Lomax Anthony (Orcid ID: 0000-0002-7747-5990)
Savvaidis Alexandros (Orcid ID: 0000-0001-6373-5256)

Improving absolute earthquake location in west Texas using probabilistic, proxy ground-truth station corrections

Anthony Lomax[1], and Alexandros Savvaidis[2]

[1]ALomax Scientific, Mouans-Sartoux, France.

[2]Bureau of Economic Geology, The University of Texas at Austin, Austin, TX.

Corresponding authors: Anthony Lomax (anthony@alomax.net), Alexandros Savvaidis (alexandros.savvaidis@beg.utexas.edu)

Key Points:

- We show that some seismicity in the Delaware Basin, western Texas is likely caused by hydraulic fracturing.
- We use hydraulic-fracturing as proxy, ground-truth information to develop station corrections to improve absolute seismic event location.
- Spatial patterns and statistics of west Texas seismicity relocated with these corrections show possible improvement in absolute location accuracy.

Abstract

An increase in induced seismicity in the central U.S. since 2009 led to establishment of TexNet seismic-monitoring in Texas. Accurate, absolute seismic-event location is critical to TexNet, allowing quantitative evaluation of possible association of seismicity with human activity. For the Delaware Basin, western Texas, relocation using different velocity models and TexNet station subsets shows absolute location error up to 4 – 5 km. The preferred method to reduce absolute error, ground-truth calibration, is not available in this area. Alternatively, we used industrial well activity as proxy, ground-truth for developing probabilistic, proxy ground-truth (PPGT) station corrections for relocation. Assuming well activity causes seismicity, we defined a distance – time probability associating events and well activity. We used these associations and other evidence to show some seismicity in the Delaware Basin is more likely due to hydraulic-fracturing than salt-water disposal. We then probabilistically accumulated PPGT station corrections using event hypocenters constrained to associated fracturing-well locations. We applied this procedure within 12 km of TexNet station PB02, optimizing the procedure through comparison of rates of causal and acausal associations. Relative to the initial locations, final PPGT relocations show smaller residuals and shifts in epicenter as much as 3 km, predominantly toward the north and northwest. PPGT residuals are similar to those from relocation with standard station corrections. The initial hypocenters showed an unreasonable deepening with distance from station PB02, whereas PPGT relocations produced an overall flattening of event depths. These results are consistent with PPGT corrections giving real improvement in absolute location accuracy.

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1029/2019JB017727

Plain Language Summary

The TexNet seismic-monitoring program in Texas was established in response to an increase in human-induced earthquakes in the central U.S. since 2009. Accurate determination of the geographic location and depth in the Earth of earthquakes is critical to TexNet monitoring, allowing potential association of earthquakes with human activity. But accurately locating earthquakes is difficult, due to lack of geologic knowledge and sparsity of seismic monitoring stations. For a study area in western Texas, errors in TexNet earthquake location and depth may be as much as 4 – 5 km. Ideally, accuracy would be improved by calibration of the seismic network using *ground-truth* seismic events, such as quarry blasts, but such information is unavailable in this area. Instead, assuming that some earthquakes could be caused by hydraulic-fracturing in the study area, we statistically associate earthquakes in space and time to fracturing activity. We then use the known locations of this associated activity as *proxy* ground-truth to calibrate the seismic network. Our results suggest some earthquakes in west Texas are more likely due to hydraulic-fracturing than salt-water disposal. Quality measures and spatial patterns of earthquakes relocated using this proxy ground-truth calibration suggest a real improvement in absolute location accuracy.

1. Introduction

Induced earthquakes have been recorded recently in the southern mid-continent of the U.S., including Texas (Ellsworth, 2013; Frohlich et al., 2016; Keranen and Weingarten, 2018). These events, potentially associated with development of hydrocarbons and associated disposal of wastewater, have led to substantial public discussions regarding cause, public safety and potential risks of damage to infrastructure. In an effort to monitor and explain these events, and earthquake activity in general, the 84th Texas Legislature funded the creation of a statewide, seismic monitoring program, known as TexNet (Savvaidis et al., 2019). The goal of TexNet is to provide authenticated data to evaluate the location, frequency and likely causes of natural and induced earthquakes.

TexNet, as of December 2018, had installed 68 new broadband seismic stations in Texas. Of these new stations, 26 are permanent and form, along with 18 existing broadband stations, a 70km to 120km spaced, backbone seismic network in the state; 42 of the new stations are portable and have been deployed in 5 priority areas of the state having recent seismicity and which are of high socioeconomic importance. Implementation of TexNet has reduced the magnitude of completeness (Mc) for west Texas from 2.7 to 1.3, along with a large decrease in uncertainties of earthquake-source parameters (Savvaidis et al., 2019).

High-accuracy, absolute location of seismic events with realistic uncertainty estimates is critical to TexNet monitoring, especially when seismic sources must be precisely associated with geographic locations of human activities. Absolute-location accuracy of as little as a few hundreds of meters is desirable for monitoring anthropogenic sources (Foulger et al., 2018), but local-scale, absolute seismic-event location with station spacing of tens of kilometers typically has absolute errors of several kilometers, as we demonstrate here for TexNet in west Texas. An understanding of, and improvement in accuracy and error in the absolute location of seismic events beneath the Earth's surface is thus critical for monitoring of anthropogenic sources such as induced seismicity.

Absolute location of an underground seismic event is directly or indirectly controlled by time-based information (e.g., arrival time of energy onset or peak of waveform correlation on seismogram waveforms), and errors in location arise from errors in detecting and estimating this timing. Additionally, a conversion between this time-based information and distance is required in estimating an event's spatial location while simultaneously

constraining its origin time. Because this conversion is provided by a reference model of seismic wave velocity, which always differs from an ideal model that produces the same seismic-phase travel times or waveforms as in the true Earth, errors always occur in prediction of travel times. These sources of error, plus inadequacies in the network geometry, such as poor station coverage or lack of seismic stations near and above the events, are some of the primary causes of errors in absolute locations (e.g., Pavlis, 1986; Billings et al., 1994; Husen and Hardebeck, 2010; Buehler and Shearer, 2016).

For improvement in location accuracy, one or a combination of several approaches is typically chosen, such as a 1D velocity model is replaced with a more complicated and, presumably, more accurate 3D model (e.g., Ryaboy et al., 2001; Darold, et al., 2014) and/or regional or source-specific empirical corrections are applied to travel-times from the 1D model (e.g., Myers and Schultz, 2000; Richards-Dinger and Shearer, 2000; Nicholson et al., 2008). If available, ground-truth sources, such as artificial explosions with known locations or natural events with hypocenters highly constrained by a dense local network, are used to calibrate arrival time corrections (e.g., Bondár et al., 2001, 2004; Bondár and McLaughlin, 2009). Additional approaches to improving absolute event locations include using stations close to or above the target source zone (e.g., Darold et al., 2014; Eaton et al., 2018); adding additional constraint, especially for depth control, from secondary phases such as Moho reflections (regional - Wagner et al., 2013) or depth phases (teleseismic - Engdahl et al., 1998); obtaining a centroid depth from CMT waveform inversion (Engdahl et al., 1998), which may be quasi-independent of the arrival-time location depth, but which depends on the velocity model; or using the zone of greatest intensity of shaking as a proxy for an epicentral location (e.g., Holland, 2013b). In practice, a combination of these may be best, although in general, use of ground-truth sources is considered the best approach and a necessary component for improving accuracy of absolute seismic-event locations.

