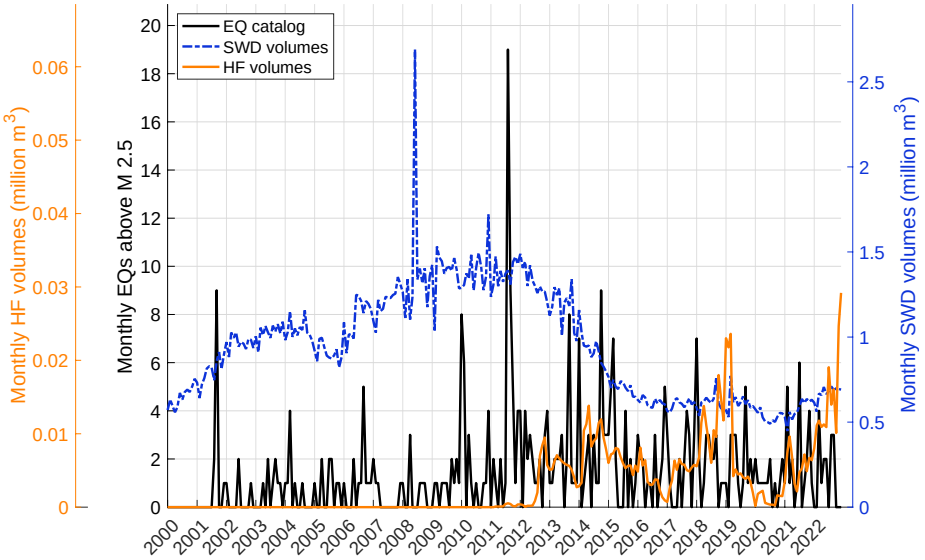

Figure 20. Time history of monthly HF & SWD volumes, and non-declustered seismicity rates above M 2.5, for the given investigated area. The HF data are incomplete before 2012.

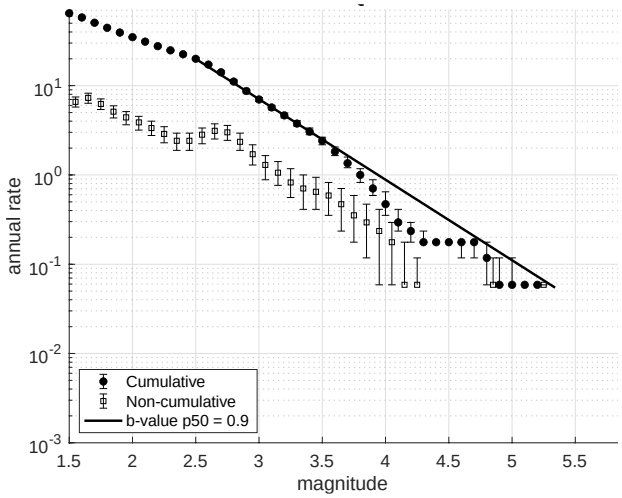

Figure 21. Bootstrapped (non-declustered) magnitude-frequency-distribution for the given investigated area, along with its fitted (median) b-value. Data between 2005 and 2021. The catalog completeness varies greatly over time and space.

4.3.1 HF

The San Juan basin hosted HF activity throughout its area, with larger total volumes in the south, while the Raton basin had HF stimulations mostly in its western side (Figure 22). It is notable that all this HF activity in the San Juan basin does not coincide with felt seismicity. That said, the seismicity in the Raton basin shows no correlation to the HF activities there (Figure 23; 24), so neither basin overall seems to have earthquakes triggered by HF. Notably, this is the first time that HF is quantitatively examined as a causal factor for the seismicity in the Raton basin.

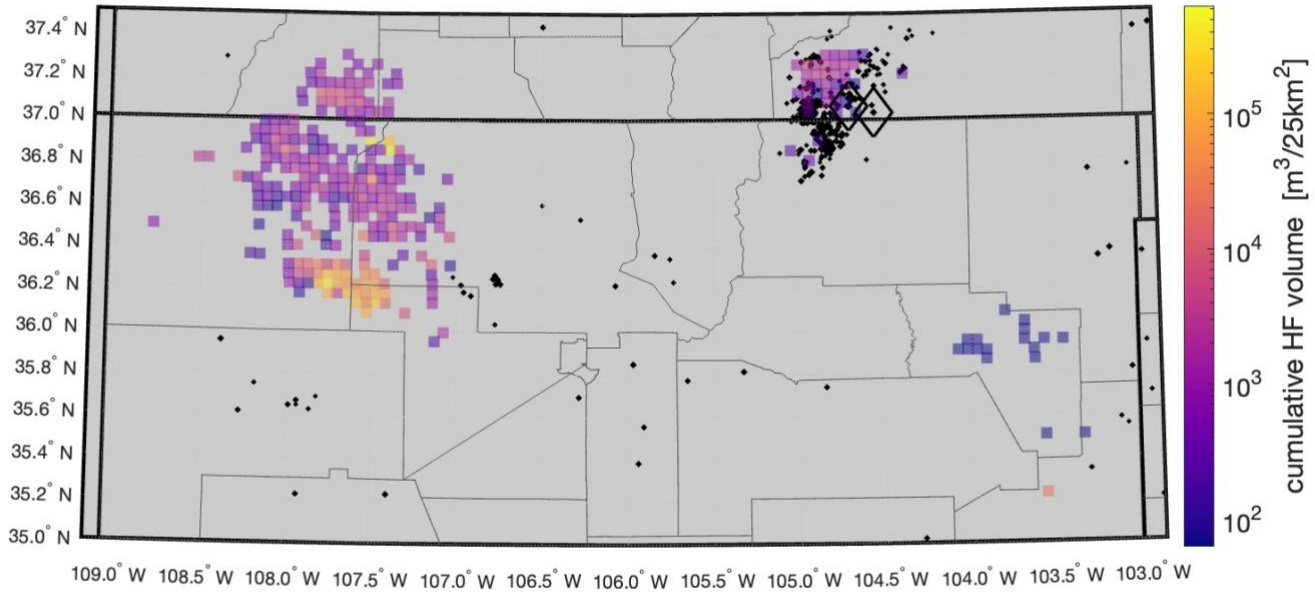

Figure 22. Map of the given investigated area showing declustered seismicity (M ≥ 2; black dots), gridded cumulative HF volumes and county borders. Data between 2011 and September 2022. The black diamonds indicate the largest earthquakes.

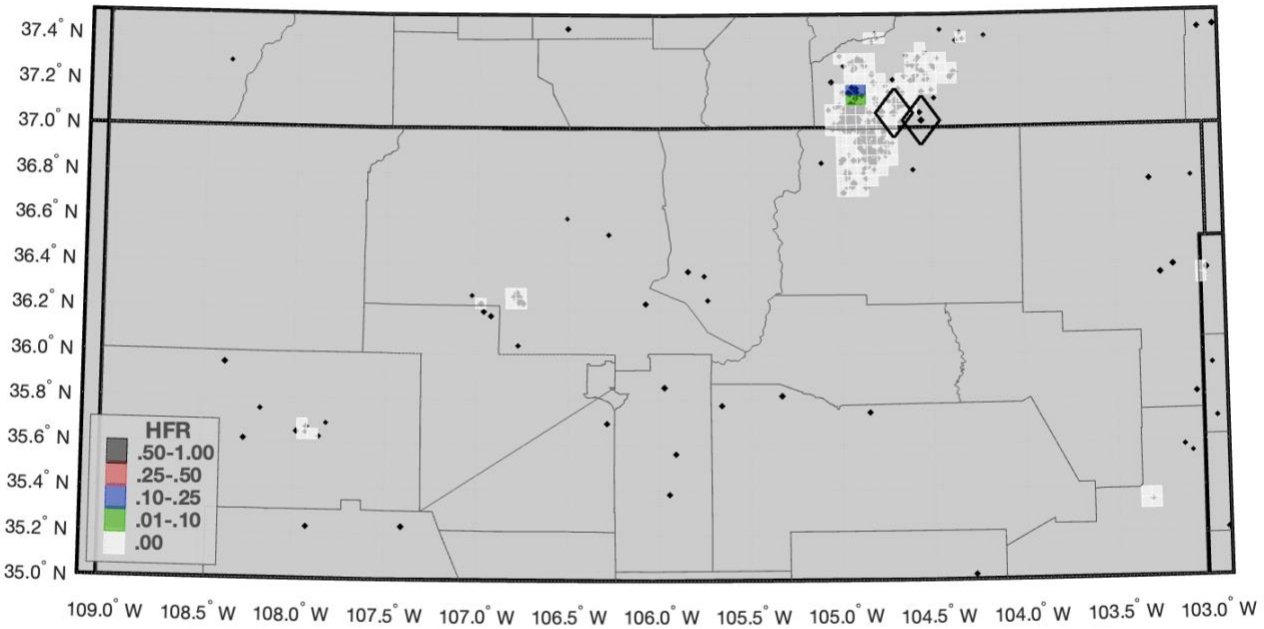

Figure 23. HFR values mapped on a 5 km grid, with all declustered earthquakes between 2011 and September 2022 above M 2 overlapped. The black diamonds indicate the largest earthquakes. Only blocks with at least 3 events are color-coded. Country borders are also mapped.

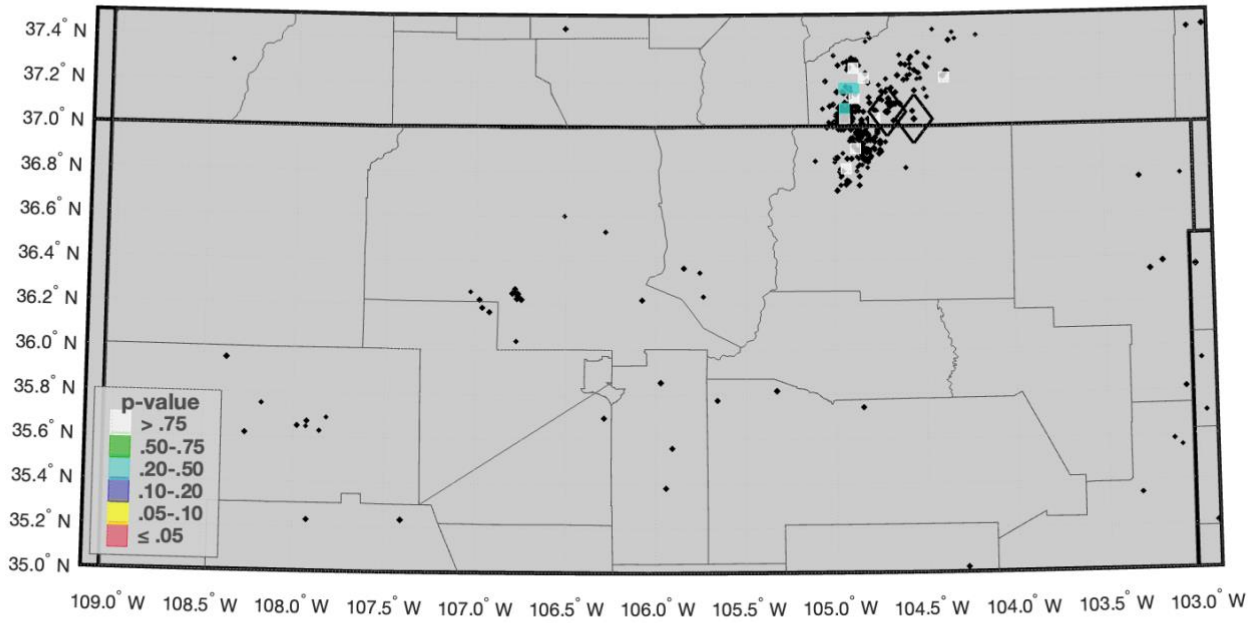

Figure 24. Statistical p-values for HF between 2011 and September 2022, with all declustered earthquakes above M 2 overlapped. Only blocks with at least 3 events are color-coded. The black diamonds indicate the largest earthquakes. Country borders are also mapped.

4.3.2 SWD

SWD has been identified as the sole causal factor behind the seismicity in the Raton basin (Rubinstein et al. 2014). Figure 25 shows that similar levels of SWD activity in the San Juan basin however do not coincide with any seismicity.

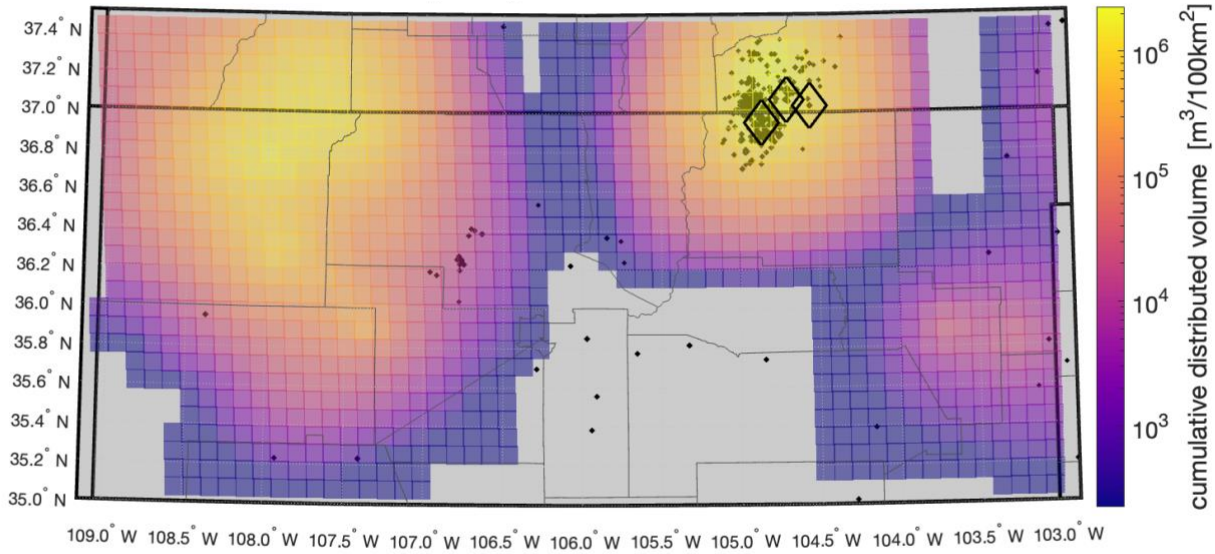

Figure 25. Map of the investigated area showing non-declustered seismicity (M ≥ 2.5; black dots), gridded diffused cumulative SWD volumes (D=1m²/s) and county borders. The black diamonds indicate the largest earthquakes. Data between 2000 and 2021. The CO SWD data within the San Juan basin are incomplete.

Figure 26 shows the statistical p-values for SWD (p_{SWD}) in the region, for the period between 2005 and 2021, for non-declustered earthquakes with M ≥ 2.5 and for D set to 1 m²/s. Most of the seismicity occurs in blocks with high p-values (low confidence for association), with only two blocks have p_{SWD} below 0.05. The results do not change much if we use a declustered catalog (Figure 27), with the confidence level decreasing further if we use D equal to 0.3 m²/s (Figure 28). Overall, neither SWD, not HF, appear to be able to explain the rise in seismicity in

the last two decades. Perhaps more consistent catalogs could benefit our analyses. In the future we plan on using the catalog of Glasgow et al. (2023), which was just published. That said, there is always the chance that these earthquakes are of mainly tectonic origin with slow progressing earthquake-to-earthquake interactions keeping the activity elevated over decades. This does seem unlikely given the evidence presented by Glasgow et al. (2021) and Rubinstein et al. (2014), but not impossible.

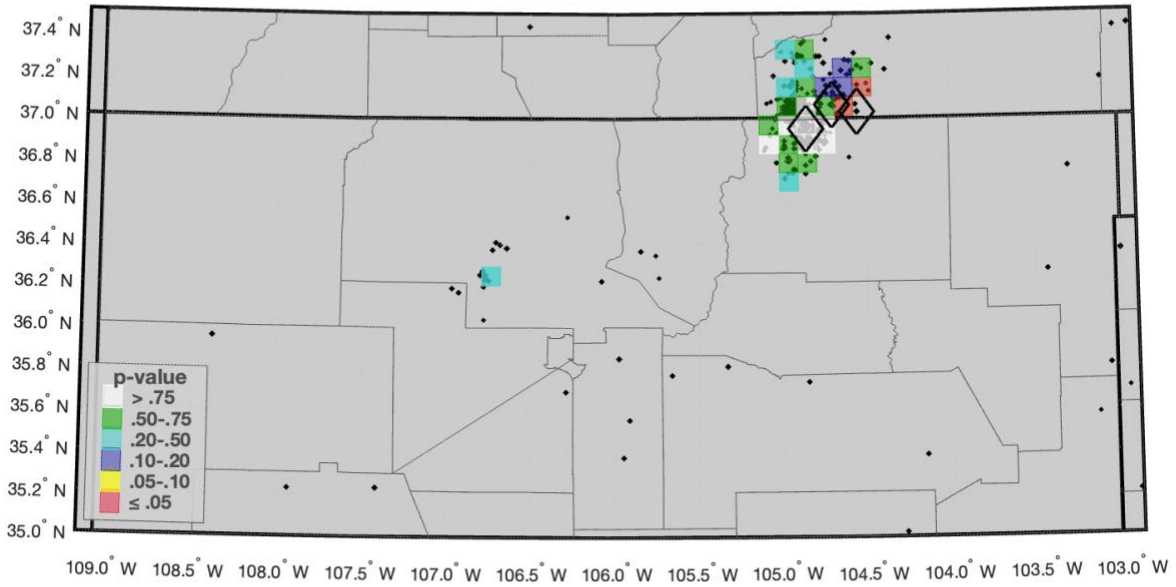

Figure 26. Statistical p-values for SWD between 2005 and 2021, for D equal to 1 m^2/s, with non-declustered earthquakes above M 2.5 overlapped. Only blocks with at least 3 events are color-coded. The black diamonds indicate the largest earthquakes. Country borders are also mapped.

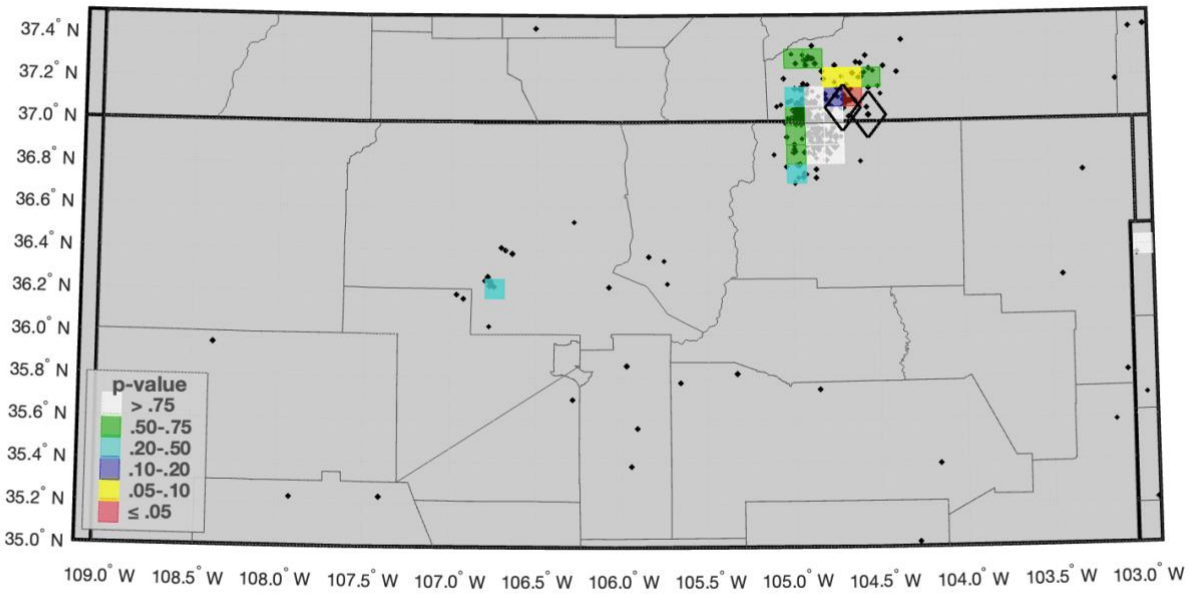

Figure 27. Same as Figure 26, but using declustered earthquakes above M 2.5.

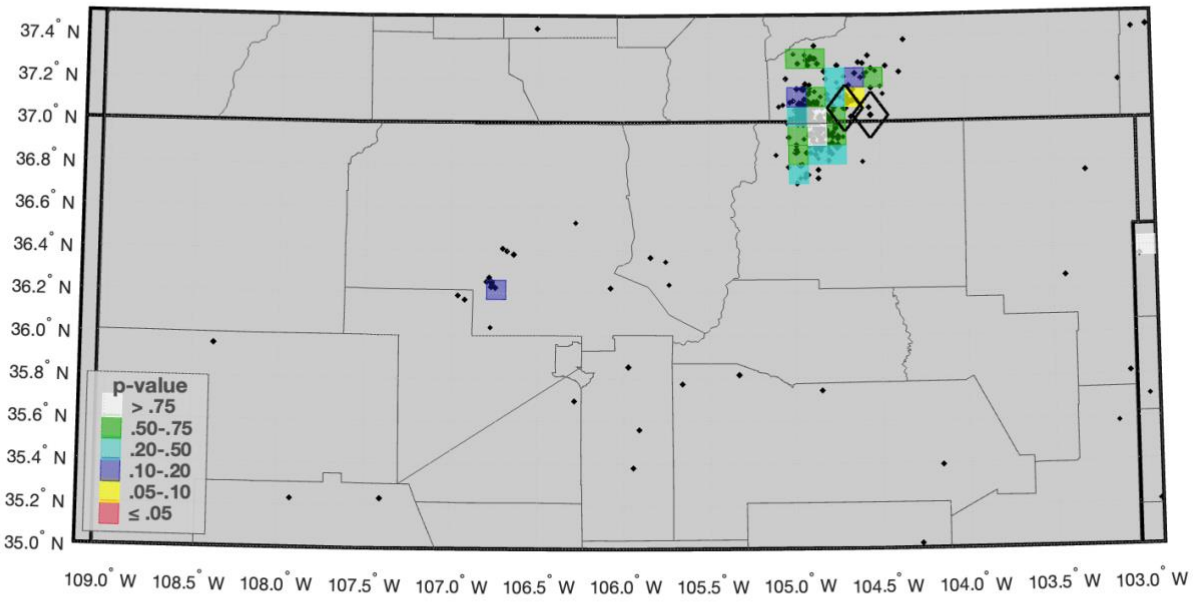

Figure 28. Same as Figure 26, but using declustered earthquakes above M 2.5 and a D of 0.3 m^2/s.

4.4 Fort Worth basin

The Fort Worth basin (FWB) encompasses the Barnett Shale formation, which has served as a major shale-gas producing unit since 2004, and benefiting from HF operations. Co-produced wastewater during production is primarily disposed into the underlying Ellenburger dolomitic limestone formation (Hornbach et al., 2016), which lies just above the crystalline Precambrian basement. The basin hosts the Dallas–Fort Worth (DFW) metropolitan area, with a population over 6 million.

There are six main earthquakes sequences in the basin. The 2008–2015 DFW Airport sequence, the 2009-2012 Cleburne earthquakes in southwest Johnson County (south of the DFW metroplex), the Azle-Reno sequence (northwest of Fort Worth), the Irving-Dallas sequence, the Lake Lewisville sequence and the 2015 Mw 4.0 Venus sequence in northeast Johnson County, which hosted the May 2015 Mw 4.0 earthquake, the largest ever recorded in the basin. Each of these sequences exhibits swarm-like behavior, rather than typical mainshock–aftershock characteristics (Quinones et al., 2019). Overall, seismicity in the FWB does have a strong spatial and temporal correlation with SWD with the majority of seismicity occurring within 15 km of SWD wells (Quinones et al., 2019). The main exceptions to this are the Irving–Dallas and Lake Lewisville sequences, which have no SWD wells within 15 km. The spatial proximity of these near-well sequences, along with the strong temporal correlation between the onset of seismicity and increasing disposal rates within the basin, suggests that pore-pressure diffusion is a key driving force behind these sequences (Quinones et al., 2019). Furthermore, ever decreasing disposal rates from their peak levels in 2014 do appear to coincide with lower rates of seismicity across the basin.

From 2013 to 2018, local seismic research networks operated by Southern Methodist University (SMU) have provided basic earthquake data for the Fort Worth basin (FWB), an area that was aseismic until 2008. Their design aimed at resolving the depth and of focal mechanism of small magnitude nearby earthquakes for the Azle–Reno, Irving–Dallas, and Venus sequences. The networks are not designed for regional earthquake monitoring within the FWB (Quinones et al., 2019). Despite these limitations, the estimated regional M_c is estimated to be around 2.1 after the year 2013, with the earthquakes rupturing normal faults located primarily within the Precambrian basement formation (Quinones et al., 2019). These local studies have been unified into a single dataset, which has been termed the North Texas Earthquake Study Catalog (NTXES; Quinones et al., 2019; Heather et al., 2019). The EarthScope USArray data passed through the FWB with a 70-km station spacing between November 2009 and September 2011. Frohlich (2012) documented more than earthquakes down to magnitude 1.5 occurring across the basin in discrete clusters, nearly eight times the rate reported in ComCat over the same time period.

Seismicity at the DFW Airport began in October 2008 with a series of earthquakes with magnitudes up to m_{bLg} 3.0, with a larger m_{bLg} 3.3 event occurring in May 2009. They were located just south of the Dallas-Fort Worth Airport at a depth of 4.8 km (Frohlich et al., 2010). Since January 2008, 13 wells have been drilled and hydraulically fractured in the Barnett Shale within 3 km of the cluster. These wells began production on dates ranging from June 2008 to March 2009 (Frohlich et al., 2011), so they could have been stimulated at any point between their drilling and production-start dates. Two SWD wells were also drilled, one northeast and one southwest of the DFW airport.

The southwest SWD well is about 200 m south of the mean DFW focus, much closer than any of the production wells where HF occurred. The permitted disposal interval (3.3–4.2 km) is in the Ellenberger formation, only slightly above the initial focal depths of the earthquakes (Frohlich et al., 2010). Disposal at the southwest SWD well commenced on 12 September 2008, just seven weeks before the first DFW earthquakes were detected. Both SWD wells reached disposal rates in excess of 250,000 bbls/month (Frohlich et al., 2011). 3D pore-pressure diffusion modeling by Ogwari et al. (2018) suggests that the high-rate nearest SWD well is the primary driver of the seismicity. Since the shut-in of the nearest SWD well, the sequence continued into at least 2015 slowly migrating northeast along the 6 km causative fault (Ogwari et al., 2018).

On 9 June 2009, an m_{bLg} 2.8 earthquake near Cleburne was felt by a community that had never felt earthquakes in the past. The located events of this sequence were within 3km from two SWD wells, less than a few hundred meters from the injection depth, and not near other types of production activity (Justinic et al., 2013). One of these wells frequently had over 300,000 bbls/month average rate since 2005. The felt earthquakes started a number of years after the initiation of the high-rate injector and between 6 and 8 months after the disposal rate was nearly doubled. Additionally, the earthquake sequence decreased significantly six months after this injector was shut-in while the closer injector continued to operate but with significantly smaller volumes of fluids (Justinic et al., 2013).

Beginning in November 2013, a series of earthquakes occurred near the city of Azle and Reno, about 25 km northwest of Fort Worth. As of 2015, the sequence includes eight earthquakes having magnitudes of 3 or greater; with the two largest and widely felt having m_{bLg} equal to 3.6 (Frohlich et al., 2016). Two high-volume wastewater injection wells and more than 70 production wells that produce gas and brine are situated within 10 km of the Azle earthquake sequence (Hornbach et al., 2015). Hornbach et al. (2015) concluded that on the basis of 3D pore-pressure modelling results and the absence of historical earthquakes near Azle/Reno, brine production combined with wastewater disposal represent the most likely cause of seismicity there. That said, coupled linear poroelastic modeling by Zhai & Shirzaei (2018) found that the impacts of brine production on stress change and seismicity rate were negligible.

The Irving–Dallas sequence generated significant felt earthquakes within the NTXES catalog, including two m_{bLg} 3.5 and 3.6 events on 6 January 2015 (Heather et al., 2019). This sequence remains the most enigmatic one in the basin; it occurred on an optimally oriented for failure within the given stress field, but it is unclear whether poroelastic stress changes alone would be enough to induce slip on the fault (Quinones et al., 2019). It is very far from both SWD and HF wells (Figure 29).

On 7 May 2015, a M_w 4.0 earthquake occurred in Johnson County, near Venus, not far from the Cleburne sequence. Five SWD wells were within 5 km the Venus cluster and at a small vertical distance from its hypocenters. Scales et al. (2017) found that SWD between 2006 and 2015 at these wells led to an increase in subsurface pore fluid pressure that contributed to the triggering of this long-lived sequence. They also saw no definitive spatial or temporal correlation between individual SWD wells (in terms of start times, rates, or volumes) and the earthquake history.

Next, Quinones et al. (2019) notes that in the geomechanical studies of Ogwari et al. (2018), Hornbach et al. (2016) and Zhai and Shirzaei (2018), stress changes associated with direct pore-pressure effects are highly concentrated at close distances to the wells, with poroelastic effects being a prerequisite for any far-field triggering of seismicity. They question however if poroelastic stress changes alone are large enough to have induced slip far-field, given that the actual magnitude of the poroelastic stress changes is lower than the near-field pore-pressure effects (Segall and Lu, 2015).

As far as the regional behaviour of faults is concerned, Hennings et al. (2019) noted strong heterogeneity in the response of the FWB faults, with many non-seismogenic faults experiencing higher pore pressure increase compared to those that became seismogenic. Faults proximal to SWD operations became seismogenic at appreciably higher mean pore pressure increase than those in distal areas (their approach ignores poroelastic effects). Crucially, they found no straightforward relationship between pore pressure increases or FSP values, and the rate or maximum magnitude of earthquakes hosted on the same faults. FSP stands for fault slip potential and employs a linearized Mohr Coulomb failure criterion, which defines critically stressed conditions by comparing the ratio of shear stress to effective normal stress on the fault before and after injection, and by considering the orientation and geometry of the fault relative to the regional stress field (Walsh & Zoback, 2016).

Finally, despite the focus on SWD, the FWB hosts significant levels of HF activity that (at least) spatially overlaps with some of the aforementioned clusters (Figure 29).

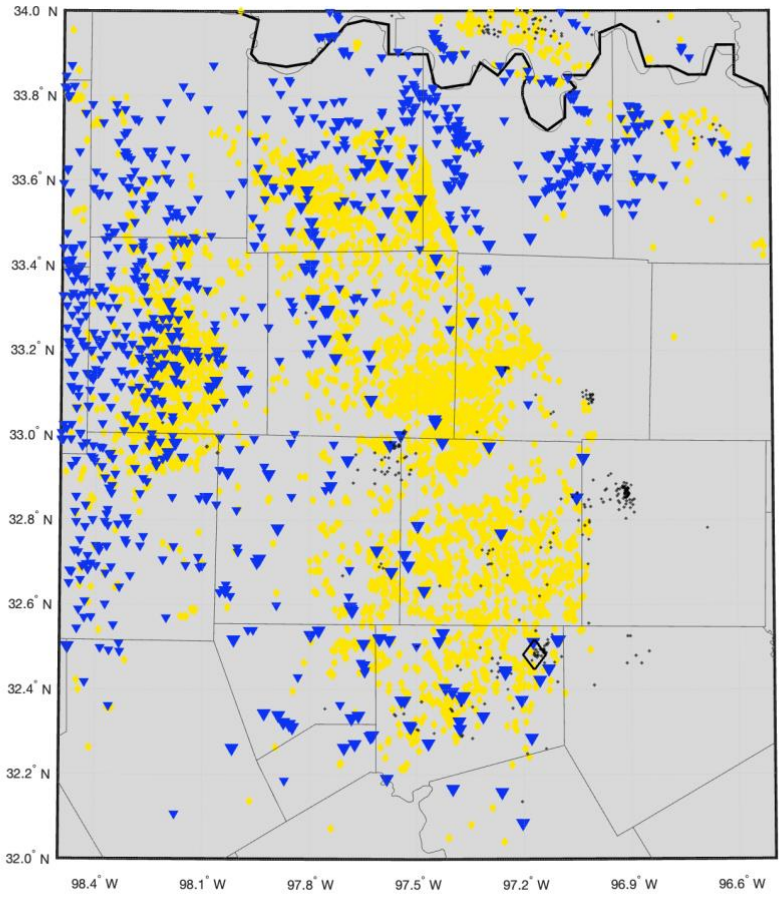

Figure 29. Map of the given investigated area showing declustered seismicity (M ≥ 2; black dots), wells (blue: SWD; yellow: HF) and county borders. Data between 2000 and Sep 2022. The black diamond indicates the Venus earthquake (M_w 4.0).

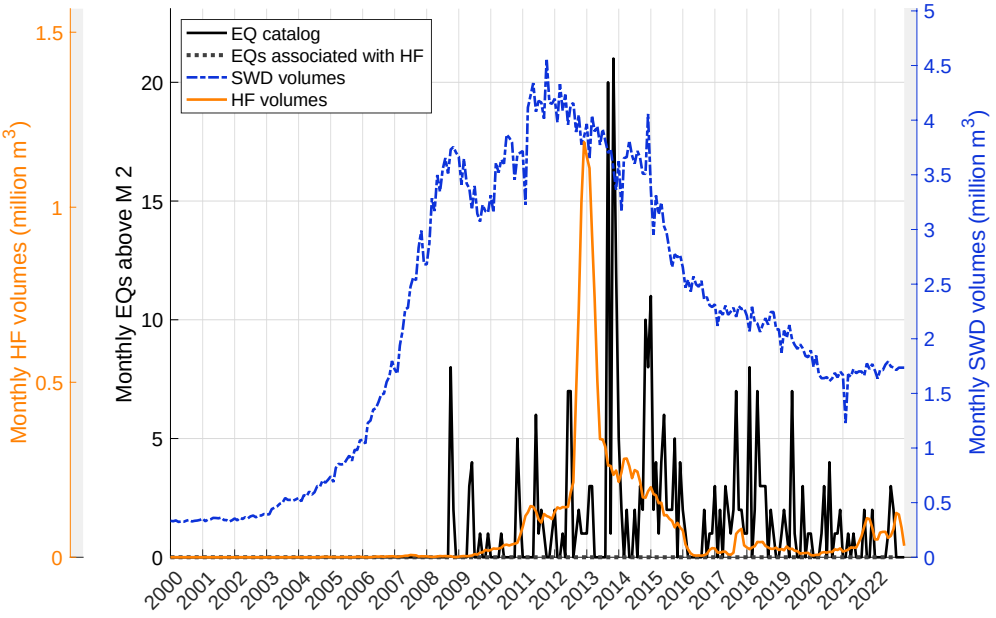

Figure 30. Time history of monthly HF & SWD volumes, and declustered seismicity rates above M 2, for the given investigated area. The HF data are incomplete before 2012.

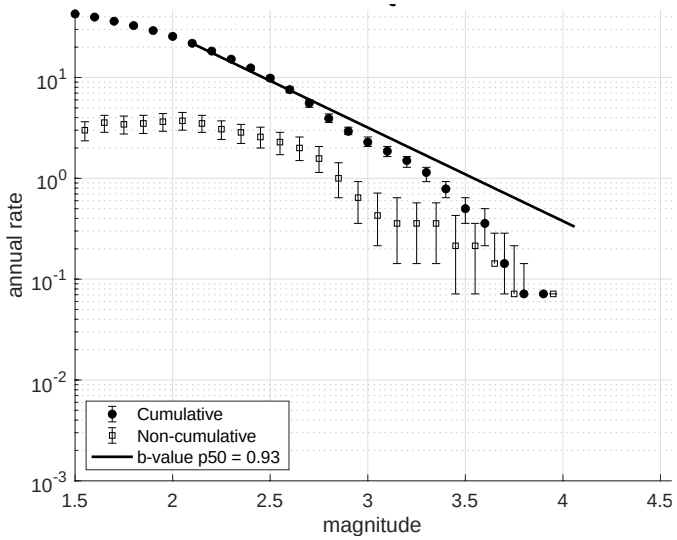

Figure 31. Bootstrapped (non-declustered) magnitude-frequency-distribution for the given investigated area, along with its fitted (median) b-value. Data between 2008 and 2021.

4.5.1 HF

In this section we will analyze the potential link between the observed seismicity in the FWB and HF. The gridded spatial distribution of total HF stimulation-days and the earthquake epicenters are mapped in Figure 32. The Irving-Dallas sequence is the only that is obviously not related to HF. The most intense HF activities occurred close to the border with OK and north of the Venus earthquake.

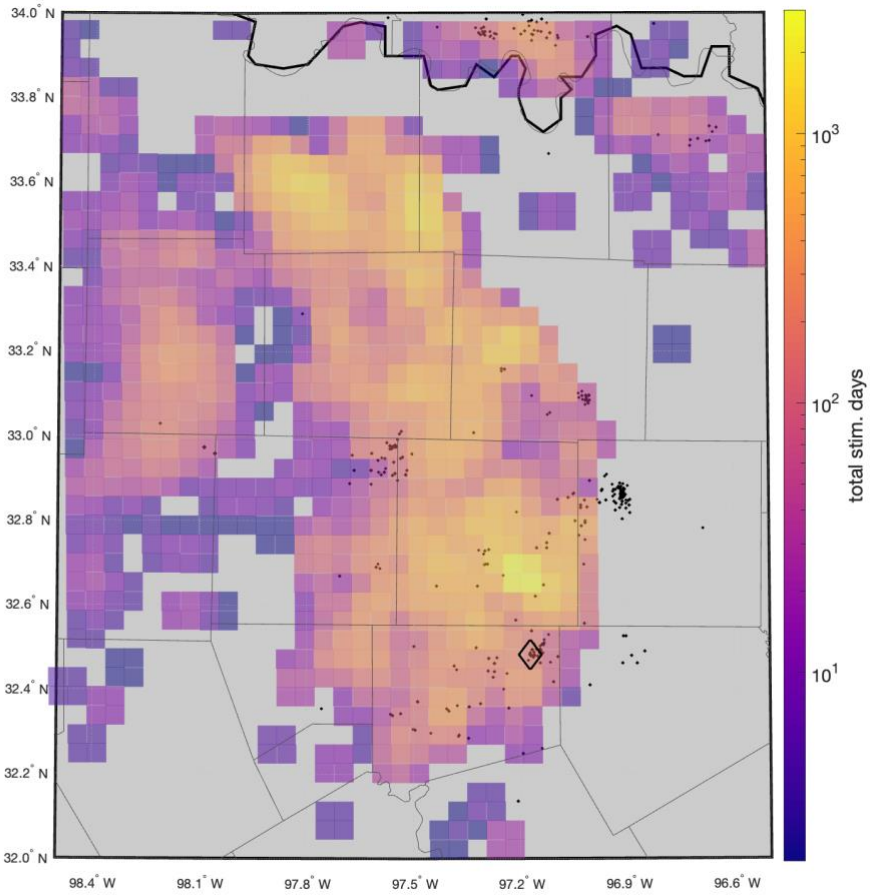

Figure 32. Map of the given investigated area showing declustered seismicity (M ≥ 2; black dots), gridded cumulative HF stimulation-days and county borders. The black diamond indicates the M_w 4.0 Venus earthquake. Data between 2008 and September 2022.

Figure 33 shows HFR values for the region, for the period between 2008 and September 2022, and for declustered earthquakes with magnitudes above 2. The majority of blocks have low values (below 0.10), validating the little focus paid to HF. According to our own association-criteria (section 3.4), just 4% of declustered earthquakes are associated with only 0.2% of the HF stimulations. There is one unnamed cluster at the border with OK that has very high HFR value and is likely HF-induced. This is a new finding that has not been reported thus far. The p-values tell the same story (Figure 34), including the fact that the Lake Lewisville sequence is not affected by HF.

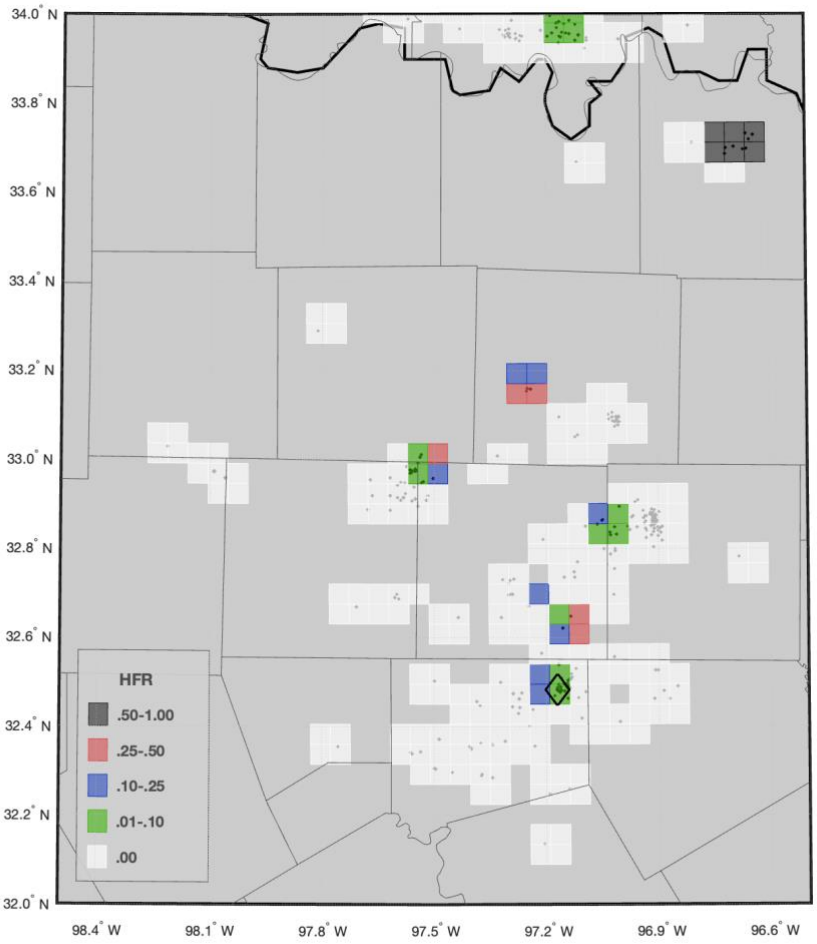

Figure 33. HFR values mapped on a 5 km grid, with all declustered earthquakes between 2008 and September 2022 above M 2 overlapped. The black diamond indicates the M_w 4.0 Venus earthquake. Country borders are also mapped.