To avoid adverse effects of the unavoidable error in absolute location using first-arrival times, many studies extend absolute locations using waveform cross-correlation, differential times and high-precision, multi-event, relative location (e.g., Got et al., 1994; Shearer, 1997; Waldhauser and Ellsworth, 2000; Shearer et al., 2005; Lin et al., 2007; Trugman and Shearer, 2017, and references therein). Resulting, precise images of localized seismicity patterns, including clustering and planar features, are useful for studying natural seismicity and the temporal and spatial details of induced seismicity (e.g. Yoon et al. 2017; Holland, 2013a; Bao and Eaton 2016; Schoenball and Ellsworth, 2017). Although theory and numerical experiments suggest relative location procedures using high-precision differential times can produce more accurate absolute locations than single-event absolute location (Waldhauser and Ellsworth, 2000; Menke and Schaff, 2004), this is likely not true for real data sets and velocity-model error (e.g. Michelini and Lomax, 2004; Waldhauser and Schaff, 2008); the centroid of relative location clusters usually preserve the centroid of the underlying, absolute locations (e.g. Shearer et al., 2005). Thus, in general, relative location algorithms alone are not directly applicable to improving absolute location accuracy, though the tighter clustering of events provided by relative location gives information on absolute location bias and error caused by error in absolute, arrival time picks.

We herein consider TexNet monitoring in a Delaware Basin study area in the Permian Basin of west Texas and introduce and apply a novel technique to reduce the absolute-location error in the absence of true ground-truth information in the study area. First, we describe the study area, station distribution, seismic-velocity models, and reference event locations. Next, we show our estimation of the absolute-location error for seismic events surrounding TexNet station PB02 in the study area through analysis of relocations with different velocity models and station subsets. Then we present a probabilistic, space-time

likelihood model that industrial well activity, such as hydraulic-fracturing (HF) or saltwater-disposal (SWD), caused a seismic event observable to TexNet. We show calibration of the model parameters using the estimated absolute-location error, information from previous studies of induced seismicity, and by comparison of rates of causal and acausal HF-event associations. We next apply this model to probabilistically associate reference event locations surrounding station PB02 with HF and SWD activity, fixing each event hypocenter at the associated activity location to probabilistically accumulate station phase residuals forming *probabilistic, proxy ground-truth* (PPGT) station corrections. We use these results, other evidence and geologic considerations to show that HF activity is more likely than SWD to be causing some seismicity in the study area. Finally, using HF PPGT corrections developed under different ground-truth depth assumptions and SWD PPGT corrections, we assess residuals, epicenters and variation in the depths of seismicity with regards to reference locations, HF and SWD well locations, and expectations for spatial distribution of induced seismicity.

2. Hydraulic fracturing and saltwater-disposal datasets

The HF dataset contains HF activity reported through IHS (ihsmarkit.com/products/us-well-data.html) and FracFocus (fracfocus.org) for wells in and around the Delaware Basin study area over the period September 2011 through May 2019. In Texas, HF reporting can be delayed by as much as 5 months. The dataset includes information on HF jobs, consisting of one or more HF stages, describing each job with a well identifier (UWI_API) and the times and spatial positions of the first (toe) and last (heel) limits in the horizontal production well of the stages in the job. We set job start and end times to the first and last stage times, respectively, although for the IHS data, there is no reported end time for most jobs. For the FracFocus data, most jobs have reported end times; the difference between the last and first stage times gives a mean job duration of 14 days with most durations being between 5 and 20 days.

We also had information on SWD well activity in the study area over the period 1983 through January 2019 (Lemons et al., 2019), this activity may have caused some of the observed seismicity. However, temporal association of this SWD activity with seismic events provides little information on causality because of a lack of start and end time precision and very long durations (up to decades) of operation. Also SWD reporting can lag by as much as 1 year and a possible 3 months for processing gives total delay up to 15 months. SWD well activity may be causing some seismicity, which would add noise to our analyses using HF wells only; to mitigate adverse effects of this type of noise, we use a probabilistic association procedure. The SWD dataset also adds information for interpreting seismic event depths. The dataset describes each job with a well identifier (UWI_API), the time limits of periods of injection/operation, spatial positions of the injection, and total volumes per month. There is very low precision on time and volumes.

The reliability and completeness of the HF and SWD datasets are unknown. Skoumal et al. (2018) found that ~8% of HF wells between 2014 and 2016 in the Oklahoma Corporation Commission well completion database were not reported to FracFocus and note apparent errors in reporting.

3. Station network, velocity models and reference locations

Our study area (Delaware Basin study area; Figure 1) encompasses part of the Permian Basin in west Texas in which the majority of events recorded by TexNet in January 2017 – May 2019 are located (Savvaidis et al., 2019). We used a subset of TexNet permanent and temporary stations (Savvaidis et al., 2019) up to ~300km away from the Delaware Basin

study area and available between January 2017 and December 2018 (Figure 1; Table S1). Station spacing in and near the study area is around 40 km.

3.1. Velocity models

We used two seismic P- and S-wave velocity models developed for the west Texas area (Figure 2): PB3D - a smooth, 3D, tomographic model defined and developed in spherical coordinates (Huang et al., 2017), and DB1D - a 1D, constant-velocity layer model in Cartesian coordinates (Huang et al., 2017; Savvaidis et al., 2019). Although the DB1D model is used currently in TexNet monitoring in the study area, we used the PB3D model (Figure 3) for most of our analyses here because it accounts for lateral velocity variations, which should give more accurate absolute locations, and has a smooth depth profile, which reduces location artifacts (hypocenter streaking) near unrealistic layer-discontinuities (e.g. Shearer, 1997).

3.2. Reference locations

For all event locations we used the probabilistic, global search-location algorithm NonLinLoc (Lomax et al., 2000, 2014), which is also used by TexNet for reviewed, final locations. We formed a set of reference locations for the Delaware Basin study area starting from events detected and located by TexNet between January 2017 and May 2019. We located these events using NonLinLoc in spherical coordinates with travel-times from the PB3D model, and only locations with RMS $\leq$ 0.4sec, number P + S phases $\geq$ 12, ellipsoid semi-major axis $\leq$ 6.0km, and primary gap $\leq$ 120° were retained to form a set of 3103 high-quality reference locations (PB3D reference locations; Figure 4). These quality cutoff values for event selection were chosen to match values typically used in equivalent-scale seismicity studies, to remove the tails of histograms of the corresponding quality parameter for the full set of locations, and to balance the number of events removed by each quality filter.

We used subsets of PB3D reference-event locations so that our analyses and PPGT station corrections would be spatially localized and source specific, avoiding problems caused by lateral-velocity variations across the study area (e.g. Richards-Dinger and Shearer, 2000). In particular, we focused on events within 12 km (PB02 12-km event set; 686 events; $0.7 \leq ML \leq 2.8$; $Mc \approx 1.5$) and 25 km of station PB02 (PB02 25km event set; 1,413 events; $0.7 \leq ML \leq 3.1$) near the middle of the Delaware Basin study area (Figure 4). We used these event subsets around PB02 because they contain a large number of events and because event epicenters and depths are best constrained by P and S picks from nearby stations (e.g. events within a distance of a few source-depths of station PB02).

One feature of particular note in the reference locations is the clustering of shallowest events under and near stations and a general, unexpected deepening of hypocenters away from stations (Figure 4 and Figure 5a). The 12- and 25-km limits capture this zone of deepening of seismicity around PB02 and its near surroundings. A 12-km radius, given the station spacing and the size and distribution of seismicity clusters, is a good candidate size for a moving spatial filter for development of source-specific station corrections (SSST, Richards-Dinger and Shearer, 2000) over the monitoring area. SSST station corrections (Table S3) are derived from the mean of the reference-location residuals.

The zone of deepening of seismicity around PB02 and its near surroundings (Figure 4 and Figure 5a) may be related to the detection threshold improving nearer to stations or more accurate picking of closer events, but more likely suggests errors in the velocity model that affect travel times along steeply up-going rays to nearby stations with opposite polarity to times along near-horizontal to down-going rays to more distant stations. Compensating for such a model error and improving absolute-event depth determinations have been one of the

main motivations for the current study. This deepening away from PB02 is less apparent for relocation using SSST corrections within 12km of PB02, but persists beyond 12km (Figure 5b). For relocations using the DB1D model (Figure 5c), relative to the reference, PB3D model locations, the pattern of deepening is similar, while the events within 12 km of PB02 are overall about 1km deeper.

4. Estimation of absolute-location errors

In order to develop PPGT station corrections on the basis of probabilistic association of seismicity with well activity, we required a realistic estimate of the absolute-location errors in TexNet event locations. Velocity-model error and non-optimal station coverage are primary causes of absolute-location error (e.g., Billings et al., 1994; Husen and Hardebeck, 2010). To account for these sources of error, we compared PB3D model reference locations with multiple relocations in the DB1D model using different station subsets, a form of "jackknife" resampling (Tichelaar and Ruff; 1989). This set of relocations helped define absolute error by including and excluding source-station paths that may have had large model and travel-time error, by exploring variations in station coverage in distance and azimuth and by changing the velocity model and geometry. Because the PB3D and DB1D models were based on seismic analyses in the west Texas area, they form different but valid representations of the true velocity structure in the area. PB3D is a smooth, 3D model in spherical coordinates, whereas DB1D is a constant-velocity, 1D layered model in Cartesian coordinates, introducing differences in parameterization and geometry into relocations.We performed multiple relocations in the DB1D model with the PB02 12-km event set (DB1D/PB02 relocations), as well as, for each relocation, a random subset of half the available stations, except that station PB02 was never excluded. These relocations show significant shifts in epicenter and depth relative to the corresponding PB3D reference locations (Figure 6).

The histograms in Figure 6 show a typical difference in epicenter (dh) of as much as 4 km, with 68% (1-σ) of the differences dh < 2.0 km, and a broad range in differences in depth (dz) of ~–4 to 5 km. A similar set of relocations allowing exclusion of station PB02 shows a typical difference in epicenter (dh) of as much as 5 km, with 68% of the differences in epicenter dh < 2.2km, and differences in depth (dz) of ~–5 to 7km.

From these results (which include relocations with different models, different station subsets and with and without excluding station PB02) we retained a typical epicentral error of as much as 4km, a representative standard deviation (1-σ) for epicenter of about 2.0km, and absolute depth error of as much as 5km

5. Probabilistic, proxy ground-truth station corrections

In this section, we introduce association of seismic events with industrial well injection activity as proxy ground-truth information to develop probabilistic, proxy ground-truth (PPGT) station corrections for seismic sources in the Delaware Basin study area. To calibrate the association procedure, we provisionally assume that most seismic events located by TexNet in the study area in 2017-2019 may have been caused by HF. We later show and discuss evidence to support this assumption. We calibrate the association procedure by maximizing causal relative to acausal HF-event associations. We cannot calibrate with SWD activity because of the lack of start and end time precision and reporting of only monthly and not daily injection volumes in the available SWD information. Additionally, for these reasons, association with SWD is only semi-quantitative with high temporal uncertainty.

We associate seismic events with HF and SWD well activity and accumulated station corrections probabilistically because the large absolute error in event locations and high density of wells necessarily produce incorrect associations, and because some seismicity associated to HF or SWD may be due to SWD or HF, respectively, or other industrial activity. We defined a likelihood of association between each event and HF activity at a well on the basis of their horizontal-separation distance and on the difference between event origin time and HF end time. For each event-well association, we relocated the event with its hypocenter constrained at the well location. We accumulated the resulting residuals for each observed station/phase arrival probabilistically using the association likelihood as a weight over all event-well associations so as to form station/phase corrections.

5.1. Algorithm

We probabilistically associate seismic events with well activity (HF jobs or SWD) through a likelihood of association, L_A,

$$L_A(\Delta t, \Delta d) = L(\Delta t) L(\Delta d), \quad (1)$$

where Δt is the difference in time between the event origin and activity start time and Δd is the horizontal distance between the event epicenter and the center point between the toe and heel coordinates of a HF job or SWD well location.

$L(\Delta t)$ (Figure 7) has the form

$$\begin{aligned} L(\Delta t) &= t/t_{prec} && \Delta t \le t_{prec} \\ L(\Delta t) &= 1 && t_{prec} < \Delta t \le t_{prec} + \Delta t_{crit} \\ L(\Delta t) &= e^{-\left(\Delta t - \left(t_{offset} + \Delta t_{crit}\right)\right)^2 / 2\sigma_t^2} && \Delta t > t_{offset} + \Delta t_{crit} \end{aligned}, \quad (2)$$

where t_{prec} defines a linear ramp likelihood to accommodate job-start time precision; t_{prec} = 1day or 1 hour for start times reported to day or hour precision, respectively; Δt_{crit} is a critical time difference set to the activity duration or to a minimum Δt_{crit} value if the activity duration is less than this minimum or a activity end time is not available; and σ_t is a characteristic time uncertainty defining a Gaussian taper from L=1 toward L=0.

$L(\Delta d)$ (Figure 7) has the form

$$\begin{aligned} L(\Delta d) &= 1 && \Delta d \le \Delta d_{crit} \\ L(\Delta d) &= e^{-\left(\Delta d - \Delta d_{crit}\right)^2 / 2\sigma_d^2} && \Delta d > \Delta d_{crit} \end{aligned}, \quad (3)$$

where Δd_{crit} is a critical distance related to the typical well toe to heel half-length and a characteristic maximum distance for causality between fluid injection activity and induced seismicity σ_d is a characteristic distance uncertainty defining a Gaussian taper from L=1 toward L=0.

The processing procedure includes:

1. Forming an event catalog of NonLinLoc locations for a set of events.
2. Probabilistically associating events in the event catalog to well activity using Equations 1–3. For computational efficiency, we used only event-activity

associations with $L_A > L_{cutoff}$ = 0.05 for further processing; we assumed that this cutoff is low enough to result in insignificant information loss.

3. For each associated event-well activity pair, the event is re-located with its hypocenter constrained at the heel-toe center point for a HF job or well location for SWD. Resulting phase residuals for all constrained locations are accumulated through a weighted mean using each event-well activity association probability as a weight to form PPGT station/phase corrections.

See Supplementary material S1 for details.

We used the PB02 25km event subset of the PB3D reference locations and all jobs in the HF dataset or all SWD wells within 12 km of PB02 for probabilistic association with well activity and to accumulate PPGT station corrections.

5.2. Selection of association parameters

To select association parameters Δt_{crit}, σ_t, Δd_{crit}, and σ_d in Equations 1–3, we used the HF dataset and performed a grid search over a range of each of the parameters and maximized the number of causal relative to acausal HF-event associations. We expected coincidental associations, because of the high spatial and temporal rates of HF activity and seismicity in the study area. Acausal application of the probabilistic association procedure, in which the time relation between HF start and stop dates and seismic-event origin time is reversed, can produce only coincidental associations. In contrast, normal, causal association will produce a combination of coincidental and true, causal associations.

For the time parameters, given that most durations in the HF dataset are >5 days, we searched for minimum Δt_{crit} over 4, 5, 6, 7, and 8 days and for σ_t over 2, 3, 4, 5, and 6 days. For distance parameters, given the typical well toe to heel half-length of 1 km and roughly 3-km HF-event causality distances found in previous studies of HF induced seismicity (e.g. Holland, 2011, 2013a; Bao and Eaton, 2016), we searched Δd_{crit} over 1, 2, 3, and 4 km. Given our estimate of the absolute epicentral standard-deviation (1-σ) of 2.0km, we search σ_d over 1, 2, 3, 4km. This gives a total of 400 sets of association parameters for the grid search.

To define a target fitness function for selecting optimal association parameters during the grid search, we considered for each search case the sums of likelihood over all associations for $L_A > 0.5$,

$$S = \sum L_A . \qquad (4)$$

We defined the fitness function, *F*, as

$$F = (S_{causal} - S_{acausal})[(S_{causal} - S_{acausal})/S_{acausal}] , \qquad (5)$$

where the first term on the right is the absolute increase in the causal sum, S_{causal}, over acausal sum, $S_{acausal}$, and the second term in brackets is the relative increase in S_{causal} over $S_{acausal}$.