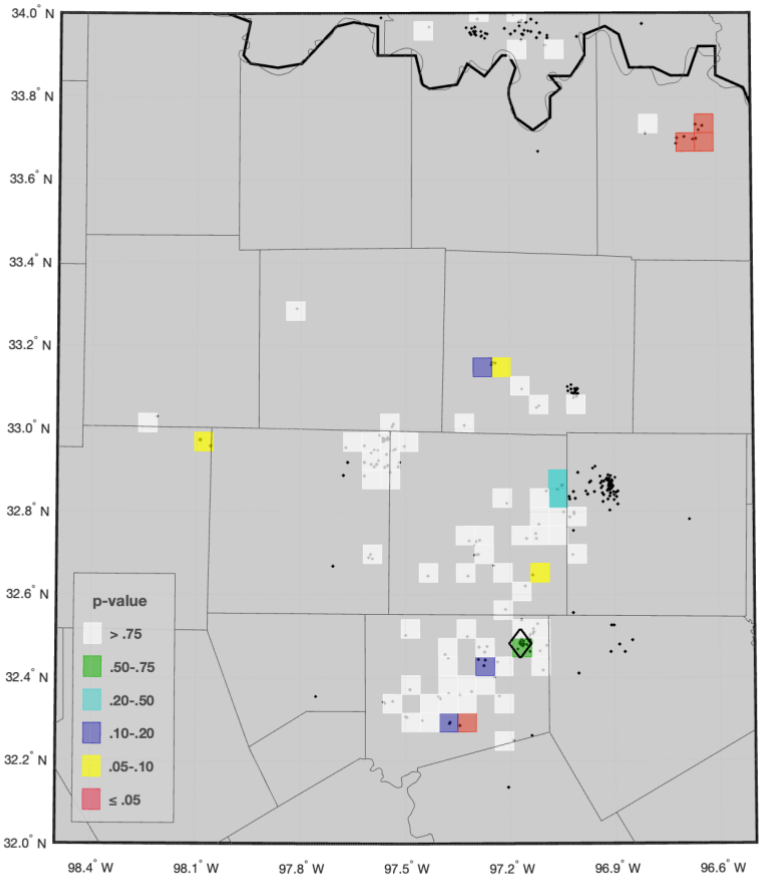

Figure 34. Statistical p-values for HF between 2008 and September 2022, with all declustered earthquakes above M 2 overlapped. Only blocks with at least 3 events are color-coded. The black diamond indicates the M_w 4.0 Venus earthquake. Country borders are also mapped.

4.5.2 SWD

The Ellenburger formation and underlying basement are fractured (Ewing, 1990), enhancing permeability. This is why Zhai & Shirzaei (2018) opted for a D of 0.7 m^2/s, which is the upper value from the range of Hornbach et al. (2015). Here we show results for D equal to 2 m^2/s, because this value leads to the highest confidence intervals for association with SWD. Lower values (1 or 0.3) lead to higher p-values (lower confidence). The non-declustered catalog also leads to higher p-values than the declustered one.

Figure 35 shows that the cumulative injected SWD volumes across the basin. It is notable that many earthquakes occur in blocks with no SWD wells. Figure 36 shows that the same volumes after diffusion in space and time. After diffusion the SWD volumes have reached (to a varying degree) all the blocks.

Figure 37 shows the statistical p-values for SWD (p_{SWD}) in the region, for the period between 2008 and 2021, for $M \geq 2$ and the large-scale diffusivity value D set to 2 m^2/s. SWD appears to be the sole causal factor behind the Irving-Dallas and Lake Lewisville sequences. These are novel findings, not previously reported in the literature. Notably, no correlation whatsoever is observed for the Venus or Cleburne earthquakes. A relatively low, but inconclusive, p-value is computed for the Azle sequence.

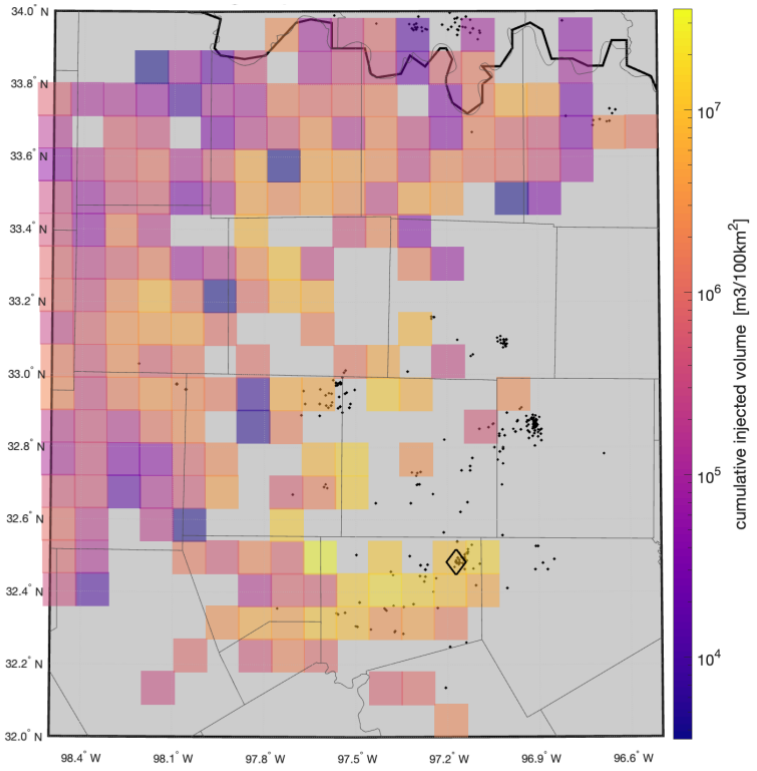

Figure 35. Map of the investigated area showing declustered seismicity (M ≥ 2; black dots), gridded cumulative SWD volumes and county borders. Data between 2006 and 2021. The black diamond indicates the M_w 4.0 Venus earthquake.

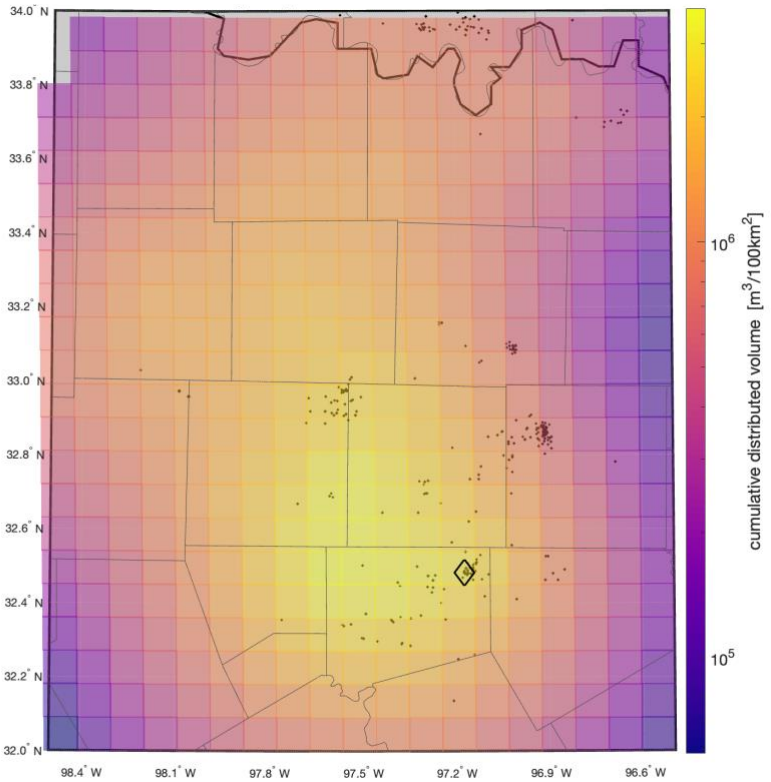

Figure 36. Map of the investigated area showing declustered seismicity (M ≥ 2; black dots), gridded diffused cumulative SWD volumes (D=2m²/s) and county borders. The black diamond indicates the M_w 4.0 Venus earthquake. Data between 2006 and 2021.

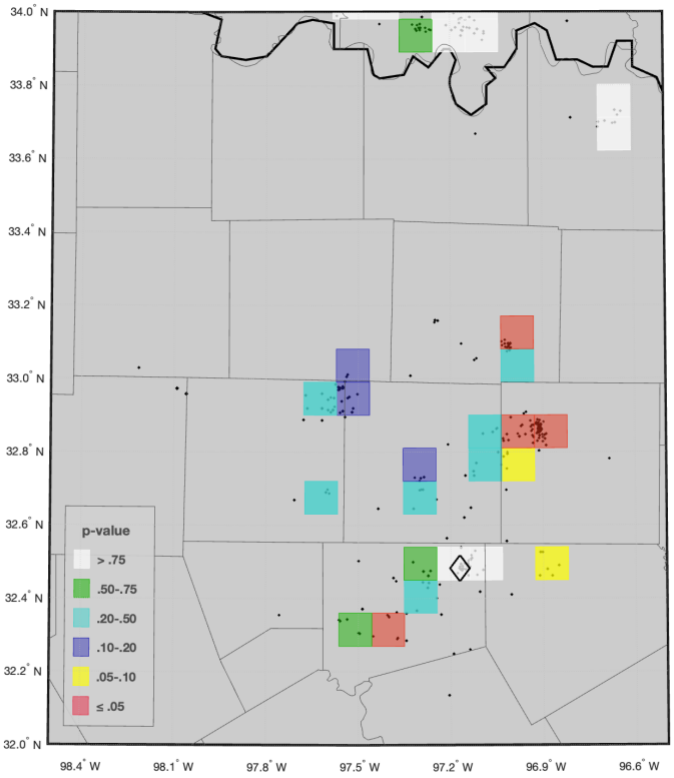

Figure 37. Statistical p-values for SWD between 2008 and 2021, for D equal to 2 m^2/s, with declustered earthquakes above M 2 overlapped. Only blocks with at least 3 events are color-coded. The black diamond indicates the Venus earthquake. Country borders are also mapped.

4.5 Oklahoma & Kansas

Historically, the seismicity rates in Oklahoma were very low, at about two to four earthquakes above M_w 3 per year. However, around 2009 these rates started to increase, then accelerated rapidly around the year 2014, and peaked at the end of 2015 (Figure 38). The largest event was a M_w 5.8 in Pawnee (Yeck et al., 2017), with four events larger than M 5 in total (one in 2011 and three in 2016). At the beginning the epicenters were mostly concentrated in the center of the state but then migrated further north and northwest. Since 2016, the seismicity rates have been steadily decreasing are now much closer to the historical values. Several scientific studies attributed this surge in seismicity to the disposal of large quantities of coproduced wastewater into deep laterally extensive aquifers (e.g. Arbuckle) overlying the Precambrian basement (e.g. Grigoratos et al., 2020b). The SWD rates have been gradually increasing since at least the year 2000, with accelerated growth between 2011 and 2015 (Figure 38). Since then, they have been steadily decreasing over time. Notably, even though disposal continuously increased between 2000 and 2015, the earthquake rates did not increase until about 2009 and then accelerated in about 2014 (Grigoratos et al., 2020b). Thus, there is an apparent time lag of months to years between increases in the disposal rates and the onset of seismicity. That said, a similar time lag was not observed in the response of the seismicity after the disposal rates started their downward trend. Furthermore, a small spike in the SWD rates in 2018 coincided without lag with a similar spike in the seismicity rate. This issue requires further research.

Around 2013 the seismicity had migrated further north into southern Kansas, with the peak later in 2015 coinciding with the one in Oklahoma. Kansas had also experienced very little seismicity until that time, with only 15 events above M 3 since 1973 (Rubinstein et al., 2018). SWD within the Arbuckle formation has been the identified causal factor for Kansas as well (Rubinstein et al., 2018; Ansari and Bidgoli, 2020). However, this conclusion has been based on qualitative observations and is mostly deterministic. Here, we apply our probabilistic framework to validate it. We also re-examine the findings of Grigoratos et al. (2020b) for SWD in OK. Finally, Figure 39 demonstrates that the link in space between SWD and earthquake occurrence in OK and KN is not straightforward. In some regions, the seismicity is collocated with SWD wells, but in other regions there is no clear spatial correlation.

SWD is the primary but not the only source of increased seismicity within OK. Skoumal et al. (2018) highlighted that some earthquake clusters within OK have been induced by HF, especially within the SCOOP/STACK oil play.

To validate that claim, we also compute p-values for the HF stimulations, which in some areas overlap with SWD wells (Figure 39).

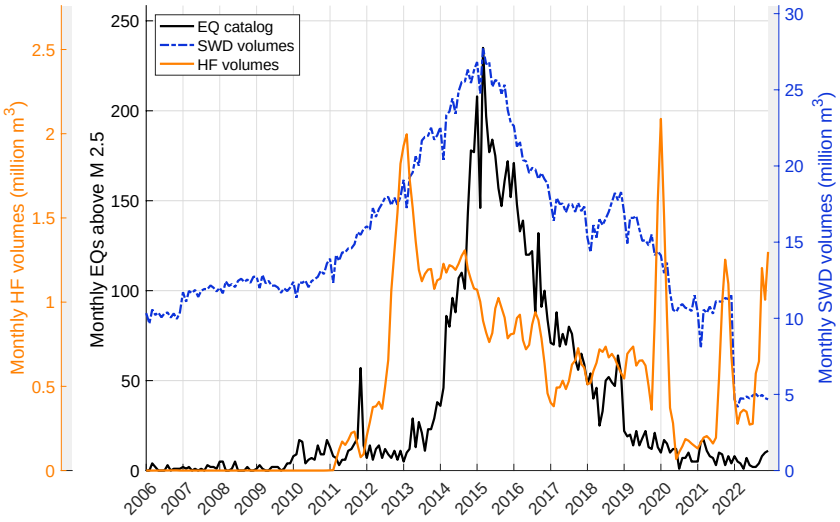

Figure 38. Time history of monthly HF & SWD volumes, and declustered seismicity rates above M 2, for OK and southern KN (latitude below 37.5). The HF data are incomplete before 2012, the seismicity data are incomplete below M 3 before 2011 and the SWD data are incomplete for 2022.

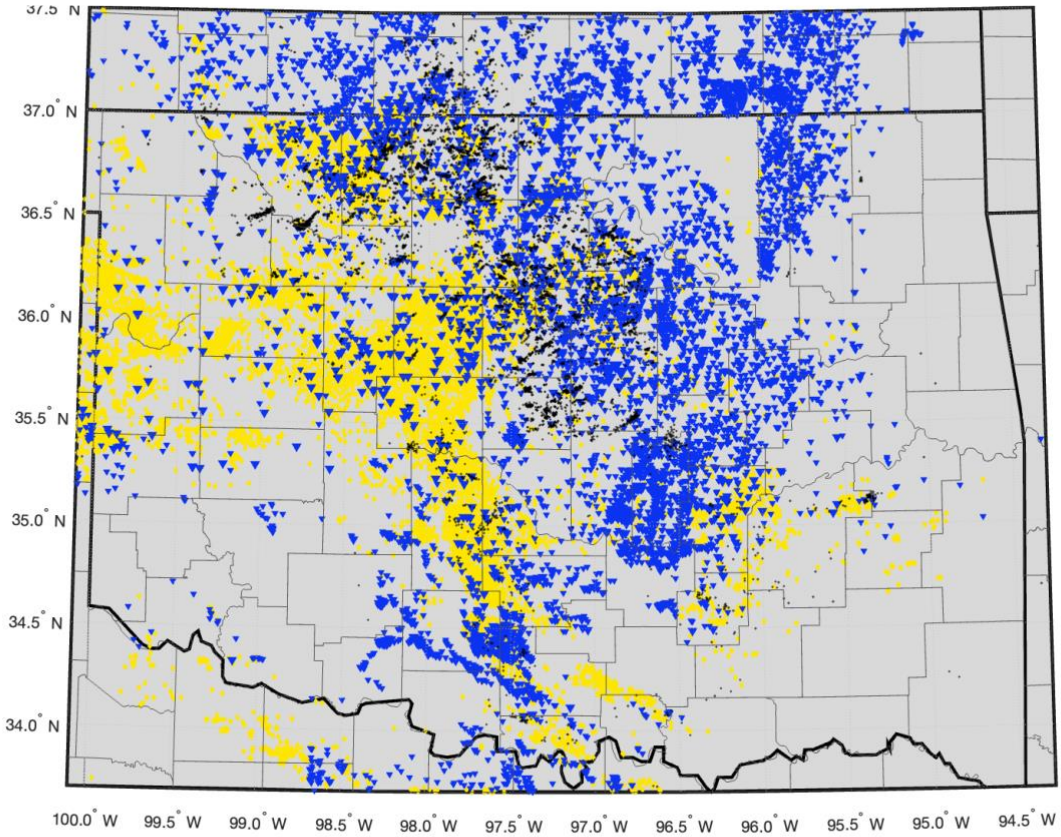

Figure 39. Map of the given investigated area showing declustered seismicity ($M \geq 2.5$; black dots), wells (blue: SWD; yellow: HF) and county borders. Data between 2006 and 2022.

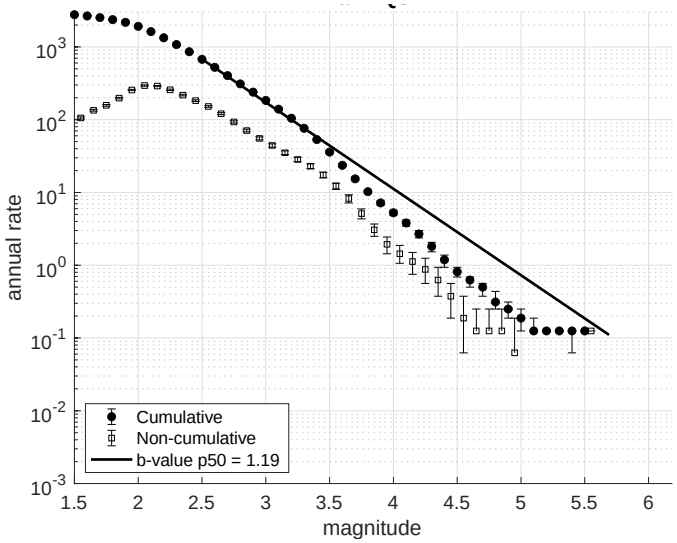

Figure 40. Bootstrapped (non-declustered) magnitude-frequency-distribution for the given investigated area, along with its fitted (median) b-value. Data between 2006 and 2021.

4.5.1 SWD

Within Oklahoma the D value ranges between 0.1 and 2 m^2/s (e.g. Goebel *et al.*, 2017). Grigoratos et al. (2020a) opted for *D* equal to 0.3 m^2/s. Here, we first show results for D equal to 2 m^2/s.

Figure 41 shows the statistical p-values for SWD (p_{SWD}) in the region, for the period between 2006 and 2021, for $M \geq 2.5$ and the large-scale diffusivity value D set to 2 m^2/s. In particular, 40% of blocks have $p_{SWD} \leq 0.05$, and 50% of blocks have $p_{SWD} \leq 0.10$. Furthermore, 54% of earthquakes occurred within a block with $p_{SWD} \leq 0.05$ and 65% within a block with $p_{SWD} \leq 0.10$. Lastly, one in three active SWD wells was within 5km of an earthquake. The results validated previous studies that have identified SWD as the main driver of seismicity in central and north western OK and in southern Kansas. That said the p-values for the blocks around Oklahoma city, Prague and south of Cushing have higher p-values than expected. This was also evident in the results of (Grigoratos et al., 2020b). This area is where the seismicity first started popping up in the early 2010s, and hosted the Prague M_w 5.7 event in 2011. It does not appear to be affected by HF either. Therefore, this issue requires further research. It does seem unlikely however that all the other clusters to the north are linked to SWD, but this area is unaffected. Perhaps, it there is a more complex interplay of pore pressure and background stress, because it had already accumulated enough tectonic strain. The Cushing earthquake is close to blocks with low p-values for its own block has an inconclusive p-value. The other two M5+ events (Fairview; Pawnee) occurred in blocks with $p < 0.05$ (95+ confidence interval for association).