The maximum F value determined through the grid search was 76, and rounded, optimal association parameters, derived from the mean of results for the 10 best solutions, are

Δd_{crit} =2.0km, σ_d =1.0km, Δt_{crit} =5.0days, σ_t =3.5days (Table 1; Figure 7). For all search cases, $S_{acausal} < S_{causal}$, evidence that the probabilistic association procedure is capturing real causality between HF activity and seismic events and supporting our basic assumption that most seismic events in the study area could be caused by HF activity.

Association of well activity with all PB3D reference locations for Delaware Basin study area using the rounded, optimal parameters (Table 1) of produces 2399 probabilistic associations of 1217 reference events with HF jobs and 5596 associations of 1572 events with SWD activity (Figure 8). The HF well associations are numerous around station PB02 due to the high rates and spatial density of both HF and seismicity in this area. Elsewhere, the associations often cluster around individual wells or groups of close wells, usually with the wells positioned to one side of its associated events. Many HF wells have no associated seismicity, which may be due in part to real lack of induced seismicity, and in part due to the grid search finding a low, optimal values of Δd_{crit} or Δt_{crit} in response to causal-acausal ambiguity of the high HF and seismicity rates and density.

The SWD well associations show isotropic, star-like pattern of associations with events, and almost all wells have associations with any nearby seismicity. These patterns are likely a direct consequence of the long duration of activity of the SWD wells, so that each well is a associated to most nearby events solely based on proximity. It is thus difficult to asses if there is real causality between SWD well activity and seismicity, or to make meaningful comparison between these and the HF well associations.

6. Association of seismicity with hydraulic-fracturing

In the section, we present evidence of space-time causality between HF activity and seismicity, and discuss why HF activity is more likely than SWD to be causing some of the seismicity in the study area. In the Delaware Basin numerous hydrocarbon extraction and injection processes are applied in close space-time proximity; in such a case, it is difficult or impossible to unambiguously relate seismicity to specific wells and activities (Foulger et al., 2018). So our main goal is to identify HF as a possible and likely source for causing much of the seismicity. Then, the probabilistic formulation of our association procedure allows application with only HF sources, while remaining robust to noise from other causes of seismicity (e.g. SWD and other hydrocarbon extraction related activities).

The assumption of causality between HF activity and seismicity in the study area is justified by the findings of causality between HF and seismicity in other areas (e.g. Green and Styles, 2012; Holland, 2013a; Atkinson et al., 2016; Bao and Eaton 2016; Yoon et al. 2017; Skoumal et al., 2018). Hydraulic fracturing is an oil and gas well development process for injecting fluids under high pressure into a rock formation via the well. This process is intended to create new fractures in the rock as well as increase the size, extent and connectivity of existing ones. Thus, HF by design, should cause micro-seismicity during each stage of stimulation as the rock is being fractured. This micro-seismicity has magnitude typically M<-2.5, with largest events M<1.0 according to Warpinski et al. (2012), but since our PB02 12- and 25-km event sets have ML≥0.7, they are likely too large to be due directly the HF stimulation process. Besides micro-seismicity, HF can induce earthquakes of larger magnitudes due to reactivation of nearby faults (Maxwell et al., 2010). Skoumal et al. (2018) find earthquakes induced by HF of up to M3.5 in Oklahoma, and note larger HF associated events in Canada and China. The assumption is also justified by a general space-time

correlation of seismicity in and around the study area with HF activity (Figure 8a), and less so with SWD activity (Figure 8a), especially for outlying clusters of events.

Some of the strongest evidence for causality between HF and seismic events in the study area is found to the southwest of Pecos and station PB02 where relatively isolated HF wells form clusters of associations to most of the nearby seismicity (clusters A-F in Figure 8a), and there are few SWD wells (Figure 8b). For clusters A and BCD, the associated HF wells capture nearly all nearby events, and the onset of associated seismicity following soon after the onset of HF activity at one of the nearby wells (Figure 9ab). In contrast, the associated events for clusters E and F overlap in space (Figure 8a) and both show associations at the end of the HF activity (Figure 9c). Both of these temporal relations suggest a close proximity of the true absolute locations of the associated events to the associated well activity.

There are three SWD wells in our dataset in this area. One, in cluster X, associates with a subset of the events associated to HF well activity forming cluster F in Apr 2018 (Figures 8b and 9c), another SWD, in cluster Y, associated to events north of cluster F in Jul 2018 (Figures 8a and 9d). The third SWD well (in cluster Z in Figures 8b and 9d) associates with a subset of the events associated in HF cluster G in Dec 2017 and May 2018 (Figures 8a and 9d). Only SWD cluster Y shows evidence for possible causality between SWD activity and seismicity (at end July 2018) in the absence of preceding HF activity (Figure 9c), though a small date error in the HF dataset could mask association of this seismicity with apparently later HF activity (in early August 2018). In contrast, HF clusters A, B, C, D and E show causality between HF activity and seismicity in the absence of SWD activity. Also, fewer events associate to the SWD wells than to the HF wells, and there is a similarity in HF associations occurring around the end of HF activity for clusters F, E and G.

Cochran et. al., (2018) noted that inter-event times of families within 5 to 10 km of high-rate SWD wells are near-Poissonian, while families farther from high-rate wells show a high degree of clustering that may indicate that timing of events in those sequences is dominated by earthquake-earthquake interactions along well-defined shear fractures. In contrast, HF induced seismicity is distributed with short inter-event times and a high degree of clustering close to the HF wells (Eaton, et al., 2018). Thus, the high rate of seismicity during or just after the HF that drops in a short time (e.g. clusters A, B and possibly F in Figure 9) is more likely to be associated with HF than with SWD.

For all these reasons, a causality between HF and seismicity may be favored over causality with the SWD wells in the area of isolated clusters to the southwest of Pecos and station PB02, and some causality between HF and seismicity may be inferred for the whole study area.

Additional evidence of causality between HF and seismicity comes from the trend of the seismicity, the depth of the seismicity in relation to the HF and SWD wells, and characteristics of the SWD and HF target formations. The seismicity around PB02 forms NW-SSE trends (Figure 4). This strike direction does not follow the isodepths of Woodford/Ellenburger formations or the crystalline basement which show a N-S major basin axis (Comer, 1991; Ewing, 2019). However, the Permian, post-Wolfcampian isopach map depicts a NNW-SSE to WNW-ESE major-axis (Ewing, 2019), showing that the faults in Wolfcamp shales, in which HF stimulation occurs, follow mostly a strike direction more similar to the seismicity than do the deeper formations (Woodford/Ellenburger formations and crystalline basement). This is additional evidence that the seismicity is located in the Wolfcamp shales, following the depth of HF or more shallow layers.

The reference locations and relocations using SSST corrections (Figure 5) place the hypocenter depths mainly around or greater than the HF well depths, which, in turn are mostly about 1-2km deeper than the SWD wells. In the study area, SWD occur in a permeable Mesozoic formation (Delaware Mountain Group-DMG). Since the Paleozoic most of the basin area has been stable, with only slight subsistence and uplift and no faulting (Ewing, 2016), which has preserved the trapped and faulted Paleozoic hydrocarbons below the DMG (Ewing, 2018). However, since the DMG is not expected to be fractured while the shales are heavily faulted it is more likely that the seismicity is close to the shales (i.e. HF zone and below) and not in the shallow DMG zone (i.e. SWD). In other areas, seismicity is associated with deep SWD wells when injection is close to the basement (Keranen et al., 2013; Walter et al., 2017; Yeck et al., 2017). HF induced seismicity has been favored in wells that are in proximity to the basement (Skoumal et al., 2018; Brudzinski & Kozłowska, 2019) or stimulated in highly over-pressured shale formations (Eaton & Schultz, 2018). All these cases apply to the Delaware Basin, where over-pressured shales, the HF target, are on top of the faulted basement rocks (Cook et al., 2019). Thus association of seismicity with the deepest wells above the depth of seismicity, wells closest to basement, and injection in over-pressured formations would prefer association of seismicity with HF over SWD activity in the study area.

In the following, we thus assume that HF activity is the most likely cause of some seismicity in the study area, but also analyze location results using PPGT corrections derived from association of seismicity with SWD well activity.

7. Probabilistic association results around station PB02

Association of the reference events with HF activity within 12km of PB02 using the rounded, optimal association parameters (Table 1) produces 1062 probabilistic associations of reference events with HF jobs (Figure 10). Of 1167 events that can be associated around PB02, 495 were causaly and 420 acausaly associated. The density of associations roughly follows the density of epicenters, but this may indicate either real, causal relations between well jobs and events, or coincidental associations.

Relocation of reference events associated with HF jobs is done with their hypocenters constrained at the job location. Probabilistic accumulation of PPGT station corrections (tabulated in Table S2) is performed for relocated, constrained events that have the same quality cutoff values as for the reference events (RMS $\leq$ 0.4sec, number P + S phases $\geq$ 12, ellipsoid semi-major axis $\leq$ 6.0km, primary gap $\leq$ 120°), and for P residuals $\leq$ 3.0sec and S residuals $\leq$ 6.0sec.

Application of the probabilistic association, relocation, and accumulation of correction procedures provides PPGT station P and S corrections (Figure 11cd) that are similar or larger than station P and S residuals for the reference PB02 12-km event subset locations (Figure 11a, b) but that show a similar pattern of polarity and relative sizes. Reference and PPGT P corrections show a small, negative value for station PB02, large positive values for other stations over the main seismicity in the middle of the study area, and generally small or negative values for stations farthest away. Reference and PPGT S corrections show a moderate negative value for station PB02 and mainly large negative values for other stations in the middle of the study area and for stations farthest away. Most S reference-location residuals and PPGT corrections are negative, suggesting some error in this area for the PB3D model Vp:Vs ratio of 1.75 (Vs too low, Vp:Vs too high), or that S arrivals are generally picked too early.

Most station P and S residuals (Figure 11e, f) for PB02 12-km event subset relocation with PPGT corrections are much smaller than residuals for the PB02 12-km reference locations (Figure 11a, b). PPGT P and S residuals (Figure 11e, f) are of similar size to residuals (Figure 11g, h) for PB02 12km relocations using reference-location residuals (Figure 11a, b) as SSST station corrections (Table S3). We expect PPGT relocation residuals to be larger than SSST relocation residuals, because SSST corrections explicitly remove mean, reference-location residuals, and, unlike the PPGT corrections, they do not need to satisfy (proxy) ground-truth information.

8. Discussion

The statistics of causal and acausal fracturing-event associations indicates that the association procedure is performing correctly, in a probabilistic sense. Comparisons of residuals for PPGT relocation with other location scenarios indicate that combined probabilistic associations and PPGT corrections procedures are reasonably accounting for absolute errors in the PB3D velocity model and contributing to real improvement in absolute locations. We further investigated and assessed improvement in absolute location through comparisons and analyses of epicenters and depths of reference locations and relocations with SSST and alternate scenarios of PPGT corrections.

Figure 12 shows the magnitude distribution for associated and all PB02 12-km events. There is no clear difference in relative distribution for associated and all events with magnitude, suggesting that most events could be induced, especially since $M<<1.0$ is expected for hydraulic fracturing micro-seismicity (Warpinski et al. (2012). Many induced events may not be associated due to conservative, grid-search optimized association parameters needed to avoid causal-acausal ambiguity produced by the high density of well activity.

Figure 13 shows PB02 12-km events we relocated in the PB3D model using PPGT P and S corrections, and Figure 14 shows the shifts in epicenter and depth between the reference locations and PPGT relocations. Shifts in the epicenter are typically <2 km and as much as 3 km. The azimuth of the epicenter shift is predominantly toward the north and northwest, with few epicenters shifting southward. A dominant azimuth range for epicentral shift is compatible with real improvement in absolute location by PPGT corrections if the travel-time effects of unmodeled geological structures and seismic-wave-velocity variations change slowly with azimuth around the area of station PB02.

Some of the most important and interesting effects of using PPGT corrections relate to changes in depth of PPGT relocations relative to reference locations and SSST relocations. The changes in depth fall mainly between a shallowing and deepening of ~2 km, with a mean shallowing of ~0.7 km (Figure 14). An examination of event depths with respect to distance from station PB02 (Figure 15) provides more explanation of these depth changes.

As noted earlier, reference event locations (Figures 4 and 5) show a clustering of shallowest events under and near stations and a general deepening of hypocenters away from stations, possibly because of velocity model error. This pattern is evidenced for station PB02 in Figure 5a through a broad clustering of depth around fracturing-well depths within ~5 km of the station, a deepening of hypocenters to ~15 km distance, and a broad distribution of event depths for distances ≤12 km. In contrast, relative to the reference events, relocations with PPGT corrections (Figure 15a) are deeper out to ~5 km from station PB02 and generally shallower beyond, resulting in an overall flattening of event depths with distance and a narrow distribution of depths peaking at ~2 to 3 km. This flattening is consistent with PPGT

corrections correctly accounting for velocity model error, both at shallow depth above the seismicity and below station PB02, and also at greater depths along the ray paths to more distant stations. Less flattening occurs with SSST relocations (Figure 5b), evidenced in section view with a broader distribution of event depths for distances ≤12 km and generally higher depths for distances >12 km.

Relative to these PPGT relocations, relocation using alternate PPGT corrections generated with HF GT depths fixed at 4 and 6km instead of the well depth show a broader distribution of depths and still show deepening of hypocenters with distance from PB02, even out to 12-25km (Figure 15bc). Similarly, relocation using alternate PPGT corrections generated using associated SWD wells within 12km of PB02 show a shallowing of hypocenters with distance from PB02 out to 12km and no clear pattern in depths at 12-25km (Figure 15d). These results with alternate PPGT corrections are not as simple or reasonable as the stronger flattening of event depths and narrow depth distribution of HF PPGT relocations around HF wells depths, further supporting the assumption of causality between HF well activity and seismicity.

The strong peak in event depths around ~2 to 3 km with PPGT corrections for distances ≤12 km (Figure 15, middle left) corresponds to fracturing-well depths, but this is an expected consequence of constraining event depth to the well depths for calculation of PPGT corrections. There is moderate density of events at shallower depth (1–2 km) and greater density deeper (3–6 km) which may partly reflect error in depth determinations. But this spread of depths is also compatible with the range of expected and observed depths for fracturing induced events (e.g., Yoon et al. 2017; Holland, 2013a; Bao and Eaton 2016) above and below fracturing wells, extending down toward basement rocks, which have a depth of ~6 km beneath PB02 (Lemons et al., 2019). Some of the shallower seismicity could also be related to SWD-well activity.

Increased depth of PPGT relocations relative to reference locations near PB02 is consistent with the small negative PPGT P corrections for PB02 (Figure 11c, Table S2; arrival times are corrected later, rays are up-going to the station, so events are pushed downward) and larger, positive P PPGT corrections for nearby surrounding stations (Figure 11c, Table S2; arrival times are corrected earlier, rays are down-going towards the stations, so events are pulled towards the station and downward). Although a similar effect may hold for S arrivals, the pattern of PPGT corrections (Figure 11d) is more complicated than for P, and additional error may be introduced by unmodeled variations in the PB3D model Vp:Vs ratio.

Beyond ~12 km away from PB02, event depths with PPGT corrections are generally shallower than those for reference locations and SSST relocations, though all sets of locations show broad scatter in depth, and many PPGT relocations locate at the surface (around -0.8 km depth). The persistent scatter and surface locations indicate a limiting distance for applicability of the PPGT corrections, corresponding to the 12-km distance cutoff for HF jobs we used to generate corrections. We would not expect real lateral changes in shallow velocity structure beyond ~12 km from PB02 and not present in the PB3D model to be accommodated by these PPGT corrections.