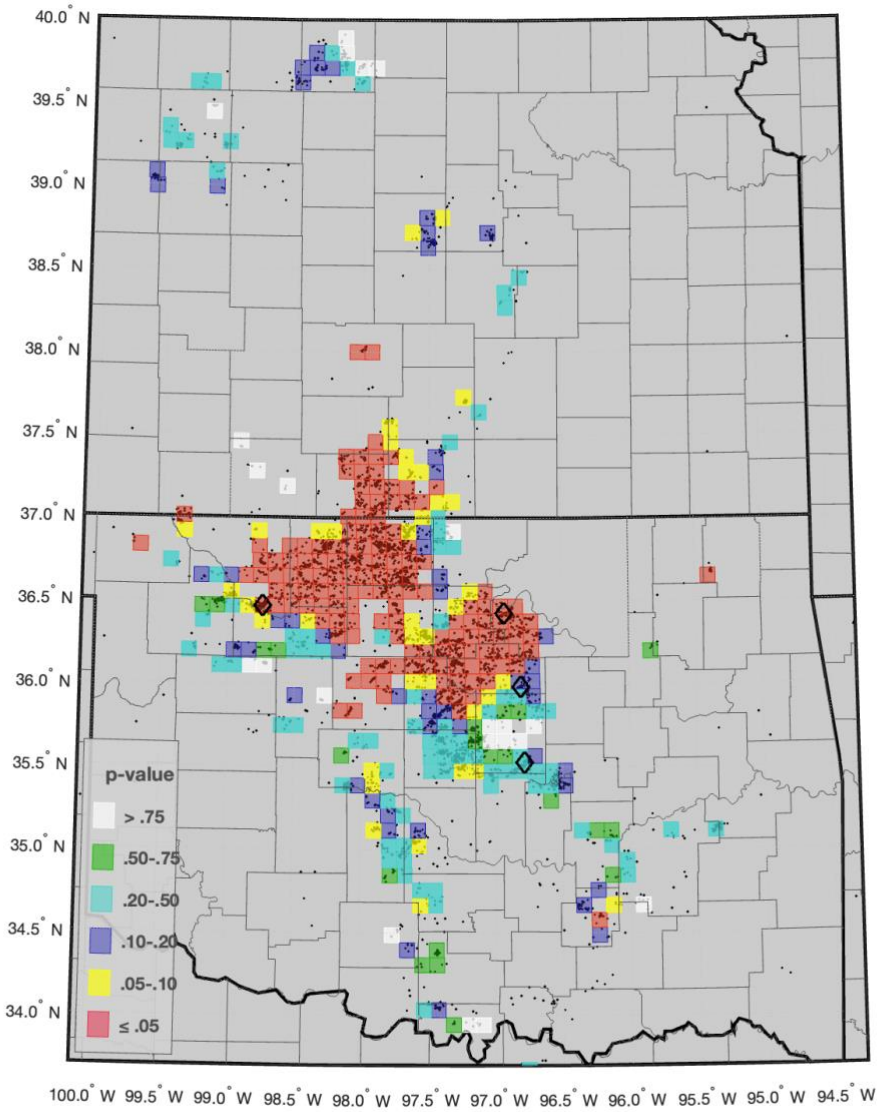

Figure 41. Statistical p-values for SWD between 2006 and 2021, for D equal to 2 m^2/s, with declustered earthquakes above M 2.5 overlapped. Only blocks with at least 3 events are color-coded. The black diamonds indicate the four events above M_w 5. Country borders are also mapped.

The results also validate Skoumal et al. (2018) who found that HF is the main driver in the SCOOP/STACK play, by not flagging SWD are a high confidence causal factor there. Our results also flag a few clusters in central Kansas that might be associated with SWD.

For D equal to 0.3 m^2/s (Figure 42) the results do not change dramatically, but they do flag a few blocks in SCOOP/STACK and lower the p-value in Fairview, where the SWD wells were very far from the cluster.

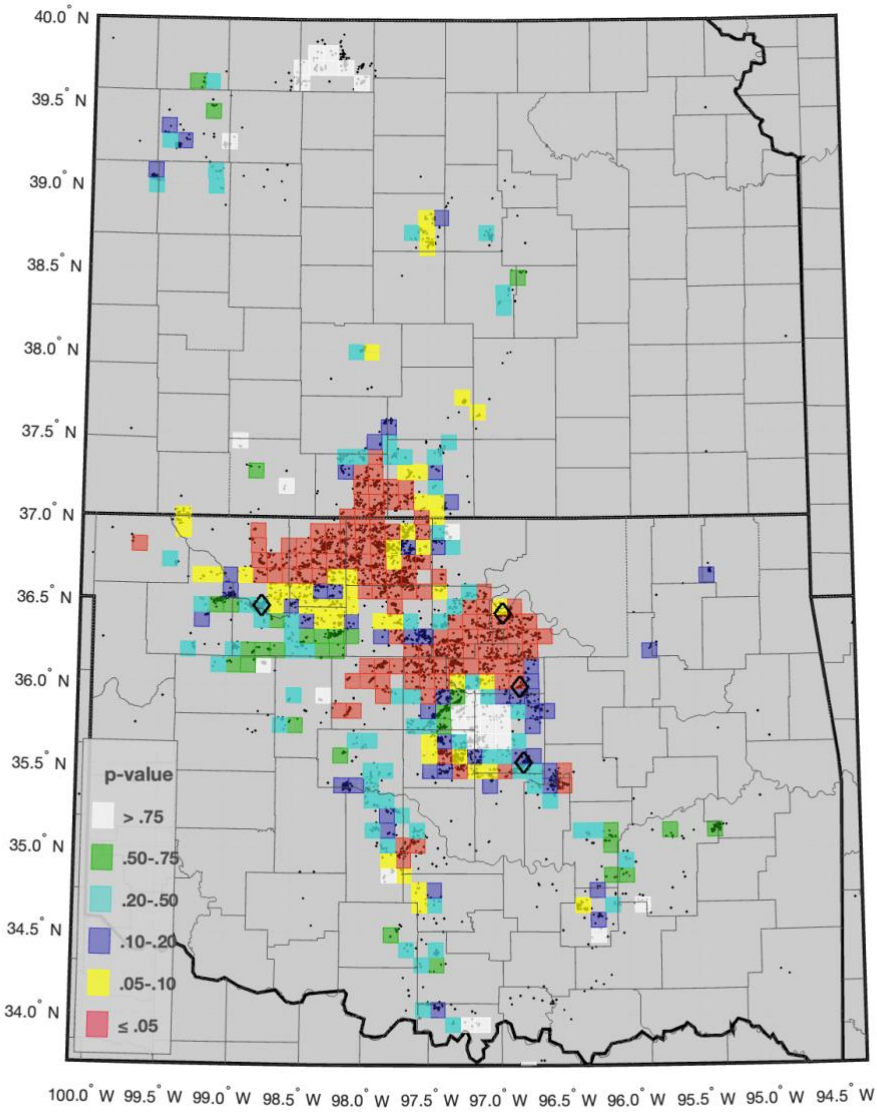

Figure 42. Statistical p-values for SWD between 2006 and 2021, for D equal to 0.3 m^2/s, with declustered earthquakes above M 2.5 overlapped. Only blocks with at least 3 events are color-coded. The black diamonds indicate the four events above M_w 5. Country borders are also mapped.

4.5.2 HF

In this section we will analyze the potential link between the observed seismicity in OK and southern KN, and HF. The gridded spatial distribution of total HF volumes and the earthquake epicenters is mapped in Figure 43. The largest volumes are in the SCOOP/STACK play, and in general in the south and west part of OK. Southern KN has very limited HF operations.

Figure 44 shows HFR values for the region, for the period between 2014 and 2022, and for declustered earthquakes with magnitudes above 2.5. According to our association-criteria, 7% of declustered earthquakes are associated with only 3% of the HF stimulations. Finally, 82% of HF stimulations that eventually induced seismicity according to our criteria, started doing so during the stimulation period. The vast majority of blocks in the south and within the SCOOP/STACK play have HFR above 0.25, as expected (Skoumal et al., 2018). The vast majority of blocks with low HFR values also have low p_{SWD} values, indirectly validating the results of the previous section. That said, there are several blocks with HFR between 0.10 and 0.25 and a few with HFR above 0.25 within zones affected by SWD. Those blocks also have low p_{HF} values in Figure 45, since once again HFR and p_{HF} values tell the same story. Therefore, there are areas within OK where the seismicity is likely occurring at different depths, with shallow earthquakes triggered by HF and deeper ones (within the basement) triggered by SWD. This is a novel finding of this study.

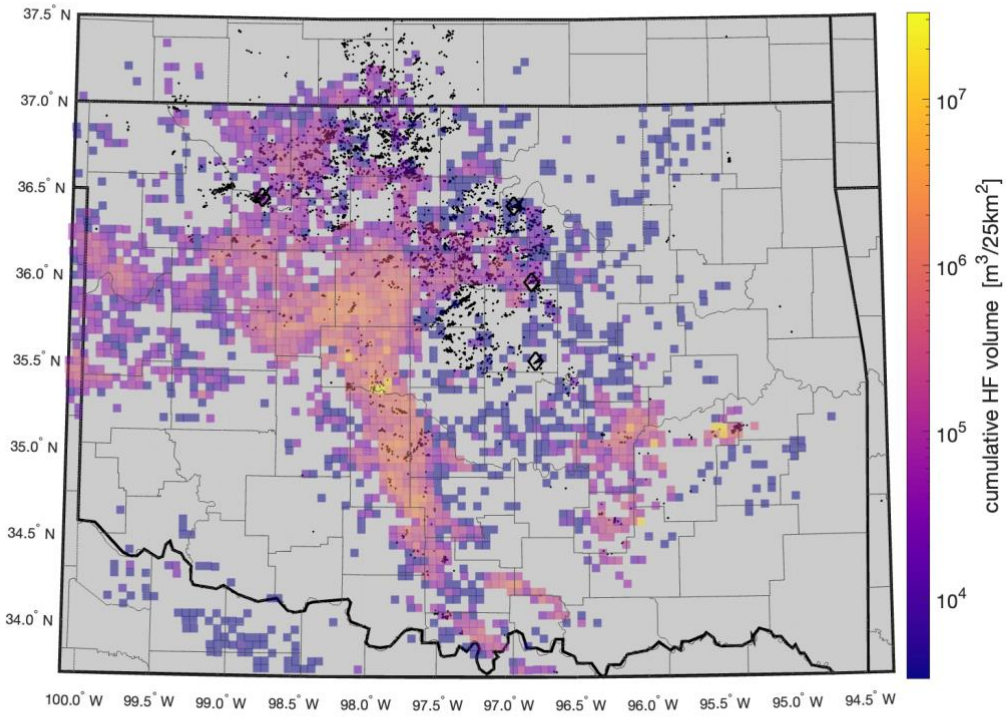

Figure 43. Map of the given investigated area showing declustered seismicity (M ≥ 2.5; black dots), gridded cumulative HF volumes and county borders. Data between 2014 and 2022. The black diamonds indicate the four events above M_w 5. Country borders are also mapped.

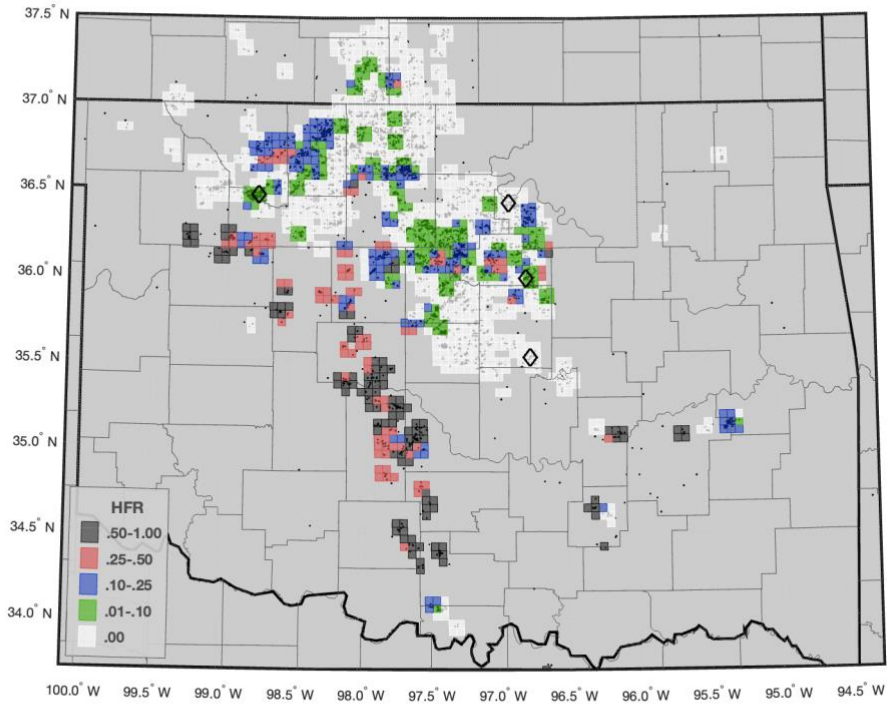

Figure 44. HFR values mapped on a 5 km grid, with all declustered earthquakes between 2014 and 2022 above M 2.5 overlapped. Only blocks with at least 3 events are color-coded. The black diamonds indicate the four events above M_w 5. Country borders are also mapped.

Figure 45 shows the statistical p-values for HF (p_{HF}) in the region, for the period between 2014 and 2022. In particular, 22% of blocks have $p_{HF} \leq 0.05$, and 35% of blocks have $p_{HF} \leq 0.10$. Furthermore, 8% of earthquakes occurred within a block with $p_{HF} \leq 0.05$. When we combine SAF with p_{HF}, 3% of declustered earthquakes above

M 2 are both linked to HF and occurred within a block with $p_{HF} \leq 0.05$. These earthquakes are almost certainly induced by HF. As far as the HF wells are concerned, only 1% of stimulations are both linked to seismicity and occurred within a block with $p_{HF} \leq 0.05$. On the other hand, 92% of stimulations are not linked to seismicity and occurred within a block with $p_{HF} \geq 0.05$. Lastly, 82% of blocks have a fitted time-lag (equation 4) of 3 days or less and the median Σ value among blocks with $p_{HF} \leq 0.05$ is -2.6.

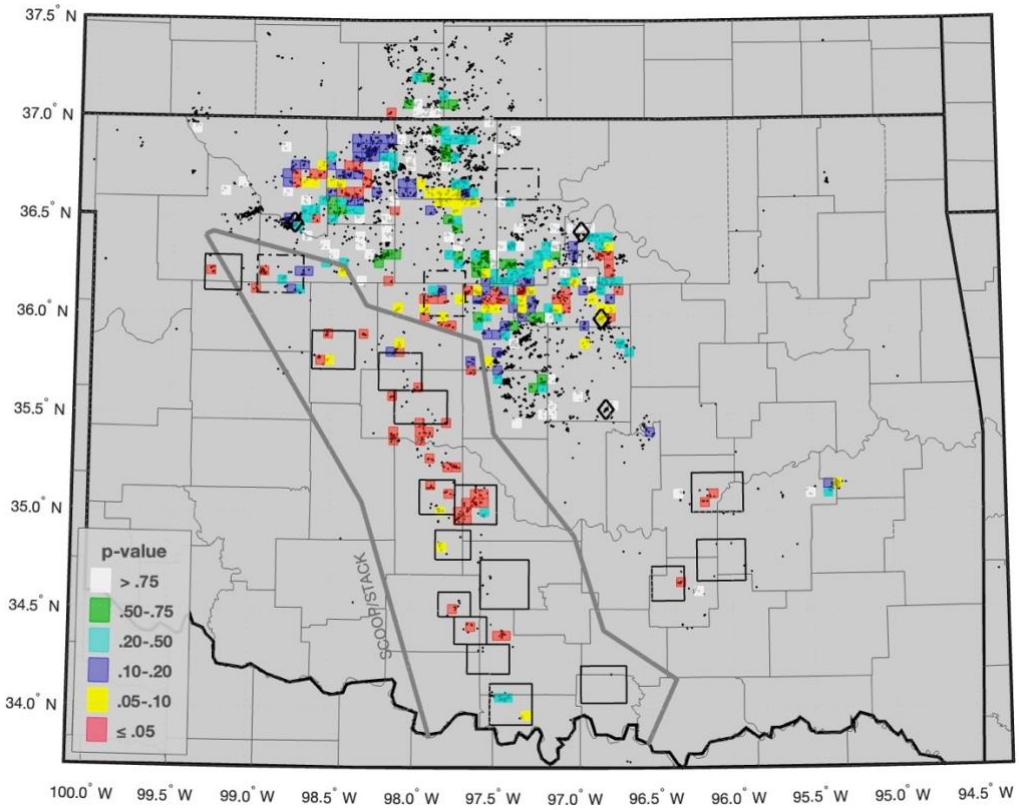

Figure 45. Statistical p-values for HF between 2014 and 2022, with all declustered earthquakes above M 2.5 overlapped. Only blocks with at least 3 events are color-coded. The black diamonds indicate the four events above M_w 5. Country borders are also mapped. Black rectangles are taken from Skoumal et al. (2018) and represent areas with seismicity linked to hydraulic fracturing (HF).

4.6 Permian basin

The Permian basin has three major sub-basins, the Delaware basin to the west, the Central Basin Platform (uplift buffer) and the Midland basin to the east. We examine the Delaware basin and the Midland basin separately, since there is no horizontal hydraulic communication between them.

Thanks to a recent study by Frohlich et al. (2020), we now know that seismicity rates in the Delaware basin have been gradually rising undetected since 2009, with further acceleration after 2013. Earthquake catalog completeness improved significantly after January 2017 with the establishment and continuous expansion of the Texas Seismological Network (TexNet; Savvaidis et al., 2019). The area has historically experienced very low seismicity levels, with only three events above M 4.8 in the last 100 yr. Namely, a damaging M_w 6.3 in 1931 near Valentine, a m_b 4.8 in 1975 near the city of Pecos, and a two recent M_w 5.0+ events after 2020 that might be linked to deep SWD (Tung et al., 2020; Grigoratos et al., 2022). A relation between several smaller historical earthquakes ($M < 3$) and oil and gas operations has been frequently suspected, but the evidence has been mostly inconclusive (Doser et al., 1992). That said, ongoing seismicity in the town of Kermit between 1966 and 1978 ($M < 4.1$) is likely related to enhanced oil recovery (EOR; Davis et al., 1985). In recent times, the attention has shifted almost entirely to SWD and HF operations, both within the Delaware basin (Savvaidis et al., 2020; Skoumal, Barbour, et al., 2020) and in neighboring states (Rubinstein et al., 2014; Skoumal et al., 2018), with alternative causal factors being flagged as less likely. EOR operations aim to stabilize, rather than increase, the pore pressure and given that the formations targeted by HF in the Delaware basin are already overpressured (Lee and Williams, 2000), we do not expect EOR wells to cause detectable seismicity (with the caveat that sometimes SWD wells are misclassified as EOR for licensing

purposes). The case for production (fluid extraction) was briefly examined by Skoumal, Barbour, et al. (2020), who concluded that the total stress change and the observed focal mechanisms do not support a causal link. The effects of production (fluid extraction) are generally also counteracted by the HF activity (fluid injection) within the same interval. Within the Delaware basin, the aggregate injection volumes of both SWD and HF wells have been increasing since 2013, with a notable acceleration after 2016, and are thus well correlated with the increasing seismicity rates in the time domain (Figure 49; 50; 51). The vast majority of SWD wells inject into the Delaware Mountain Group (DMG), composed of the Bell Canyon, Cherry Canyon, and Brushy Canyon formations. The HF wells, on the other hand, are targeting the Bone Spring and Wolfcamp formations, which is about 1 km below the DMG and close to an overpressured zone (Lee and Williams, 2000). Only a small number of SWD wells are close to seismic clusters while injecting into the deeper Devonian, Silurian, and Ordovician rocks, relatively close to the basement (Figure 46). That said, some of these wells have very large median injection rates, in excess of 10^6 bbls/month. This complex hydrogeologic profile (Hennings et al., 2021) is much more challenging than the two-layer assumption (aquifer directly above basement) adopted in Oklahoma. Although vertical flow within the DMG formations is to be expected, it is assumed that the interbedded low-permeability shale formations targeted by HF restrict vertical flow (Lee and Williams, 2000; Woo & Ellsworth, 2023), potentially insulating deeper faults from shallow SWD (excluding rare and often isolated poroelastic effects). The basement may also inhibit the vertical migration of fluids, although probably to a lesser extent (assuming it is locally fractured). We should note that all formations are trending shallower in the western parts of the basin, with their relative distance thereby decreasing.

Within Reeves county, almost all well-constrained hypocenters are shallow, a couple of kilometers below the HF wells (Grigoratos et al., 2022). In areas with improved station coverage or local velocity models, however, the hypocenters are even shallower (Sheng et al., 2022), within the DMG (shallow SWD) and the Bone Spring formation (HF). Grigoratos et al. (2022) hypothesized that there is systematic bias toward deeper depths within the catalog, perhaps by about 2 km. This is also somewhat evident in the data from Culberson County, an area with historically poor station coverage and complex 3D geologic effects not captured by the 1D velocity models. In Culberson County, there are no hypocenters within the first 2 km of the basement, even though the seismicity there is likely triggered by deep SWD wells injecting just above the basement top (Grigoratos et al., 2022). Finally, relatively lower levels of seismicity have been detected close to the border with New Mexico, with the exact reason yet to be determined.

Our catalog for the Delaware basin includes template-matched events by Skoumal, Barbour, et al. (2020) for 2017 and 2018. Because no template matching was available for 2019 and beyond, we incorporated an extension of the public TexNet catalog that includes earthquakes of limited manual review due to their smaller size. This enabled us to match the low magnitude of completeness (M_c) of the previous two years, that is, M_L 1.5. After 2021, the network was dense enough for the standard TexNet catalog to be sufficient.

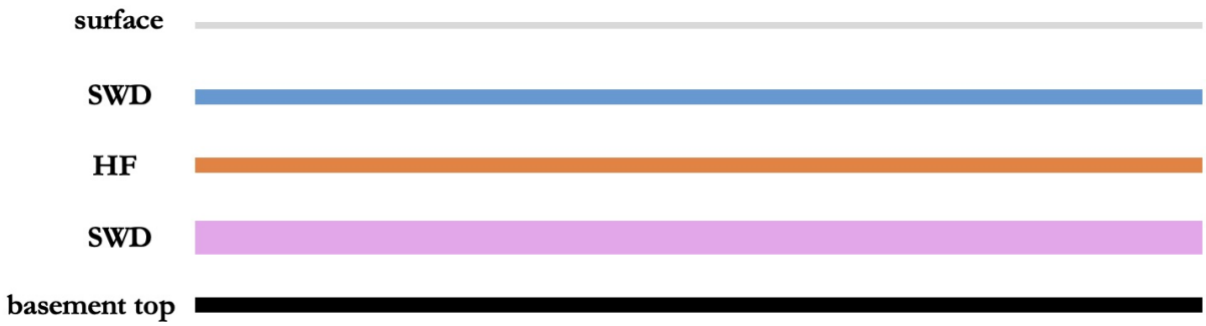

Figure 46. Schematic depth cross-section for the Permian basin.