9. **Conclusions**

High-accuracy and realistic uncertainty estimates for absolute location of seismic events are critical to TexNet monitoring because of the need to identify associations of seismic sources with geographic locations of human activities. For the west Texas, Delaware Basin study area surrounding TexNet station PB02, we estimated an absolute epicentral error of as much as 4 km and absolute depth error of as much as 5 km. Calibration using ground-

truth information is widely considered the best approach to improving absolute earthquake locations, but such information is not available to TexNet in this study area.

To improve absolute location in the study area, we assembled proxy ground-truth information by assuming a causal relation between industrial well activity and TexNet catalog events and probabilistically associated events with well activity. Assuming that HF activity is the most likely cause of seismicity in the study area, we re-located each associated event constrained to the mean location of the associated HF job and accumulated the resulting phase residuals probabilistically to make probabilistic, proxy ground-truth (PPGT) P and S station corrections. We apply this procedure in a preliminary test using HF wells within 12km of station PB02 to generate PPGT corrections for TexNet stations applicable to event location around PB02. The number of causal associations was larger than that of acausal associations, indicating that the procedure for associating events with HF activity performs correctly. We use these results and other evidence to justify our basic assumption that seismic events in the study area could be caused by HF activity. More definitive establishment of a causal relation between HF or SWD and seismicity is not possible until complete, well-activity datasets overlapping several years of the improved catalog of seismicity provided by TexNet are available.

Locations with PPGT station corrections around station PB02 show much smaller mean residuals than do reference locations, and similar mean residuals to relocations using SSST corrections derived from reference-location residuals. Relative to reference locations, PPGT relocations show shifts in epicenter typically <2 km and as much as 3km, with azimuths of epicenter shifts predominantly toward the north and northwest. PPGT relocations using associated HF wells as GT show depths concentrating around the HF-well depths (Figure 15a), though this is an expected consequence of the PPGT procedure. The depths spread from just above to some distance below the HF wells, corresponding to expected and observed depths for HF-induced events. The reference locations show a somewhat unreasonable deepening of hypocenters with distance from station PB02 (Figure 5), whereas the PPGT relocations show an overall flattening of event depths, consistent with the PPGT corrections correctly accounting for velocity model error. Alternate PPGT relocations with GT depths fixed at 4 and 6km and using associated SWD wells (Figure 15b-c) do not show the strong flattening of event depths and a narrow depth distribution of HF PPGT relocations around the HF wells depths, supporting the assumption that most seismic events in the study area could be caused by HF activity.

An array of 25 3-component geophones is under deployment around the city of Pecos, and we hope to perform explosion shots. We expect that these datasets will provide high-quality GT sources in the area and help validate our PPGT procedures. These efforts should greatly improve epicentral accuracy, and may help reduce the absolute depth uncertainty.

We conclude that our probabilistic, PPGT association and correction procedures can correctly account for absolute errors in seismic-velocity models and produce real improvement in absolute locations. We expect that further development of this procedure, perhaps combined with high-precision, relative-location information, and application throughout the west Texas Delaware Basin study area and other TexNet priority areas, will help improve the accuracy of TexNet location of seismic sources and reliable association of seismic events with geographic locations of human activities.

10. Data and Resources

The TexNet catalog is available at http://www.beg.utexas.edu/texnet-cisr/texnet/earthquake-catalog. TexNet phase data will be available through TexNet FDSN

web-services, as soon as TexNet becomes an authoritative network from USGS/ANSS. Prior to this phase data can be available by request via the TexNet website. Tables with hypocenters plotted in Figures 5 and 15 are available through the Texas Data Repository at https://doi.org/10.18738/T8/BT9FMG. The NonLinLoc software package is available at http://alomax.net/nlloc.

Hydraulic stimulation information is available through the IHS Markit Well Database (https://ihsmarkit.com/products/us-well-data.html), Frac Focus Chemical Disclosure Registry (http://fracfocus.org) and Texas Railroad Commission. SWD data are available through the Texas Railroad Commission (https://www.rrc.state.tx.us) and through IHS Markit. Data from FracFocus and IHS are of restricted access and available only through a subscription.

We used SeismicityViewer (http://www.alomax.net/software), Folium (http://python-visualization.github.io/folium), Leaflet (https://leafletjs.com) and Generic Mapping Tools (GMT; http://gmt.soest.hawaii.edu) for seismicity maps and charts, and OpenOffice.org (http://www.openoffice.org) for spreadsheet calculations and charts.

11. Acknowledgments

The authors thank two anonymous reviewers for extensive and helpful suggestions for improving and clarifying this work. We would like to thank the State of Texas for funding this research. We thank Casee Lemons and Caroline Breton for accessing and compiling the well information used in this study.

12. References

Atkinson, G.M., Eaton, D.W., Ghofrani, H., Walker, D., Cheadle, B., Schultz, R., Shcherbakov, R., Tiampo, K., Gu, J., Harrington, R.M. & Liu, Y. (2016). Hydraulic fracturing and seismicity in the Western Canada Sedimentary Basin. *Seismological Research Letters*, *87*(3), 631-647.

Bao, X., & Eaton, D. W. (2016). Fault activation by hydraulic fracturing in western Canada, *Science*, *354*, 1406-1409. https://doi.org/10.1126/science.aag2583

Billings, S. D., Sambridge, M. S., & Kennett, B. L. N. (1994). Errors in hypocenter location: Picking, model, and magnitude dependence. *Bulletin of the Seismological Society of America*, 84(6), 1978–1990. https://pubs.geoscienceworld.org/ssa/bssa/article/84/6/1978/119873

Bondár I. & McLaughlin, K. L. (2009). A new ground truth data set for seismic studies, *Seismological Research Letters*, *80*(3), 465-472. https://doi.org/10.1785/gssrl.80.3.465

Bondár I., Myers, S. C., Engdahl, E. R., & Bergman E. A. (2004). Epicentre accuracy based on seismic network criteria, *Geophysical Journal International*, *156*(3), 483-496. https://doi.org/10.1111/j.1365-246X.2004.02070.x

Bondár I., Yang X., North R., & Romney C. (2001). Location calibration data for CTBT monitoring at the Prototype International Data Center, *Pure and Applied Geophysics*, *158*(19). https://doi.org/10.1007/PL00001155

Brudzinski, M., & Kozłowska, M. (2019). Seismicity induced by hydraulic fracturing and wastewater disposal in the Appalachian Basin, USA: a review, *Acta Geophysica*, *67*, 351-364. https://doi.org/10.1007/s11600-019-00249-7

Buehler, J. S., & Shearer, P. M. (2016). Characterizing Earthquake Location Uncertainty in North America Using Source–Receiver Reciprocity and USArray. *Bulletin of the*

Seismological Society of America, 106 (5), 2395–2401. https://doi.org/10.1785/0120150173

Cochran, E. S., Ross, Z. E., Harrington, R. M., Dougherty, S. L., & Rubinstein, J. L. (2018). Induced earthquake families reveal distinctive evolutionary patterns near disposal wells. *Journal of Geophysical Research: Solid Earth*, 123, 8045–8055. https://doi.org/10.1029/ 2018JB016270

Comer, J.B. (1991). *Stratigraphic analysis of the Upper Devonian Woodford Formation, Permian Basin, West Texas and southeastern New Mexico*, The University of Texas at Austin, Bureau of Economic Geology, Report of Investigation, No 201., 63p.

Cook, S.W., McKeee, M., & Bjorlie, S. (2019). Delaware Basin Horizontal Wolfcamp Case History: HTI Fracture Analysis to Avoid H_2S and Extraneous Water Linked to Graben Features. Unconventional Resources Technology Conference, 22-24 July, Denver, CO. https://www.abstractsonline.com/pp8/#!/7868/presentation/705

Darold, A., Holland, A. A., Chen, C., & Youngblood, A. (2014). *Preliminary Analysis of Seismicity Near Eagleton 1-29, Carter County, July 2014*, Oklahoma Geolical Survey Open-File Report, OF2-2014, 16p. http://ogs.ou.edu/docs/openfile/OF2-2014.pdf

Eaton, D. W., & Schultz, R. (2018). Increased likelihood of induced seismicity in highly overpressured shale formations. *Geophysical Journal International*, *214*(1), 751–757.

Eaton, D.W., Igonin, N., Poulin, A., Weir, R., Zhang, H., Pellegrino, S. & Rodriguez, G. (2018). Induced Seismicity Characterization during Hydraulic-Fracture Monitoring with a Shallow-Wellbore Geophone Array and Broadband Sensors. *Seismological Research Letters*, *89*(5), 1641-1651.

Engdahl, E. R., van der Hilst, R. D., & Buland, R. P. (1998). Global teleseismic earthquake relocation with improved travel times and procedures for depth determination, *Bulletin of the Seismological Society of America*, *88*(3), 722–743.

Ellsworth, W. L. (2013). Injection-induced earthquakes. *Science*, *341*(6142), 1225942. https://doi.org/10.1126/science.1225942

Ewing, E. T., (2016) Texas through time. Bureau of Economic Geology, Udden Series No. 6, Bureau of Economic Geology, 431 pp.

Ewing, E.T., (2019) Tectonics of the West Texas (Permian) Basin-Origins, Structural Geology, Subsidence, and Later Modification (Chapter 4), in *Anatomy of a Paleozoic Basement: The Permian Basin, USA (Volume 1)*, Editor Stephen C. Ruppel, Bureau of Economic Geology and American Association of Petroleum Geologists,

Foulger, G. R., Wilson, M. P., Gluyas, J. G. , Julian, B. R. , & Davies, R. J. (2018). Global review of human-induced earthquakes, *Earth-Science Reviews*, *178*, 438-514, https://doi.org/10.1016/j.earscirev.2017.07.008.

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., H. DeShon, B. Stump, C. Hayward, M. Hornbach, and J. I. Walter (2016). A historical review of induced earthquakes in Texas, Seismol. Res. Lett. *87*(4), 1022–1038, doi: 10.1785/0220160016

Got, J.-L., Frèchet, J., Klein, F.W. (1994). Deep fault plane geometry inferred from multiplet relative location beneath the south flank of the Kilauea, *Journal of Geophysical Research*, *99*(B8), 15375–15386, https://doi.org/10.1029/94JB00577

Green, C. A., & Styles, P. (2012). *Preese Hall shale gas fracturing: Review and recommendations for induced seismicity mitigation*, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/15745/5075-preese-hall-shale-gas-fracturing-review.pdf

Holland, A. (2011). *Examination of Possibly Induced Seismicity from Hydraulic Fracturing in the Eola Field, Garvin County, Oklahoma*, Oklahoma Geological Survey, Open File Rept. OF1-2011, 28 p. *www.ogs.ou.edu/pubsscanned/openfile/OF1_2011.pdf*

Holland, A. A. (2013a), 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

Holland, A. A. (2013b), Preliminary Analysis of the 2013 Love County Earthquake Swarm, *Oklahoma Geological Survey Open File Report*, *OF1-2013* (2013.9.30), 19. http://ogs.ou.edu/docs/openfile/OF1-2013.pdf

Huang, G.-C. D., Aiken, C., Savvaidis, A., Young, B., & Walter, J. (2017). Improving the velocity structure in the Delaware basin of West Texas for seismicity monitoring, *Eos Transactions AGU*, S23C-0836, https://agu.confex.com/agu/fm17/meetingapp.cgi/Paper/287401

Husen, S., & Hardebeck, J. L. (2010). Earthquake location accuracy, Community Online Resource for Statistical Seismicity Analysis, http://dx.doi.org/10.5078/corssa-55815573. Available at http://www.corssa.org.

Keranen, K. M., Savage, H. M., Abers, G. A., & Cochran, E. S. (2013). Potentially induced earthquakes in Oklahoma, USA: Links between wastewater injection and the 2011 Mw 5.7 earthquake sequence. *Geology*, 41(6), 699–702. https://doi.org/10.1130/G34045.1

Keranen K.M., & Weingarten, M. (2018). Induced seismicity. *Annual Review of Earth and Planetary Sciences*, *46*(1), 149-174.

Lemons, C., McDaid, Acevedo, J., Breton, C.L.., Hennings, P. (*2019*). Geologic, Geographic, and Temporal Variations in Saltwater Disposal Practices Within the Permian Region, Texas and New Mexico, U.S.A., *AAPG ACE 2019*. https://www.abstractsonline.com/pp8/#!/6795/presentation/6108

Lin, G., Shearer, P. M., & Hauksson E. (2007), Applying a three-dimensional velocity model, waveform cross correlation, and cluster analysis to locate southern California seismicity from 1981 to 2005, *Journal of Geophysical Research 112*(B12). https://doi.org/10.1029/2007JB004986

Lomax, A., A. Michelini, A. Curtis (2014), Earthquake Location, Direct, Global-Search Methods. In R. A. Meyers, (Ed.)*Encyclopedia of complexity and system science*, 2nd Edition, New York: Springer, p. 1-33, http://doi.org/10.1007/978-3-642-27737-5_150-2

Lomax, A., Virieux, J., Volant, P., & Thierry, B. C. (2000). Probabilistic earthquake location in 3D and layered models: Introduction of a Metropolis–Gibbs method and comparison with linear locations. In *Advances in seismic event location*, C. H. Thurber and N. Rabinowitz (Eds.), Dordrecht: Kluwer Academic, (pp. 101-134).

Maxwell, S. C., M. B. Jones, R. L. Parker, W. S. Leaney, M. Mack, D. Dorvall, D. D'Amico, J. Logel, E. Anderson, and K. Hammermaster (2010). Fault activation during hydraulic fracturing, AAPG Search Discov. 90172, 1–4, doi: 10.1190/1.3255145.

Menke, W., & Schaff, D. (2004), Absolute earthquake locations with differential data, *Bulletin of the Seismological Society of America*, 94(6), 2254–2264, https://doi.org/10.1785/0120040033

Michelini, A., & Lomax, A. (2004), The effect of velocity structure errors on double-difference earthquake location, *Geophysical Research Letters*, *31*(L09602), https://doi.org/10.1029/2004GL019682

Myers, S.C., & Schultz, C. A. (2000). Improving sparse network seismic location with Bayesian kriging and teleseismically constrained calibration events, *90*(1), 199-211. http://dx.doi.org/10.1785/0119980171

Nicholson, T., Clarke, D., & Townend, J. (2008). Regional earthquake location using empirical traveltimes in a region of strong lateral velocity heterogeneity, *Geophysical Journal International*, *175*(2), 560–570. https://doi.org/10.1111/j.1365-246X.2008.03858.x

Pavlis, G.L. (1986), Appraising earthquake hypocenter location errors–A complete, practical approach for single-event locations, *Bulletin of the Seismological Society of America*, *76*(6), 1699-1717.

Richards-Dinger, K. B., & Shearer, P. M. (2000). Earthquake locations in southern California obtained using source-specific station terms, *Journal of Geophysical Research*, *105*(B5), 10939–10960, https://doi.org/10.1029/2000JB900014

Ryaboy V., Baumgardt D., Firbas P., & Dainty, A. (2001). Application of 3-D crustal and upper mantle velocity model of North America for location of regional seismic events, *Pure and Applied Geophysics*, *158*, 79-103. https://doi.org/10.1007/PL00001169

Savvaidis, A., B. Young, G-C. D. Huang, & Lomax, A. (2019), TexNet: A statewide seismological network in Texas, *Seismological Research Letters*, https://doi.org/10.1785/0220180350

Schoenball, M., & Ellsworth, W. L. (2017) Waveform-Relocated Earthquake Catalog for Oklahoma and Southern Kansas Illuminates the Regional Fault Network. *Seismological Research Letters, 88*(5), 1252–1258. https://doi.org/10.1785/0220170083

Shearer, P. M. (1997). Improving local earthquake locations using the L1 norm and waveform cross-correlation: Application to the Whittier Narrows, California, aftershock sequence, *Journal of Geophysical Research, 102(B4)*, 8269–8283, doi: 10.1029/96JB03228

Shearer, P., Hauksson, E., & Lin, G. (2005). Southern California hypocenter relocation with waveform cross-correlation, part 2: Results using source-specific station terms and cluster analysis, *Bulletin of the Seismological Society of America 95*(3), 904– 915, doi: 10.1785/0120040168.