Since 2020, the Midland basin (Midland basin) in west Texas, has experienced a rapid increase of seismicity, even in regions with no history of earthquakes (Frohlich et al., 2020). The HF rates have been mostly stable over time, while the SWD have been steadily increasing since 2011, with a reacceleration in 2017and a plateau after 2019 (Figure 47). In response to numerous felt earthquakes close to towns of more than 300,000 people, the Texas Railroad Commission (RRC) established Seismic Response Areas (SRAs) around Stanton and Gardendale to regulate both shallow and deep SWD wells in order to mitigate the seismic hazard and risk. There are four main areas of interest in the Midland, the Stanton SRA, the Gardendale SRA, the North Snyder zone and the East Snyder zone. As in the Delaware basin, also for the Midland basin we split the analysss in three, for shallow and deep SWD, and for HF (Figure 48).

On 31 December 2020, an M_L 4.2 earthquake occurred in Stanton. In the months following, there was ongoing activity with a cluster being initiated along a new fault (Woo & Ellsworth, 2023). On 28 December 2021, one year after the M_L 4.2 earthquake, a M_L 4.6 in the same area. As a result, in January 2022, the RRC designated the Stanton SRA (with the goal of reducing earthquakes greater than M 3.5. The vast majority of focal depths in the area range between 4.5 and 7.5 km, placing them below the basal Cambrian formations and the deepest SWD wells (Woo & Ellsworth, 2023). This, together with the fact that the seismicity above M 1 is continuous, implies that the most straightforward causal factor is deep SWD, as in northern OK, southern KN and the Culberson county in the Delaware basin. Woo & Ellsworth (2023) also identified a very low magnitude shallow cluster in Stanton that appears triggered by HF.

Starting in early 2020, the level of seismicity in Midland/Odessa (Gardendale) picked up significantly, including 12 events with M_L 3.5 or greater (Woo & Ellsworth, 2023). To reduce potential seismic risks in this area, the RRC established the Gardendale SRA in September 2021. In the next few months, 6 additional earthquakes above M_L 3.0 occurred. Crucially, on 16 December 2022, a M_L 5.2 earthquake occurred at the northeast edge of the SRA. The majority focal depths in the Gardendale area range from 5 to 7 km, placing the seismicity principally in the Precambrian basement (Woo & Ellsworth, 2023). This again, together with the fact that the seismicity above M 1 in most clusters is continuous, implies that the most straightforward causal factor is deep SWD.

The Cogdell oil field in north Snyder has been active for many decades and its seismicity has been associated with enhanced oil recovery operations (EOR; waterflooding), which since 2004 employ supercritical CO_2 (Davis and Pennington, 1989). This is not easily supported by the focal depths, which are in the crystalline basement (Woo & Ellsworth, 2023), and would require complex poroelastic processes to trigger seismicity that much deeper. No shallow SWD or HF wells are present within that area; only deep SWD ones (Figure 48). Curiously, the hypocenters form mostly horizontal alignments with one strike-slip focal mechanism identified by Woo & Ellsworth (2023),

Two new seismic clusters in East Snyder activated in 2020. Woo & Ellsworth (2023) claims their depths are shallow, likely not more than 4 km, although their catalog includes several shallow events there. They suggest that either HF or shallow SWD is the causal factor. East Snyder continues to be seismically active with a M_L 4.7 event in February 2023. The area has limited shallow and more deep SWD wells, as well as several HF wells.

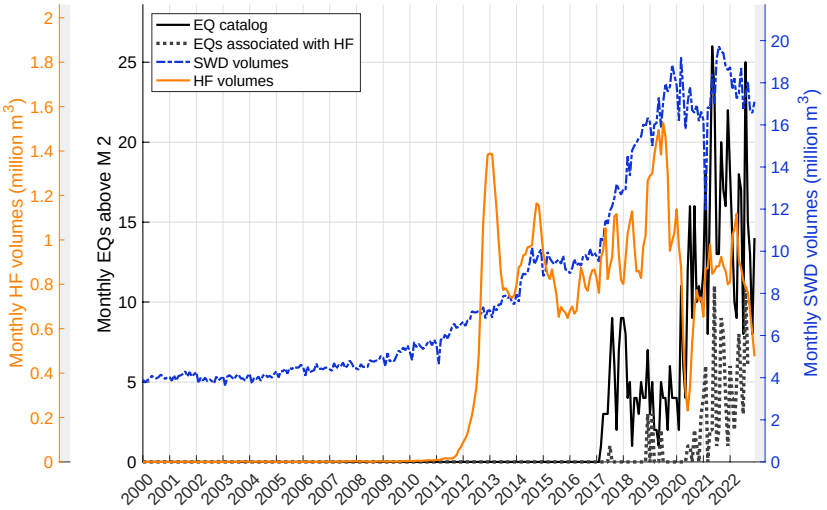

Figure 47. Time history of monthly HF & SWD volumes, and declustered seismicity rates above M 2, for the Midland basin. The HF data are incomplete before 2012, the seismicity data are incomplete below M 2.5 before 2017 and below 3.5 before 2017.

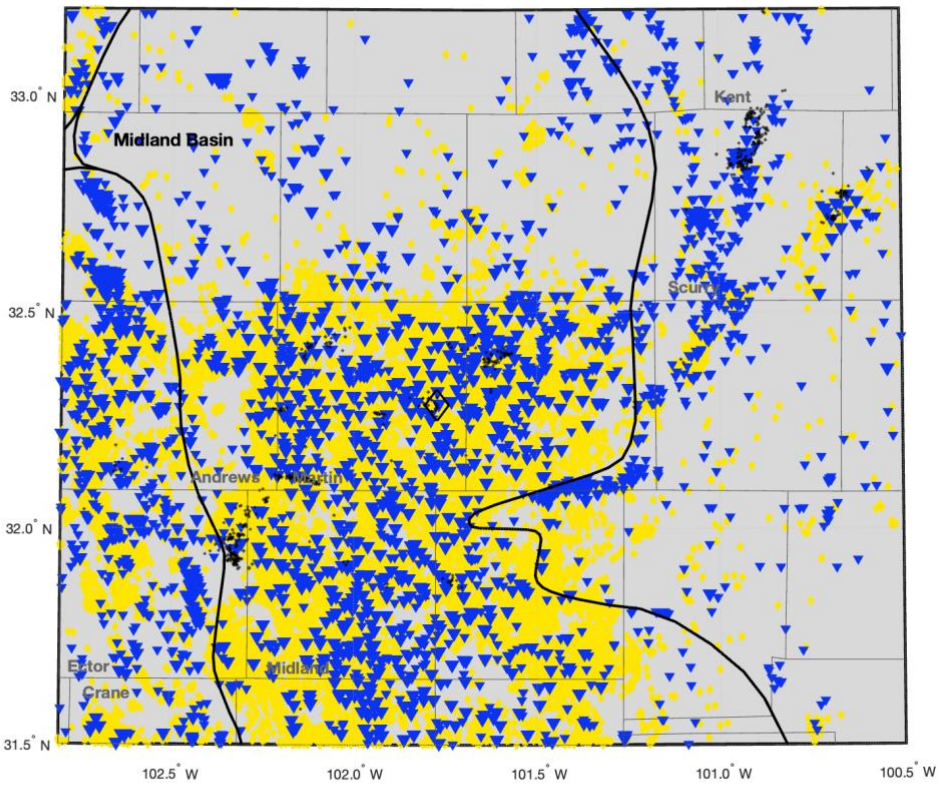

Figure 48. Map of the Midland basin showing declustered seismicity (M ≥ 2; black dots), wells (blue: SWD; yellow: HF) and county borders. Data between 2000 and September 2022. The black diamonds indicate the largest events.

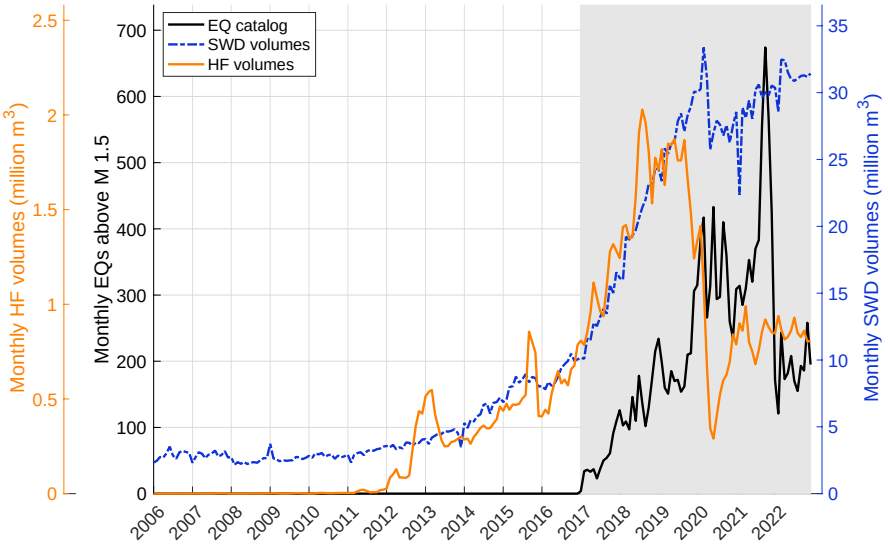

Figure 49. Time history of monthly HF & shallow SWD volumes, and declustered seismicity rates above M 1.5, for the Delaware basin. The HF data are incomplete before 2012, the seismicity data are incomplete before 2017.

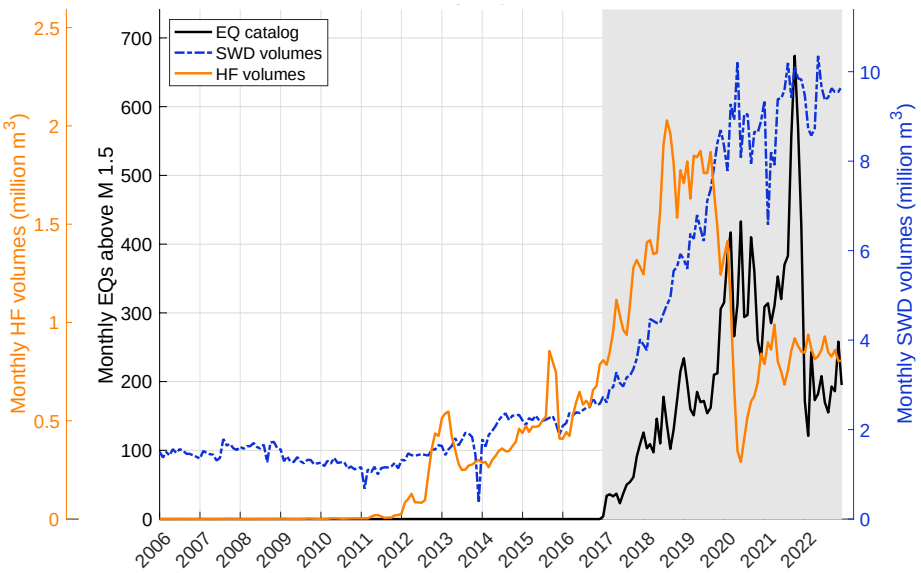

Figure 50. Time history of monthly HF & deep SWD volumes, and declustered seismicity rates above M 1.5, for the Delaware basin. The HF data are incomplete before 2012, the seismicity data are incomplete before 2017.

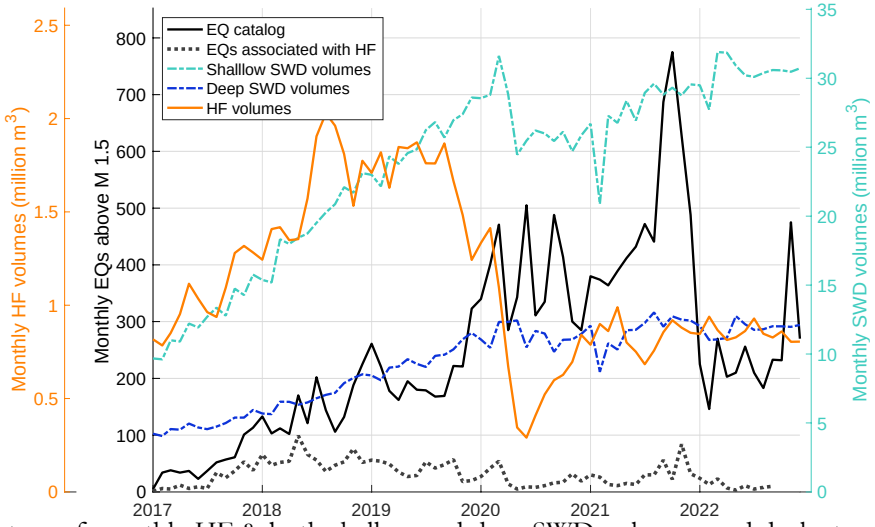

Figure 51. Time history of monthly HF & both shallow and deep SWD volumes, and declustered seismicity rates above M 1.5, for the Delaware basin.

4.6.1 Midland basin – HF

In this section we will analyze the potential link between the observed seismicity in the Midland basin and HF. The gridded spatial distribution of total HF volumes and the earthquake epicenters is mapped in Figure 52. The Stanton SRA seems be close to the largest total HF volumes in the basin. Overall, a spatial correlation between increased HF volumes and seismicity is lacking.

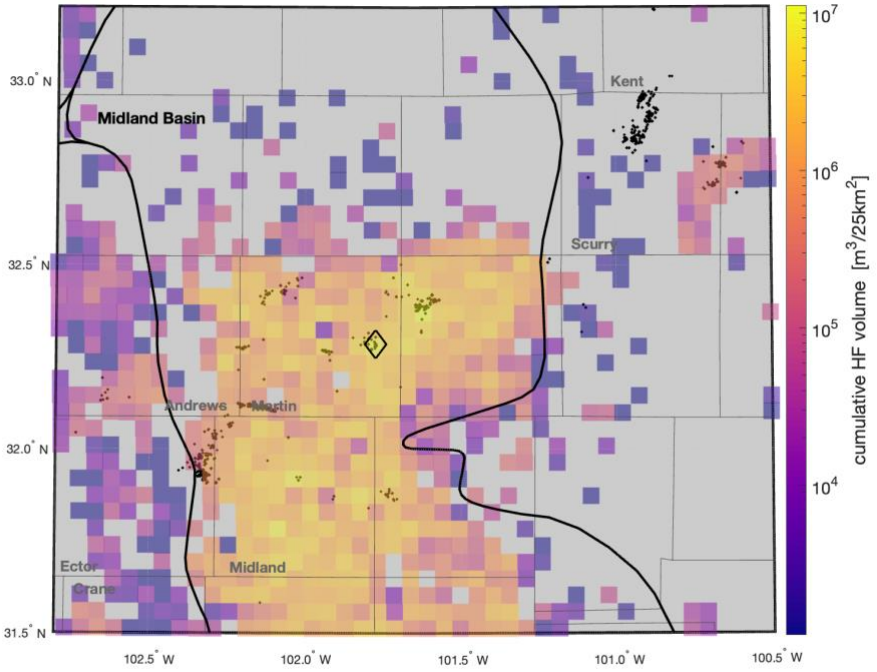

Figure 52. Map of the Midland basin showing declustered seismicity (M ≥ 2; black dots), gridded cumulative HF volumes and county borders. The black diamond indicates the largest earthquake in Stanton. Data between 2017 and September 2022.

Figure 53 shows HFR values for the region, for the period between 2017 and September 2022, and for declustered earthquakes with magnitudes above 2. Only a few blocks have high values (above 0.25), in east Snyder, east of Stanton and in the middle of the Gardendale SRA. According to our association-criteria (section 3.4), 19% of declustered earthquakes are associated with only 2% of the HF stimulations. Remarkably, 100% of HF stimulations that eventually induced seismicity according to our criteria, started doing so during the stimulation period.

Figure 54 shows the statistical p-values for HF (p_{HF}) in the region, for the period between 2017 and September 2022. In this case, they don't paint exactly the same picture as the HFR values. In particular, 14% of blocks have $p_{HF} \leq 0.05$, and 21% of blocks have $p_{HF} \leq 0.10$. Furthermore, only 10% of earthquakes occurred within a block with $p_{HF} \leq 0.05$. When we combine SAF with p_{HF}, only 7% of declustered earthquakes above M 2 are both linked to HF and occurred within a block with $p_{HF} \leq 0.05$. These earthquakes are almost certainly induced by HF. As far as the HF wells are concerned, less than 1% of stimulations are both linked to seismicity and occurred within a block with $p_{HF} \leq 0.05$. On the other hand, 97% of stimulations are not linked to seismicity and occurred within a block with $p_{HF} \geq 0.05$. Lastly, 75% of blocks have a fitted time-lag (equation 4) of 3 days or less and the median Σ value among blocks with $p_{HF} \leq 0.05$ is -3.9.

The p_{HF} analysis is inconclusive regarding HF-triggering in the Stanton SRA (Figure 54 & 55), but identified HF-seismicity east of it. No blocks with consistent HF-triggering are spotted in the Gardendale SRA and in east Snyder. Notably, all these findings are in agreement with the observations of Woo & Ellsworth (2023), despite not utilizing their catalog. They relocated hypocenters of the east clusters of Martin county and found that the vast majority of earthquakes are below the deep SWD injectors, but they also found some earthquakes within the HF interval. Their relocated hypocenters in north Scurry, Ector and west Martin counties were found to be below 4 km depth, below deep SWD.

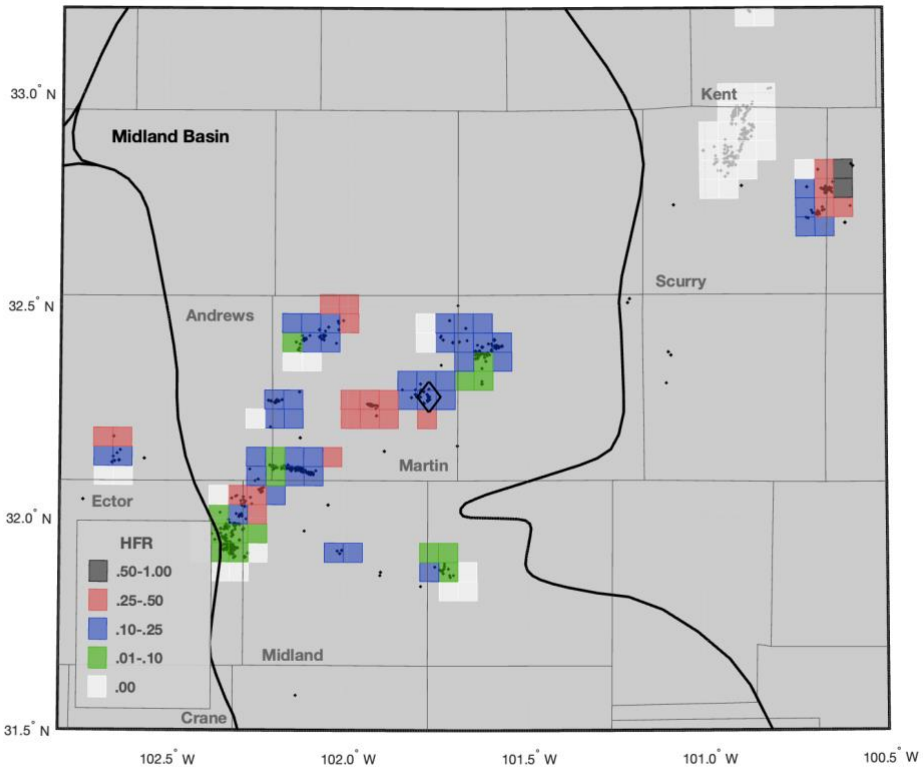

Figure 53. HFR values mapped on a 5 km grid, with all declustered earthquakes between 2017 and September 2022 above M 2 overlapped. The black diamond indicates the largest earthquake in Stanton. Only blocks with at least 3 events are color-coded. Country borders are also mapped.

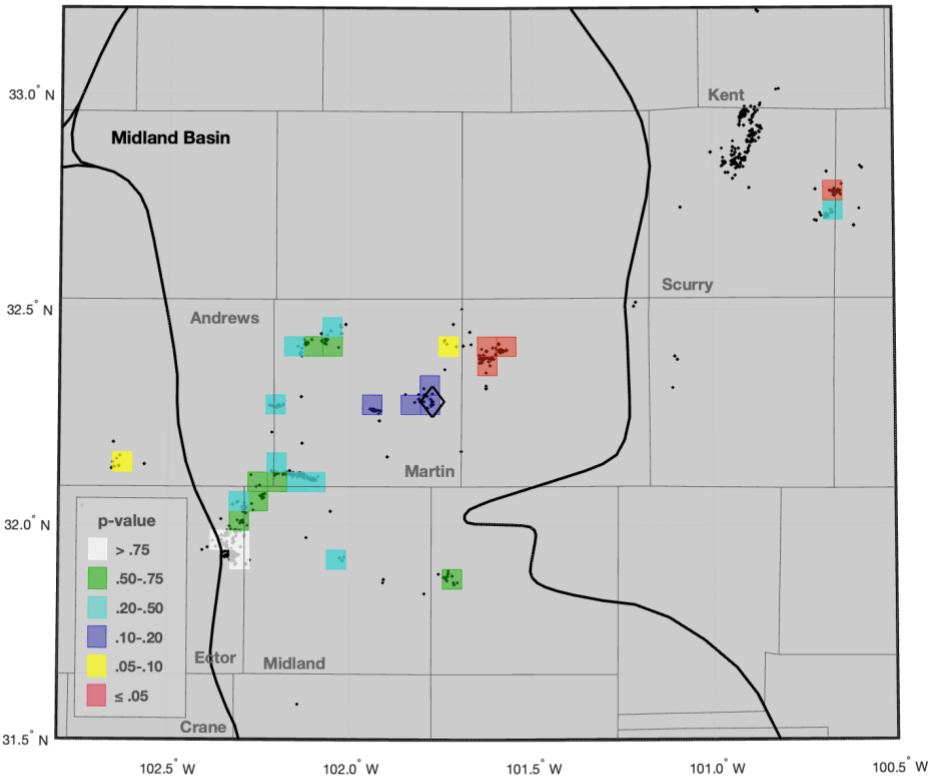

Figure 54. Statistical p-values for HF between 2017 and September 2022, with all declustered earthquakes above M 2 overlapped. Only blocks with at least 3 events are color-coded. The black diamond indicates the largest earthquake in Stanton. Country borders are also mapped.

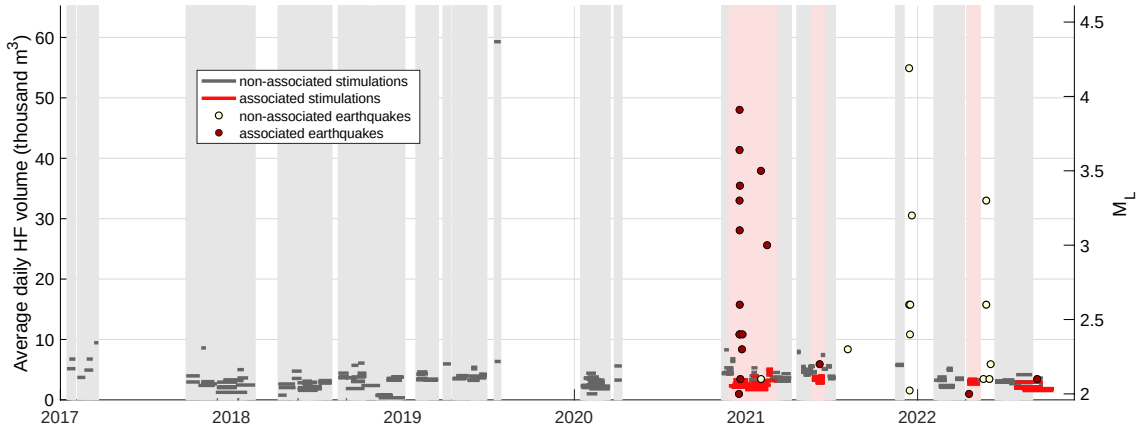

Figure 55. Time history of earthquake magnitudes and average daily HF volumes around the Stanton mainshock. The duration of each stimulation is denoted in light red color. Incomplete catalog before 2020.

4.6.2 Midland basin – SWD

Figures 56 and 57 shows that the cumulative disposed volumes across the basin (for shallow and deep SWD). There are very low and sparse volumes in the Gardendale SRA.

Figures 58 and 59 shows the statistical p-values for shallow and deep SWD in the region, for the period between 2017 and 2022, for $M \geq 2$ and the large-scale diffusivity value D set to 1 m^2/s.

Shallow SWD appears to be the sole causal factor in north Snyder, despite no such wells being close present in the oil field; this implies far-field effects that seem unlikely given the low disposal volumes there. Both deep and shallow SWD seem like a triggering factor also in east Snyder, together with HF; this would be a unique case where all three oil and gas activities contribute at the same location. The focal depths do span both shallow and deeper intervals. The Stanton SRA seems affected by both shallow and deep SWD, in contrast to the reported focal depths which point mainly at basement reactivation. The Gardendale SRA seems affected mostly by deep SWD, with some shallow SWD effects on the east end of the zone. The reported focal depths point mainly at basement reactivation, but include more shallow hypocenters than the Stanton SRA. That said, those shallow hypocenters are mostly on the west side of the zone, not the east.

Finally, the declustering algorithm removed only 19% of the events in the basin, a similar number to the Delaware basin (likely for the same reasons). This low percentage of aftershocks would indicate that deep SWD is not that dominant of a causal factor in the Midland basin, contrary to our p-value analyses. We have more faith in the p-values however.

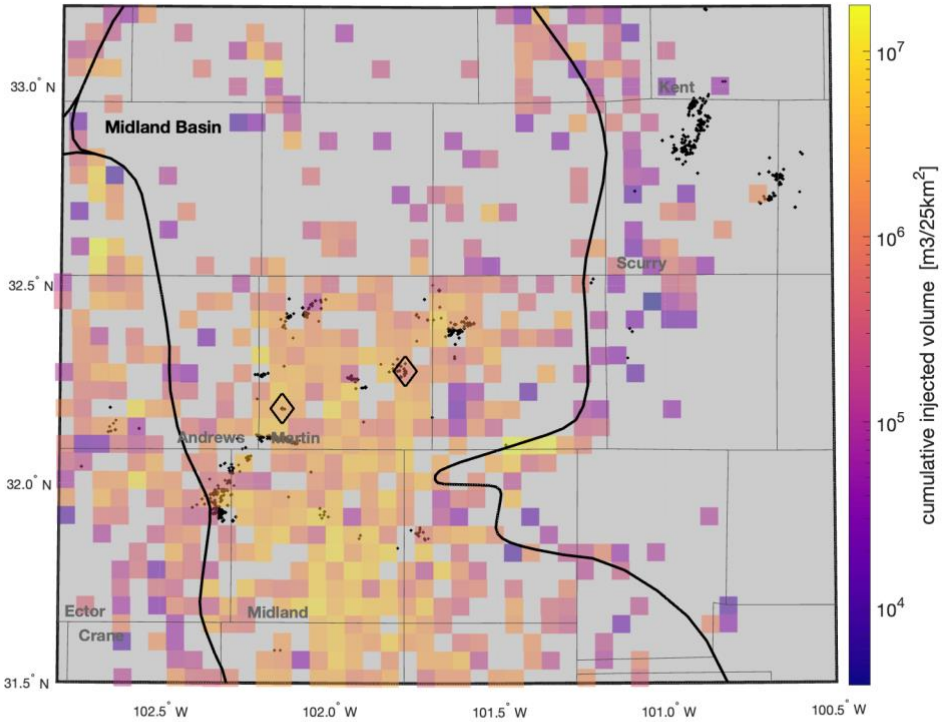

Figure 56. Map of the investigated area showing declustered seismicity (M ≥ 2; black dots), gridded cumulative shallow SWD volumes and county borders. Data between 2006 and 2022. The earthquake data are incomplete before 2017, and partly incomplete between 2017 and 2020. The black diamonds indicate the two largest earthquakes.

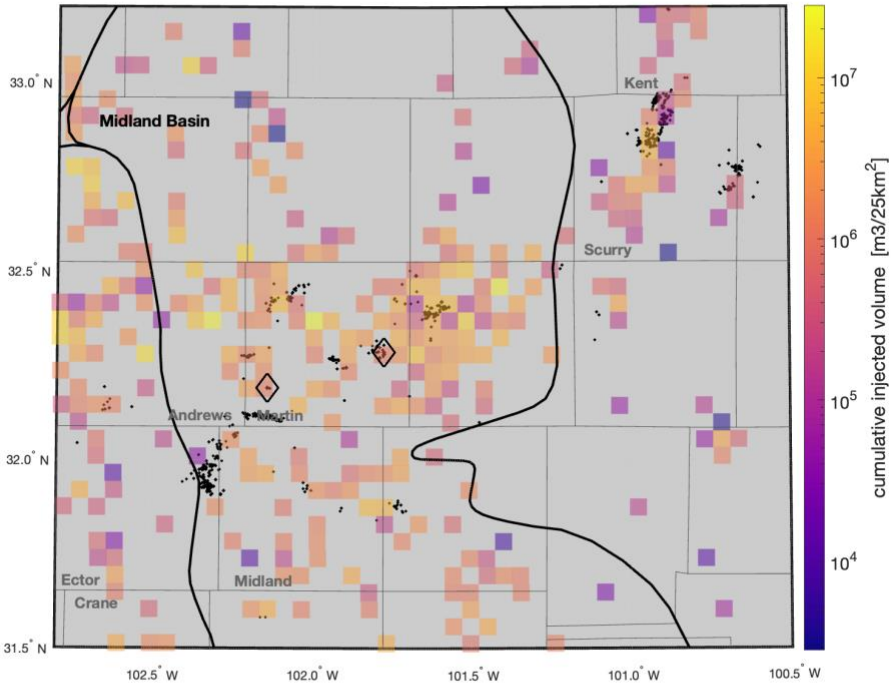

Figure 57. Map of the investigated area showing declustered seismicity (M ≥ 2; black dots), gridded cumulative deep SWD volumes and county borders. Data between 2006 and 2022. The earthquake data are incomplete before 2017, and partly incomplete between 2017 and 2020. The black diamonds indicate the two largest earthquakes.

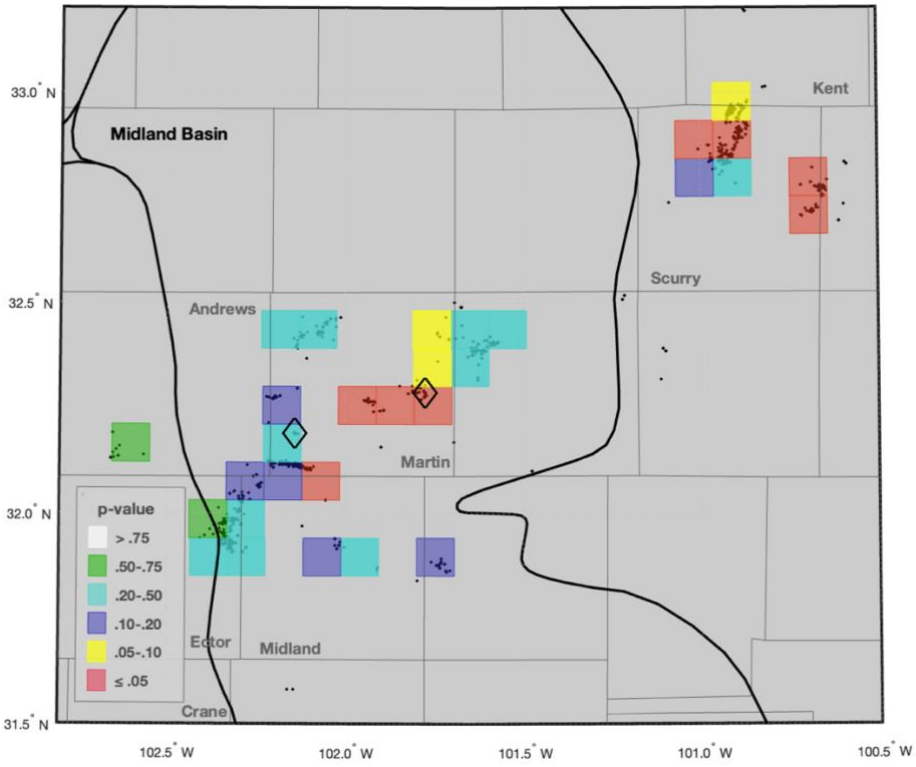

Figure 58. Statistical p-values for shallow SWD between 2017 and 2022, for D equal to 1 m^2/s, with all declustered earthquakes above M 2 overlapped. Only blocks with at least 3 events are color-coded. The black diamonds indicate the two largest earthquakes. Country borders are also mapped.

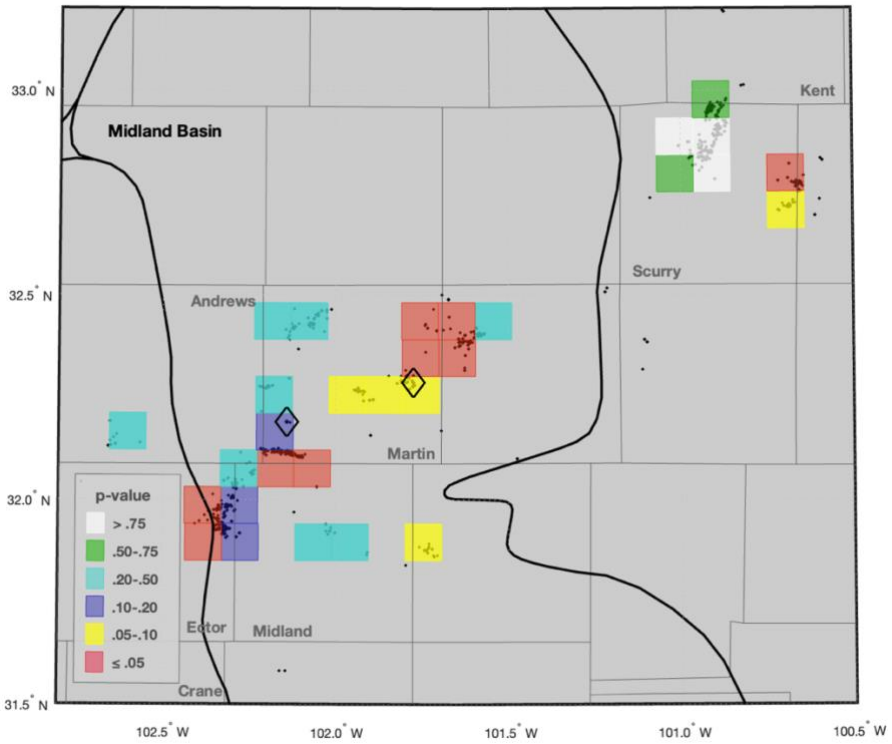

Figure 59. Statistical p-values for deep SWD between 2017 and 2022, for D equal to 1 m^2/s, with all declustered earthquakes above M 2 overlapped. Only blocks with at least 3 events are color-coded. The black diamonds indicate the two largest earthquakes. Country borders are also mapped.

4.6.3 Delaware basin - HF

The gridded cumulative HF volumes are shown in Figure 60. Although a large portion of the seismicity coincides with the presence of HF stimulations (Reeves, Ward County), a lot of earthquakes in Culberson County are very far from HF wells. We should also note that the limited seismicity on the state-border despite the presence of intense HF operations.

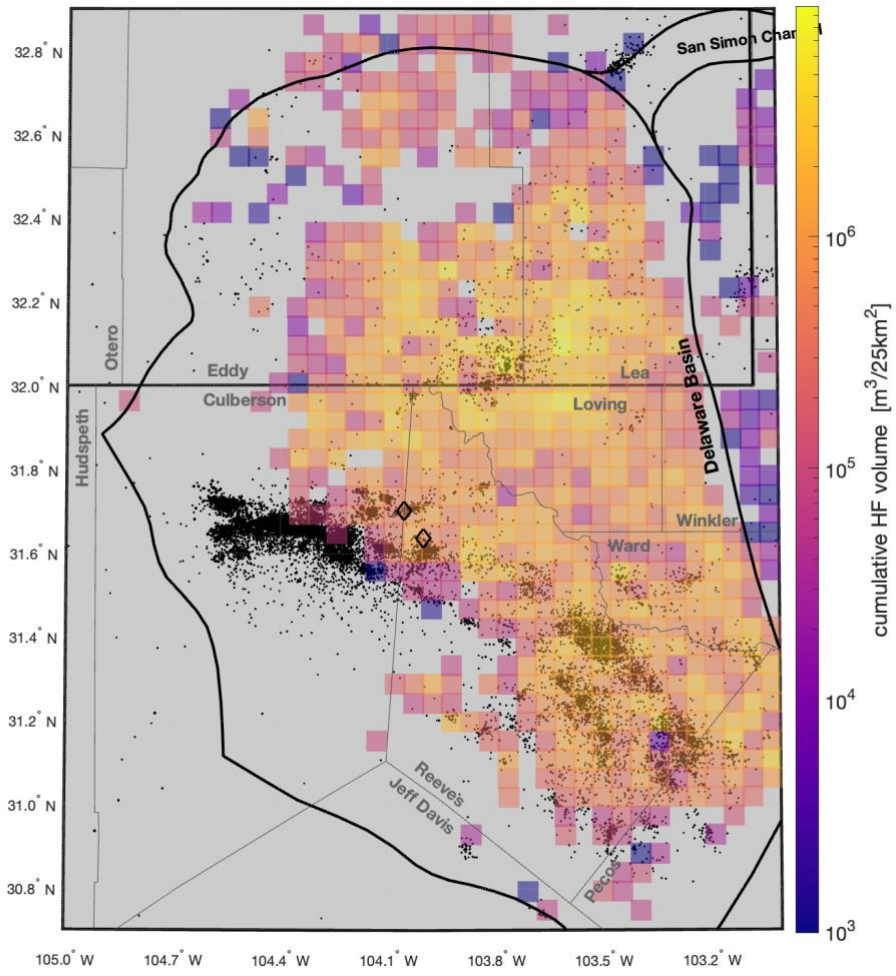

Figure 60. Map of the Delaware basin showing declustered seismicity (M ≥ 1.5; black dots), gridded cumulative HF volumes and county borders. The black diamonds indicate the two M_w 5+ earthquakes. Data between 2017 and September 2022.

Figure 61 shows HFR values for the region, for the period between 2017 and September 2022, and for declustered earthquakes with magnitudes above 1.5. They paint the same picture as the p_{HF} values we discuss below. According to our association-criteria (section 3.4), 13% of declustered earthquakes are associated with 14% of the HF stimulations. The latter is a very high number compared to other basins, and indicates coincidental linkages. Remarkably, 94% of HF stimulations that eventually induced seismicity according to our criteria, started doing so during the stimulation period.

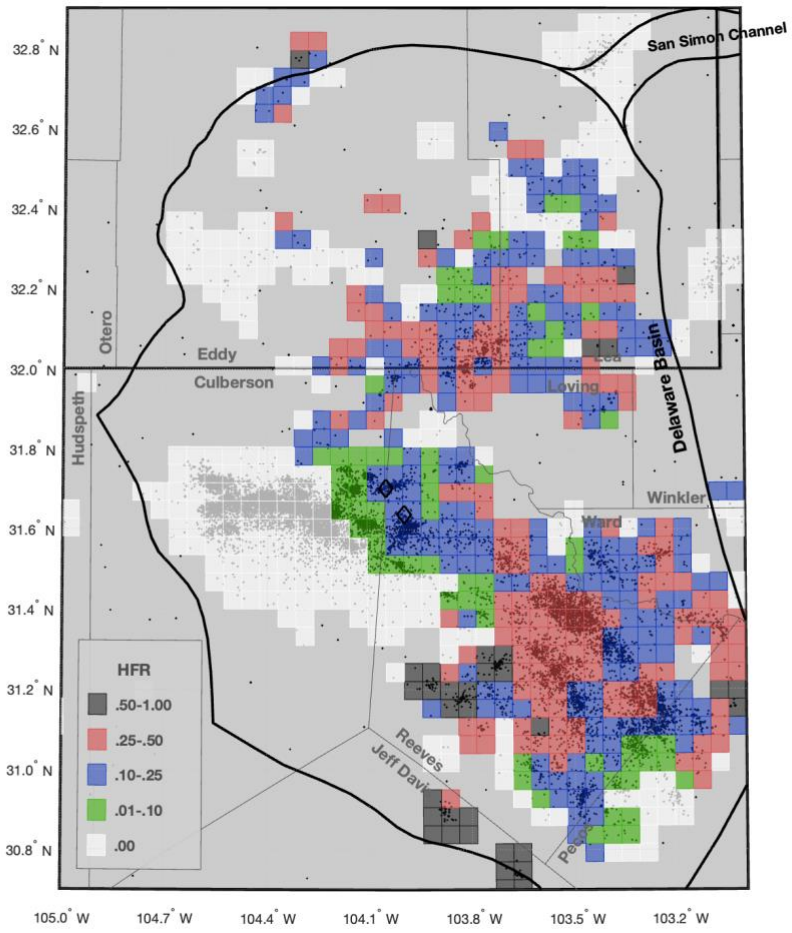

Figure 61. HFR values mapped on a 5 km grid, with all declustered earthquakes between 2017 and September 2022 above M 1.5 overlapped. The black diamonds indicate the two M_w 5+ earthquakes. Only blocks with at least 3 events are color-coded. Country borders are also mapped.

Figure 62 shows the statistical p-values for HF (p_{HF}) in the region, for the period between 2017 and September 2022. In particular, 12% of blocks have $p_{HF} \leq 0.05$, and 18% of blocks have $p_{HF} \leq 0.10$. Furthermore, only 10% of earthquakes occurred within a block with $p_{HF} \leq 0.05$. When we combine SAF with p_{HF}, only 5% of declustered earthquakes above M 1.5 are both linked to HF and occurred within a block with $p_{HF} \leq 0.05$. These are very similar numbers with the Midland basin. As far as the HF wells are concerned, 5% of stimulations are both linked to seismicity and occurred within a block with $p_{HF} \leq 0.05$. On the other hand, 97% of stimulations are not linked to seismicity and occurred within a block with $p_{HF} \geq 0.05$. Lastly, only 61% of blocks have a fitted time-lag (equation 4) of 3 days or less and the median Σ value among blocks with $p_{HF} \leq 0.05$ is -3.4.

Spatially, the observations drawn from the p_{HF} values is essentially identical to the ones made by Grigoratos et al. (2022). We will not repeat them here. We will only add that the seismicity levels in the state-border have increased lately and HF seems to have played a role in that.

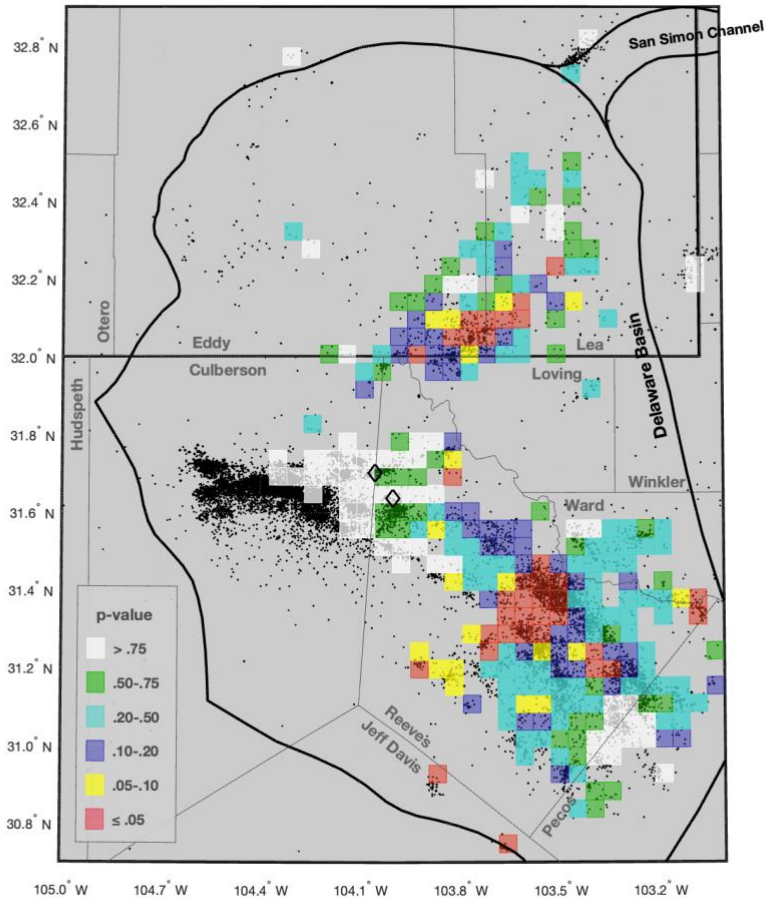

Figure 62. Statistical p-values for HF between 2017 and September 2022, with all declustered earthquakes above M 1.5 overlapped. Only blocks with at least 3 events are color-coded. The black diamonds indicate the two M_w 5+ earthquakes. Country borders are also mapped.

4.6.4 Delaware basin – deep SWD

Figure 63 shows that the cumulative deep SWD volumes disposed across the basin. It is notable that almost all earthquakes occur in blocks with no deep SWD volumes (before diffusion). Figure 64 shows that the same volumes after diffusion in space and time and how the volumes get distributed closer to the seismicity, especially in Culberson county.

Figure 65 shows the statistical p-values for deep SWD (p_{SWD}) in the region, for the period between 2017 and 2022, for M ≥ 1.5 and D set to 1 m^2/s. SWD appears to be the sole causal factor in the entire Culberson county, without requiring complex earthquake-to-earthquake interactions; direct pore-pressure diffusion suffices. This is a novel finding. The seismicity in Reeves and Ward counties, which are too far from the deep wells anyway, do not seem correlated with deep SWD rates. That said, the Culberson clusters creep slightly in the northern edge of Reeves county and there deep SWD is still a causal factor. Deep SWD overlaps as a causal factor with HF in some parts of northern Delaware basin, although the effects of deep SWD are more broad spatially. Finally, deep SWD p-values for D equal to 2 m^2/s are very similar. For D equal to 0.3 m^2/s, they are also similar, but significantly higher at west side of Culberson county, indicating that large diffusivity values are required to explain such far-field effects from the distant deep SWD wells, when ignoring poroelasticity.

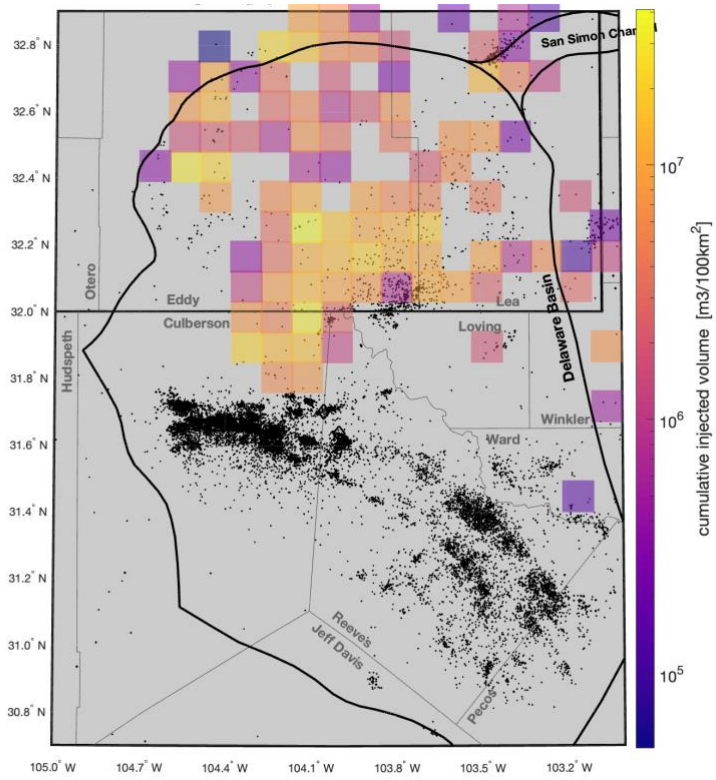

Figure 63. Map of the investigated area showing declustered seismicity (M ≥ 1.5; black dots), gridded cumulative SWD volumes and county borders. Data only between 2006 and 2022. The earthquake data are incomplete before 2017. The black diamonds indicate the two M_w 5+ earthquakes.

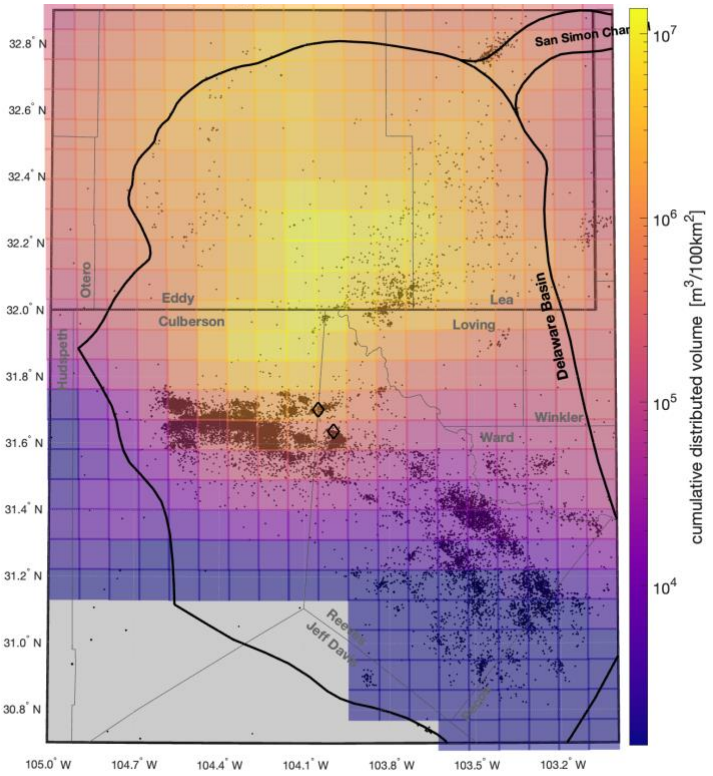

Figure 64. Map of the investigated area showing declustered seismicity (M ≥ 1.5; black dots), gridded diffused cumulative SWD volumes for D = 1 m²/s and county borders. Data only between 2006 and 2022. The earthquake data are incomplete before 2017. The black diamonds indicate the two M_w 5+ earthquakes.

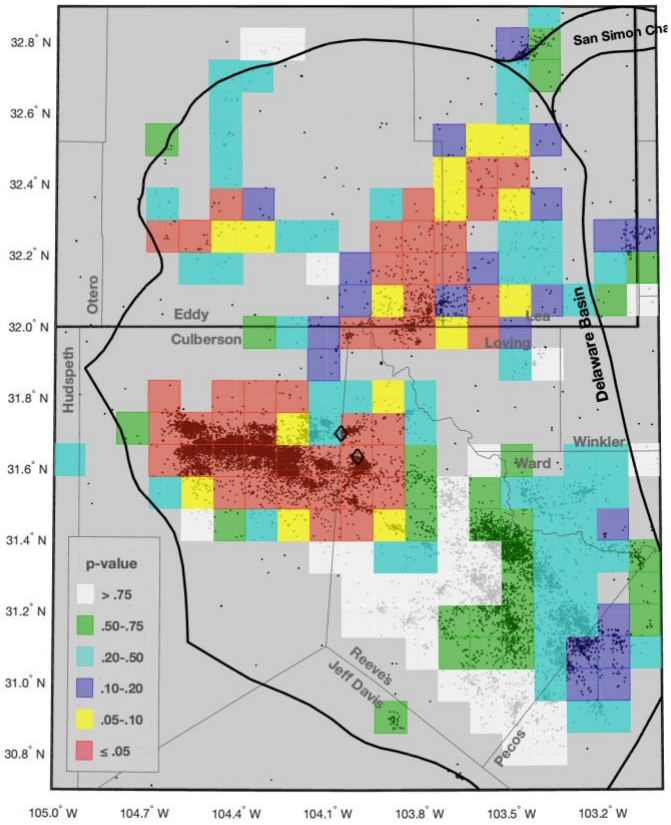

Figure 65. Statistical p-values for SWD between 2017 and 2021, for D equal to 1 m^2/s, with all declustered earthquakes above M 1.5 overlapped. Only blocks with at least 3 events are color-coded. The black diamonds indicate the two M_w 5+ earthquakes. Country borders are also mapped.

4. References

Aden-Antoniów, F., Frank, W. B., & Seydoux, L. (2022). An Adaptable Random Forest Model for the Declustering of Earthquake Catalogs. Journal of Geophysical Research: Solid Earth, 127(2). https://doi.org/10.1029/2021JB023254

Ansari, E., and T. S. Bidgoli (2020). Reply to comment by Peterie et al. on "accelerated fill-up of the arbuckle group aquifer and links to US midcontinent seismicity", J. Geophys. Res. 125, no. 12, doi: 10.1029/2019JB019275.

Ansari, E., T. S. Bidgoli, and A. Hollenbach (2019). Accelerated fill-up of the Arbuckle Group aquifer and links to US midcontinent seis- micity, J. Geophys. Res. 124, no. 3, 2670–2683.

Brown, M. R., S. Ge, A. F. Sheehan, and J. S. Nakai (2017). Evaluating the effectiveness of induced seismicity mitigation: Numerical modeling of wastewater injection near Greeley, Colorado, J. Geophys. Res. 122, no. 8, doi: 10.1002/2017JB014456.

Bao, X., and D. W. Eaton (2016). Fault activation by hydraulic fracturing in western Canada, Science 354, no. 6318, 1406–1409.

Barbour, A. J., Norbeck, J. H., & Rubinstein, J. L. (2017). The Effects of Varying Injection Rates in Osage County, Oklahoma, on the 2016 Mw 5.8 Pawnee Earthquake. Seismological Research Letters, 88(4), 1040–1053. https://doi.org/10.1785/0220170003

Brudzinski, M. R., and M. Kozlowska (2019). Seismicity induced by hydraulic fracturing and wastewater disposal in the Appalachian Basin, USA: A review, Acta Geophys. 67, no. 1, 351–364.

Chang, K. W., and H. Yoon (2020). Hydromechanical controls on the spatiotemporal patterns of injection-induced seismicity in different fault architecture: Implication for 2013–2014 Azle earthquakes, J. Geophys. Res. 125, no. 9, doi: 10.1029/2020JB020402.

ComCat. U.S. Geological Survey, Earthquake Hazards Program, 2017, Advanced National Seismic System (ANSS) Comprehensive Catalog of Earthquake Events and Products: Various, https://doi.org/10.5066/F7MS3QZH.

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

Davis, S. D., W. D. Pennington, and S. M. Carlson (1985). Historical seismicity of the state of Texas—A Summary, Guff Coast Assoc. Geol. Soc. Trans. 35, 39–44.

DeShon, H. R., Hayward, C. T., Ogwari, P. O., Quinones, L., Sufri, O., Stump, B., & Beatrice Magnani, M. (2019). Summary of the North Texas Earthquake Study Seismic Networks, 2013–2018. Seismological Research Letters, 90(1), 387–394. https://doi.org/10.1785/0220180269

Dinske, C., and Shapiro, S. (2016). Performance test of the Seismogenic index model for forecasting magnitude distributions of fluid-injection-induced seismicity, SEG Technical Program Expanded Abstracts 2016, doi: 10.1190/segam2016-13958126.1.

Doser, D. I., M. R. Baker, M. Luo, P. Marroquin, L. Ballesteros, J. Kingwell, H. L. Diaz, and G. Kaip (1992). The not so simple rela- tionship between seismicity and oil production in the Permian Basin, west Texas, Pure Appl. Geophys. 139, no. 3, 481–506.

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

Ewing, T. E. (1990). Tectonic map of Texas. Bureau of Economic Geology, University of Texas at Austin.

Fasola, S. L., Brudzinski, M. R., Skoumal, R. J., Langenkamp, T., Currie, B. S., & Smart, K. J. (2019). Hydraulic Fracture Injection Strategy Influences the Probability of Earthquakes in the Eagle Ford Shale Play of South Texas. Geophysical Research Letters, 46(22), 12958–12967. https://doi.org/10.1029/2019GL085167

Fasola, S. L., & Brudzinski, M. R. (2023). Machine Learning Reveals Additional Hydraulic Fracture Induced Seismicity in the Eagle Ford Shale. Journal of Geophysical Research: Solid Earth. https://doi.org/10.1029/2022JB025436

Frohlich, C., Potter, E., Hayward, C., & Stump, B. (2010). Dallas-Fort Worth earthquakes coincident with activity associated with natural gas production. The Leading Edge, 29(3), 270–275. https://doi.org/10.1190/1.3353720

Frohlich, C., Hayward, C., Stump, B., & Potter, E. (2011). The Dallas-Fort Worth Earthquake Sequence: October 2008 through May 2009. Bulletin of the Seismological Society of America, 101(1), 327–340. https://doi.org/10.1785/0120100131

Frohlich, C. (2012). Two-year survey comparing earthquake activity and injection-well locations in the Barnett Shale, Texas. Proceedings of the National Academy of Sciences, 109(35), 13934–13938. https://doi.org/10.1073/pnas.1207728109

Frohlich, C., W. Ellsworth, W. A. Brown, M. Brunt, J. Luetgert, T. MacDonald, and S. Walter (2014). The 17 May 2012 M 4.8 earth- quake near Timpson, East Texas: An event possibly triggered by fluid injection, J. Geophys. Res. 119, no. 1, 581–593.

Frohlich, C., DeShon, H., Stump, B., Hayward, C., Hornbach, M., and Walter, J.I. (2016). A historical review of induced earthquakes in Texas, *Seismol. Res. Lett.* 87(4), 1022-1038.

Frohlich, C., Hayward, C., Rosenblit, J., Aiken, C., Hennings, P., Savvaidis, A., Lemons, C., Horne, E., Walter, J. I., & DeShon, H. R. (2020). Onset and Cause of Increased Seismic Activity Near Pecos, West Texas, United States, From Observations at the Lajitas TXAR Seismic Array. Journal of Geophysical Research: Solid Earth, 125(1). https://doi.org/10.1029/2019JB017737

Gardner, J.K., and Knopoff, L. (1974). Is the sequence of earthquakes in Southern California, with aftershocks removed, Poissonian?, *Bull. Seismol. Soc. Am.* 64(5), 1363-1367.

Glasgow, M. E., Schmandt, B., Wang, R., Zhang, M., Bilek, S. L., & Kiser, E. (2021). Raton Basin induced seismicity is hosted by networks of short basement faults and mimics tectonic earthquake statistics. https://doi.org/10.1002/essoar.10507748.1

Glasgow, M. E., Schmandt, B., & Bilek, S. L. (2023). Cascading multi-segment rupture in an injection-induced earthquake sequence with a Mw 5.3 mainshock. Earth and Planetary Science Letters, 620, 118335. https://doi.org/10.1016/j.epsl.2023.118335

Grigoratos, I., Rathje, E., Bazzurro, P., & Savvaidis, A. (2020a). Earthquakes Induced by Wastewater Injection, Part I: Model Development and Hindcasting. Bulletin of the Seismological Society of America, 110(5), 2466–2482. https://doi.org/10.1785/0120200078

Grigoratos, I., Rathje, E., Bazzurro, P., & Savvaidis, A. (2020b). Earthquakes Induced by Wastewater Injection, Part II: Statistical Evaluation of Causal Factors and Seismicity Rate Forecasting. Bulletin of the Seismological Society of America, 110(5), 2483–2497. https://doi.org/10.1785/0120200079

Grigoratos, I., Savvaidis, A., & Rathje, E. (2022). Distinguishing the Causal Factors of Induced Seismicity in the Delaware Basin: Hydraulic Fracturing or Wastewater Disposal? Seismological Research Letters, 93(5), 2640–2658. https://doi.org/10.1785/0220210320

Grigoratos I., Poggi, V., Danciu, L., Monteiro, R. (2023). Homogenizing instrumental earthquake catalogs – a case study around the Dead Sea Transform Fault Zone. *Seismica* (under review).

Gutenberg, B., and Richter, C.F. (1956). Earthquake magnitude, intensity, energy, and acceleration: (Second paper), *Bull. Seismol. Soc. Am.* 46(2), 105-145.

Hennings, P. H., Dvory, N., Horne, E. A., Li, P., Savvaidis, A., & Zoback, M. (2021) Stability of the Fault Systems that Host Induced Earthquakes in the Delaware Basin of West Texas and Southeast New Mexico, The Seismic Record (accepted).

Hennings, P. H., Nicot, J. P., Gao, R. S., DeShon, H. R., Snee, J.-E. L., Morris, A. P., Brudzinski, R., Horne, E. A., & Breton, C. (2020). Pore Pressure Threshold and Fault Slip Potential for Induced Earthquakes in the Dallas-Fort Worth Area of North Central Texas. 23.

Herrmann, R.B., Benz, H. and Ammon, C.J. (2011). Monitoring the earthquake source process in North America, Bull. Seismol. Soc. Am., 101, 2609–2625, doi:10.1785/0120110095

Holland, A. A. (2013). Earthquakes Triggered by Hydraulic Fracturing in South-Central Oklahoma. Bulletin of the Seismological Society of America, 103(3), 1784–1792. https://doi.org/10.1785/0120120109

Homman, K., & Nyblade, A. (2020). The Pennsylvania State Seismic Network: Report on seismic event location uncertainties.

Hornbach, M. J., DeShon, H. R., Ellsworth, W. L., Stump, B. W., Hayward, C., Frohlich, C., Oldham, H. R., Olson, J. E., Magnani, M. B., Brokaw, C., & Luetgert, J. H. (2015). Causal factors for seismicity near Azle, Texas. Nature Communications, 6(1). https://doi.org/10.1038/ncomms7728

Hornbach, M. J., Jones, M., Scales, M., DeShon, H. R., Magnani, M. B., Frohlich, C., Stump, B., Hayward, C., & Layton, M. (2016). Ellenburger wastewater injection and seismicity in North Texas. Physics of the Earth and Planetary Interiors, 261, 54–68. https://doi.org/10.1016/j.pepi.2016.06.012

Horton, S. (2012). Disposal of Hydrofracking Waste Fluid by Injection into Subsurface Aquifers Triggers Earthquake Swarm in Central Arkansas with Potential for Damaging Earthquake. Seismological Research Letters, 83(2), 250–260. https://doi.org/10.1785/gssrl.83.2.250

Johnson, R. (1969). Geologic map of the Trinidad quadrangle, south-central Colorado, U.S. Geol. Surv. Misc. Geol. Investig. Map I-558.

Justinic, A. H., Stump, B., Hayward, C., & Frohlich, C. (2013). Analysis of the Cleburne, Texas, Earthquake Sequence from June 2009 to June 2010. Bulletin of the Seismological Society of America, 103(6), 3083–3093. https://doi.org/10.1785/0120120336

Kao, H., R. Visser, B. Smith, and S. Venables (2018). Performance assess- ment of the induced seismicity traffic light protocol for northeastern British Columbia and western Alberta, The Leading Edge 37, no. 2, 117–126.

Kavoura, F., Savvaidis, A., & Rathje, E. (2020). Determination of Local Magnitude for Earthquakes Recorded from the Texas Seismological Network (TexNet). Seismological Society of America, 91(6), 3223-3235.

Kettlety, T., J. P. Verdon, A. Butcher, M. Hampson, and L. Craddock (2021). High-resolution imaging of the ML 2.9 August 2019 earth- quake in Lancashire, United Kingdom, induced by hydraulic fracturing during Preston New Road PNR-2 operations, Seismol. Soc. Am. 92, no. 1, 151–169, doi: 10.1785/0220200187.

Kim, W. Y. (2013). Induced seismicity associated with fluid injection into a deep well in Youngstown, Ohio, J. Geophys. Res. 118, no. 7, 3506–3518.

Lee, M. K., and D. D. Williams (2000). Paleohydrology of the Delaware basin, western Texas: Overpressure development, hydro- carbon migration, and ore genesis, AAPG Bull. 84, no. 7, 961–974.

Lei, X., Z. Wang, and J. Su (2019). The December 2018 ML 5.7 and January 2019 ML 5.3 earthquakes in South Sichuan basin induced by shale gas hydraulic fracturing, Seismol. Res. Lett. 90, no. 3, 1099–1110.

Li, P., Huang, G.-C. D., Savvaidis, A., Kavoura, F., & Porritt, R. W. (2021). Characteristics of Seismicity in the Eagle Ford Shale Play, Southern Texas, Constrained by Earthquake Relocation and Centroid Moment Tensor Inversion. Seismological Research Letters. https://doi.org/10.1785/0220210005

Li, P., Savvaidis, A. (2021). TexNet earthquake hypoDD cross correlation relocation catalog of Permian Basin from 20170101 to 20210414. https://doi.org/10.18738/T8/RBEGTH.

McClure, M., R. Gibson, K. K. Chiu, and R. Ranganath (2017). Identifying potentially induced seismicity and assessing statistical significance in Oklahoma and California, J. Geophys. Res. 122, no. 3, 2153–2172.

McKeighan, C., P. Hennings, E. A. Horne, K. Smye, and A. Morris (2022). Understanding Anthropogenic Fault Rupture in the Eagle Ford Region, South-Central Texas, Bull. Seismol. Soc. Am. XX,1–20, doi: 10.1785/ 0120220074

Meremonte, M. E., Lahr, J. C., Frankel, A. D., Dewey, J. W., Crone, A. J., Overturf, D. E., et al. (2002). Investigation of an earthquake swarm near Trinidad, Colorado, August–October 2001. U.S Geol. Surv. Open-File Rept. 02-0073.

Mueller, C.S. (2018). Earthquake catalogs for the USGS National Seismic Hazard Maps. *Seismological Research Letters*, 90(1), 251-261.

Murray, K. (2015). Class II Saltwater Disposal for 2009–2014 at the Annual-, State-, and County- Scales by Geologic Zones of Completion, Oklahoma. Open-File Report: OF5-2015, Oklahoma Geological Survey, doi: 10.13140/RG.2.1.4841.7364.

Nakai, J. S., Sheehan, A. F., & Bilek, S. L. (2017a). Seismicity of the rocky mountains and Rio Grande Rift from the EarthScope Transportable Array and CREST temporary seismic networks, 2008-2010: Seismicity of the Rio Grande Rift. Journal of Geophysical Research: Solid Earth. https://doi.org/10.1002/2016JB013389

Nakai, J. S., Weingarten, M., Sheehan, A. F., Bilek, S. L., & Ge, S. (2017b). A Possible Causative Mechanism of Raton Basin, New Mexico and Colorado Earthquakes Using Recent Seismicity Patterns and Pore Pressure Modeling. Journal of Geophysical Research: Solid Earth, 122(10), 8051–8065. https://doi.org/10.1002/2017JB014415

Norbeck, J.H., and Rubinstein, J.L. (2018). Hydromechanical earthquake nucleation model forecasts onset, peak, and falling rates of induced seismicity in Oklahoma and Kansas, *Geophys Res Lett.* 45(7), 2963-2975.

Ogwari, P. O., DeShon, H. R., & Hornbach, M. J. (2018). The Dallas-Fort Worth Airport Earthquake Sequence: Seismicity Beyond Injection Period. Journal of Geophysical Research: Solid Earth, 123(1), 553–563. https://doi.org/10.1002/2017JB015003

Park, Y., S. M. Mousavi, W. Zhu, W. L. Ellsworth, and G. C. Beroza (2020). Machine-learning-based analysis of the Guy-Greenbrier, Arkansas earthquakes: A tale of two sequences, Geophys. Res. Lett. 47, no. 6.

Parotidis, M., & Shapiro, S. A. (2004). A statistical model for the seismicity rate of fluid-injection-induced earthquakes. Geophys. Res. Lett. 31(17).

Petersen, M. D., Mueller, C. S., Haller, K. M., Moschetti, M., Harmsen, S. C., Field, E. H., ... & Boyd, O. S. (2012). 2014 Update of the United States National Seismic Hazard Maps.

Pursley, J., Bilek, & Ruhl. (2013). Earthquake catalogs for New Mexico and bordering areas: 2005–2009. 35(1), 10. Nyblade, A., & Homman, K. (2016). Seismicity in Pennsylvania from February 2013 to June 2015, Report Submitted to the State Geologist and Director of the Bureau of Topographic and Geologic Survey, DCNR.

Quinones, L., H. R. DeShon, S. Jeong, P. Ogwari, O. Sufri, M. M. Holt, and K. B. Kwong (2019). Tracking induced seismicity in the Fort Worth basin: A summary of the 2008–2018 North Texas earthquake study catalog, Bull. Seismol. Soc. Am. 109, no. 4, 1203–1216.

Roth, M. P., A. Verdecchia, R. M. Harrington, and Y. Liu (2020). High-resolution imaging of hydraulic-fracturing-induced earth- quake clusters in the Dawson-Septimus area, Northeast British Columbia, Canada, Seismol. Res. Lett. 91, no. 5, 2744–2756.

Rubinstein, J. L., Ellsworth, W. L., McGarr, A., & Benz, H. M. (2014). The 2001-Present Induced Earthquake Sequence in the Raton Basin of Northern New Mexico and Southern Colorado. Bulletin of the Seismological Society of America, 104(5), 2162–2181. https://doi.org/10.1785/0120140009

Rubinstein, J. L., Ellsworth, W. L., & Dougherty, S. L. (2018). The 2013–2016 Induced Earthquakes in Harper and Sumner Counties, Southern Kansas. Bulletin of the Seismological Society of America, 108(2), 674–689. https://doi.org/10.1785/0120170209

Savvaidis, A., Young, B., Huang, G. C. D., & Lomax, A. (2019). TexNet: A statewide seismological network in Texas. Seismological Research Letters, 90(4), 1702-1715.

Savvaidis, A., Lomax, A., & Breton, C. (2020). Induced Seismicity in the Delaware Basin, West Texas, is Caused by Hydraulic Fracturing and Wastewater Disposal. Bulletin of the Seismological Society of America, 110(5), 2225-2241.

Scales, M. M., DeShon, H. R., Magnani, M. B., Walter, J. I., Quinones, L., Pratt, T. L., & Hornbach, M. J. (2017). A Decade of Induced Slip on the Causative Fault of the 2015 Mw 4.0 Venus Earthquake, Northeast Johnson County, Texas: Induced Seismicity in Johnson County, TX. Journal of Geophysical Research: Solid Earth, 122(10), 7879–7894. https://doi.org/10.1002/2017JB014460

Schoenball, M., and Ellsworth, W.L. (2017). A systematic assessment of the spatiotemporal evolution of fault activation through induced seismicity in Oklahoma and southern Kansas, *J. Geophys. Res. Solid Earth.* 122(12).

Shapiro, S.A. (2015). *Fluid-induced seismicity*, Cambridge University Press, doi: 10.1017/CBO9781139051132.

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

Schultz, R., V. Stern, M. Novakovic, G. Atkinson, and Y. J. Gu (2015). Hydraulic fracturing and the Crooked Lake sequences: Insights gleaned from regional seismic networks, Geophys. Res. Lett. 42, no. 8, 2750–2758.

Schultz, R., G. Atkinson, D. W. Eaton, Y. J. Gu, and H. Kao (2018). Hydraulic fracturing volume is associated with induced earthquake productivity in the Duvernay play, Science 359, no. 6373, 304–308.

Schultz, R., and R. Wang (2020). Newly emerging cases of hydraulic fracturing induced seismicity in the Duvernay East Shale Basin, Tectonophysics 779, doi: 10.1016/j.tecto.2020.228393.

Schultz, R., R. J. Skoumal, M. R. Brudzinski, D. Eaton, B. Baptie, and W. Ellsworth (2020). Hydraulic fracturing-induced seismicity, Rev. Geophys. 58, no. 3, doi: 10.1029/2019RG000695.

Sheng, Y., K. S. Pepin, and W. L. Ellsworth (2022). On the depth of earthquakes in the Delaware Basin: A case study along the Reeves– Pecos county line, Seismol. Rec. 2, no. 1, 29–37.

Skoumal, R. J., R. Ries, M. R. Brudzinski, A. J. Barbour, and B. S. Currie (2018). Earthquakes induced by hydraulic fracturing are pervasive in Oklahoma, J. Geophys. Res. 123, no. 12, 10,918–10,935.

Skoumal, R. J., Kaven, J. O., & Walter, J. I. (2019). Characterizing Seismogenic Fault Structures in Oklahoma Using a Relocated Template-Matched Catalog. Seismological Research Letters. https://doi.org/10.1785/0220190045

Skoumal, R. J., A. J. Barbour, M. R. Brudzinski, T. Langenkamp, and J. O. Kaven (2020a). Induced seismicity in the Delaware Basin, Texas, J. Geophys. Res. 125, no. 1, doi: 10.1029/2019JB018558.

Skoumal, R. J., Kaven, J. O., Barbour, A. J., Wicks, C., Brudzinski, M. R., Cochran, E. S., & Rubinstein, J. L (2020b). The Induced Mw 5.0 March 2020 West Texas Seismic Sequence. Journal of Geophysical Research: Solid Earth.

Smye, K, Banerji DA, Eastwood R, McDaid G, Hennings P (2021). Lithology and reservoir properties of the Delaware Mountain Group of the Delaware Basin and implications for saltwater disposal and induced seismicity. Journal of Sedimentary Research, 91 (11): 1113–1132.

Theis, C.V. (1935). The relation between the lowering of the piezometric surface and the rate and duration of discharge of a well using ground-water storage, *Eos Trans. AGU,* 16(2), 519-524.

Tung, S., G. Zhai, and M. Shirzaei (2020). Potential link between 2020 Mentone, West Texas M5 earthquake and nearby wastewater injection: Implications for aquifer mechanical properties, Geophys. Res. Lett. 48, no. 3, doi: 10.1029/2020GL090551.

Verdon, J. P., J. M. Kendall, A. C. Horleston, and A. L. Stork (2016). Subsurface fluid injection and induced seismicity in southeast Saskatchewan, Int. J. Greenhouse Gas Control 54, 429–440.

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

Walter, J. I., Frohlich, C., & Borgfeldt, T. (2018). Natural and Induced Seismicity in the Texas and Oklahoma Panhandles. Seismological Research Letters, 89(6), 2437–2446. https://doi.org/10.1785/0220180105

Walter, J. I., P. Ogwari, A. Thiel, F. Ferrer, I. Woelfel, J. C. Chang, A. P. Darold, and A. A. Holland (2020), The Oklahoma Geological Survey Statewide Seismic Network, Seismol. Res. Lett., 91 (2A): 611–621, doi:10.1785/0220190211.

Wang, Z., X. Lei, S. Ma, X. Wang, and Y. Wan (2020). Induced earthquakes before and after cessation of long-term injections in Rongchang gas field, Geophys. Res. Lett. 47, no. 22, doi: 10.1029/ 2020GL089569.

Weichert, D. H. (1980). Estimation of the earthquake recurrence parameters for unequal observation periods for different magnitudes. Bulletin of the Seismological Society of America, 70(4), 1337-1346.

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

Woo, J.-U., & Ellsworth, W. L. (2023). Reactivation of Precambrian Faults by Deep Wastewater Injection in Midland Basin, Texas, and Performance Evaluation of Seismic Response Areas. Bulletin of the Seismological Society of America. https://doi.org/10.1785/0120230086

Yang, H., P. Zhou, N. Fang, G. Zhu, W. Xu, J. Su, F. Meng, and R. Chu (2020). A shallow shock: The 25 February 2019 ML 4.9 earthquake in the Weiyuan shale gas field in Sichuan, China, Seismol. Soc. Am. 91, no. 6, 3182–3194.

Yoon, C. E., Y. Huang, W. L. Ellsworth, and G. C. Beroza (2017). Seismicity during the initial stages of the Guy-Greenbrier, Arkansas, earthquake sequence, J. Geophys. Res. 122, no. 11, 9253–9274.

Yeck, W. L., G. P. Hayes, D. E. McNamara, J. L. Rubinstein, W. D. Barnhart, P. S. Earle, and H. M. Benz (2017). Oklahoma experiences largest earthquake during ongoing regional wastewater injection hazard mitigation efforts, Geophys. Res. Lett. 44, no. 2, 711–717.

Yu, H., R. M. Harrington, Y. Liu, and B. Wang (2019). Induced seis- micity driven by fluid diffusion revealed by a near-field hydraulic stimulation monitoring array in the Montney Basin, British Columbia, J. Geophys. Res. 124, no. 5, 4694–4709.

Zaliapin, I., and Ben-Zion, Y. (2020). Earthquake declustering using the nearest-neighbor approach in space-time-magnitude domain. J. Geophys. Res. Solid Earth. 125(4).

Zaliapin, I., Gabrielov, A., Keilis-Borok, V., & Wong, H. (2008). Clustering Analysis of Seismicity and Aftershock Identification. Physical Review Letters, 101(1). https://doi.org/10.1103/PhysRevLett.101.018501

Zhai, G., & Shirzaei, M. (2018). Fluid Injection and Time-Dependent Seismic Hazard in the Barnett Shale, Texas. Geophysical Research Letters. https://doi.org/10.1029/2018GL077696

5. Acknowledgements and Disclaimer

This material is based upon work supported by the U.S. Geological Survey under Grant No. *G22AP00243* (FY2022) of the call *G22AS00006*. We are grateful to Alexandros Savvaidis (BEG) for sharing additional data and insights regarding the TexNet catalog, to Caroline Breton (BEG) for compiling the IHS Markit data, to Katie Smye (BEG) for classifying the Permian wells based on their target-formation, to Jun Ge (BEG) for discussions around his Permian basin pore-pressure models, to Justin Rubinstein (USGS) for sharing the catalog of the USGS network in NM, to Heather DeShon (SMU) for sharing an extended catalog of the FWB and to Margaret Glasgow (USGS) for her advice on how to use the online portals containing the Raton basin SWD data.

The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the opinions or policies of the U.S. Geological Survey. Mention of trade names or commercial products does not constitute their endorsement by the U.S. Geological Survey.

6. Project data, resources and bibliography

Sources of wastewater disposal data were the B3 database (https://www.b3insight.com/; last accessed December 11 2022), the Oklahoma Corporation Commission [OCC; https://oklahoma.gov/occ/divisions/oil-gas/oil-gas-data.html (last accessed March 2 2023); http://imaging.occeweb.com/imaging/UIC1012_1075.aspx (last accessed May 2019)], the KGS (https://www.kgs.ku.edu/Magellan/Qualified/class2_db.html; last accessed May 23 2023), the Colorado Oil & Gas Conservation Commission (https://cogcc.state.co.us/data.html#/cogis; last accessed March 1 2023), the Environmental Protection Agency (EPA; last update March 2018), Kyle Murray (OGS; last update May 2019) and the studies by Weingarten et al. (2015), Barbour et al. (2017) and Norbeck and Rubinsten (2018).

Sources of HF data were the FracFocus Chemical Disclosure Registry (last accessed February 19 2023), the IHS Markit databases (last accessed March 29 2023), and only for Oklahoma the OCC *FracNotices* (https://www.oklahoma.gov/content/dam/ok/en/occ/documents/og/isd_automated/All_Notices.csv; last accessed February 19 2023)

The following earthquake catalogs are available online: TexNet https://www.beg.utexas.edu/texnet-cisr/texnet/earthquake-catalog (last accessed February 18 2023), ANSS ComCat https://earthquake.usgs.gov/earthquakes/search/ (last accessed February 20 2023), OGS https://ogsweb.ou.edu/eq_catalog/ (last accessed February 19 2023), OhioDNR https://gis.ohiodnr.gov/MapViewer/?config=earthquakes# (last accessed August 26 2022), PASEIS (http://paseis.geosc.psu.edu/; last accessed August 2022), SLU (https://www.eas.slu.edu/eqc/eqcquakes.html; last accessed February 2023), KGS (https://www.kgs.ku.edu/Geophysics/Earthquakes/data.html; last accessed February 19 2023), NMT (https://geoinfo.nmt.edu/nmtso/events/home.cfml; last accessed February 20 2023). Publications that were used as sources of earthquake catalogs are listed in section 2.1.

The code used for the declustering is based on scripts that can be found at https://zenodo.org/record/5838353. The python code for the duplicate-removal process can be found at https://github.com/klunk386/CatalogueTool-Lite/tree/master/OQCatk. Almost all the figures were made using the Matlab software package (http://www.mathworks.com). The analyses presented in this paper greatly benefited from the high-performance computing resources available at the SED.

Supplemental material with input data and results are available as *csv* files from https://zenodo.org/ under the DOI: 10.0076/FY2022EHPg. They include earthquake catalogs (declustered and non-declustered), gridded monthly SWD data, and resulting gridded p-values. Each file-category is accompanied by its *readme* so that the end user has the appropriate documentation to understand how the data is structured. The repository also includes a full-sized *pdf* of the current report with high-resolution figures. The same data will be privately retained at least till 2030, as a back-up option.

Material resulting from the work performed under this award was part of an oral presentation titled "*Are higher HF injection rates more prone to triggering seismicity? Data from four North American basins say no.",* which was part of the 2023 *SSA Annual Meeting* in Puerto Rico.

7. Supplemental Material

Supplemental material with input data and results are available as *.csv* or *.mat* files from https://zenodo.org/record/8327414 under the DOI: 10.0076/FY2022EHPg. They include earthquake catalogs (declustered and non-declustered), gridded monthly SWD data, gridded daily HF data and resulting gridded p-values.

The following contains information about the margins adopted during the duplicate search, when creating the unified catalog. The TexNet catalog was used as the reference dataset on top of which all the rest are merged with. The other catalogs were added sequentially with time-dependent temporal and spatial margins. These margins are listed below in the following format: [first year, last year, time-window, spatial window (in km), units of time-window]. The magnitudes were not used in the duplicate-search.

```
# OGS
[[2000,2005,10.,40.,'Second'],
[2006,2024,3.1,25.,'Second']]
#--------------------------------------
# Rubinstein et al. (2018)
[[2000,2005,10.,40.,'Second'],
[2006,2024,6.1,30.,'Second']]
#--------------------------------------
# Schoenball & Ellsworth (2017)
[[2009,2024,6.1,30.,'Second']]
#--------------------------------------
# USGS network in NM
[[2000,2005,10.,40.,'Second'],
[2006,2024,3.1,20.,'Second']]
#--------------------------------------
# ComCat
[[2000,2010,6.1,35.,'Second'],
[2011,2024,5.1,30.,'Second']]
#--------------------------------------
# Ohio DNR
[[2000,2005,10.,40.,'Second'],
[2006,2010,6.1,30.,'Second'],
[2011,2024,5.1,30.,'Second']]
#--------------------------------------
# Pursley et al. (2013)
[[2000,2004,10.,40.,'Second'],
[2005,2024,6.1,30.,'Second']]
#--------------------------------------
# Holland (2013)
[[1900,2005,10.,40.,'Second'],
[2006,2024,4.1,30.,'Second']]
#--------------------------------------
# Frohlich et al. (2016)
[[1800,1899,6.,150.,'Month'],
[1900,1961,30.,100.,'Minute'],
[1962,2008,10.,35.,'Second'],
[2009,2024,7.1,30.,'Second']]
#--------------------------------------
# Walter et al. (2018)
[[2006,2024,7.1,30.,'Second']]
#--------------------------------------
# Gan & Frohlich (2013)
[[2006,2024,7.1,30.,'Second']]
#--------------------------------------
# Rubinstein et al. (2014)
[[1800,1899,6.,150.,'Month'],
[1900,1961,30.,100.,'Minute'],
[1962,2008,10.,35.,'Second'],
[2009,2024,7.1,25.,'Second']]
#--------------------------------------
# Nakai et al. (2017a)
[[2006,2024,7.1,30.,'Second']]
#--------------------------------------
# Nyblade & Homman (2016)
```

```
[[2010,2024,10.1,40.,'Second']]
#----------------------------------------
# PASEIS (>2016)
[[2010,2024,10.1,30.,'Second']]
#----------------------------------------
# AGS
[[1800,1899,6.,150.,'Month'],
[1900,1961,30.,100.,'Minute'],
[1962,2011,10.,35.,'Second'],
[2012,2024,7.1,30.,'Second']]
#----------------------------------------
# NMT
[[2010,2024,7.1,30.,'Second']]
#----------------------------------------
# SLU
[[2000,2005,10.,40.,'Second'],
[2005,2010,6.1,35.,'Second'],
[2011,2024,4.1,35.,'Second']]
#----------------------------------------
# TexNet Research
[[2010,2024,3.1,25.,'Second']]
#----------------------------------------
# NTXES (DeSchon)
[[2000,2005,10.,40.,'Second'],
[2006,2013,6.1,35.,'Second'],
[2014,2024,6.1,30.,'Second']]
#----------------------------------------
# Fasola & Brudzinski (2023)
[[2010,2024,7.1,20.,'Second']]
#----------------------------------------
# Skoumal et al. (2020)
[[2010,2024,6.1,20.,'Second']]
#----------------------------------------
# Skoumal et al. (2019)
[[2009,2024,6.1,25.,'Second']]
#----------------------------------------
# Wang et al. (2020)
[[2009,2024,4.1,25.,'Second']]
#----------------------------------------
# Mueller (2018)
[[1800,1899,6.,150.,'Month'],
[1900,1961,30.,100.,'Minute'],
[1962,1983,120,60.,'Second'],
[1984,2024,10.1,60.,'Second']]
#----------------------------------------
# KGS
[[2000,2013,12.1,40.,'Second'],
[2014,2015,12.1,30.,'Second'],
[2016,2024,12.1,25.,'Second']]
```