Skoumal, R. J., Ries, R., Brudzinski, M. R., Barbour, A. J., & Currie, B. S. (2018). Earthquakes induced by hydraulic fracturing are pervasive in Oklahoma. *Journal of Geophysical Research: Solid Earth*, *123*, 10,918– 10,935. https://doi.org/10.1029/2018JB016790

Tichelaar, B. W., and Ruff, L. J. (1989), How good are our best models? Jackknifing, bootstrapping, and earthquake depth, *Eos Transactions AGU*, 70(20), 593– 606, https://doi.org/10.1029/89EO00156.

Trugman D. T., & Shearer P. M. (2017). GrowClust: A hierarchical clustering algorithm for relative earthquake relocation, with application to the Spanish Springs and Sheldon, Nevada, earthquake sequences, *Seismological Reseach Letters*, *88*(2A), 379–391. https://doi.org/10.1785/0220160188

Wagner, M., Husen, S., Lomax, A., Kissling, E., & Giardini, D. (2013). High-precision earthquake locations in Switzerland using regional secondary arrivals in a 3D velocity model, *Geophysical Journal International*, *193*(3), 1589-1607. https://doi.org/10.1093/gji/ggt052

Waldhauser, F., & Ellsworth, W. L. (2000). A double-difference earthquake location algorithm: Method and application to the northern Hayward fault, *Bulletin of the Seismological Society of America*, *90*(6), 1353-1368. https://doi.org/10.1785/0120000006

Waldhauser, F., & Schaff, D. P. (2008), Large-scale relocation of two decades of Northern California seismicity using cross-correlation and double-difference methods, *Journal of Geophysical Research*, 113(B08311), https://doi.org/10.1029/2007JB005479

Walter, J. I., Chang, J. C., & Dotray, P. J. (2017). Foreshock seismicity suggests gradual differential stress increase in the months prior to the 3 September 2016 Mw 5.8 Pawnee earthquake. *Seismological Research Letters*, *88*(4), 1032–1039. https://doi.org/10.1785/0220170007

Warpinski, N. R., J. Du, and U. Zimmer (2012). Measurements of hydraulic-fracture-induced seismicity in gas shales, SPE Hydraul. Fract. Technol. Conf., SPE 151597, The Woodlands, Texas, 6–8 February 2012, 1–8, doi: 10.2118/151597-PA

Yeck, W. L., Hayes, G. P., McNamara, D. E., Rubinstein, J. L., Barnhart, W. D., Earle, P. S., & Benz, H. M. (2017). Oklahoma experiences largest earthquake during ongoing regional wastewater injection hazard mitigation efforts. *Geophysical Research Letters*, *44*, 711–717. https://doi. org/10.1002/2016GL071685

Yoon, C., Huang, Y., Ellsworth, W. L., & Beroza, G. C. (2017). Seismicity during the initial stages of the Guy-Greenbrier, Arkansas, earthquake sequence. *Journal of Geophysical Research*, *122*. 9253– 9274. https://doi.org/10.1002/2017JB014946

Table 1. Association parameters, results and mean values for 10 best solutions

	Δd_{crit} (km)	Δd_{crit} (km)	Δt_{crit} (day)	σt (day)	N_{causal}	$N_{acausal}$	S_{causal}	$S_{acausal}$	F
	2.0	1.0	5.0	3.0	503	332	435	287	76
	2.0	1.0	4.0	3.0	519	350	447	299	74
	2.0	1.0	4.0	4.0	497	330	429	284	74
	2.0	1.0	5.0	4.0	527	361	454	306	71
	2.0	1.0	6.0	3.0	511	347	442	297	71
	2.0	1.0	6.0	3.0	532	365	460	311	71
	2.0	1.0	3.0	4.0	512	348	441	296	71
	2.0	1.0	3.0	4.0	487	327	421	280	71
	2.0	1.0	6.0	3.0	563	391	482	330	71
	2.0	1.0	7.0	3.0	542	373	468	318	70
mean	2.0	1.0	4.9	3.7	519	352	448.1	300.9	72
rounded	2.0	1.0	5.0	3.5					

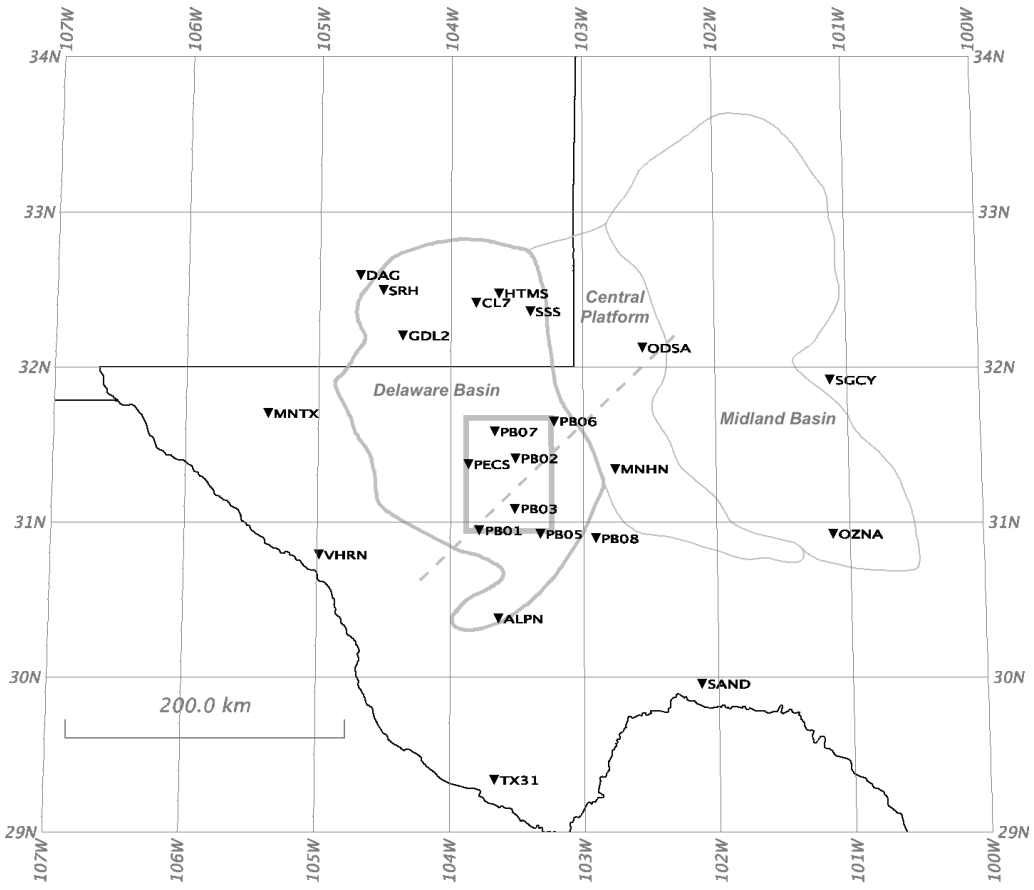

Figure 1. West Texas overview map showing the Delaware Basin (thick gray contour), other sub-basins of the Permian Basin (thin gray contours), the Delaware Basin study area (rectangle) and TexNet stations used in this study (triangles). Dashed gray line shows location of the PB3D cross section of Figure 3.

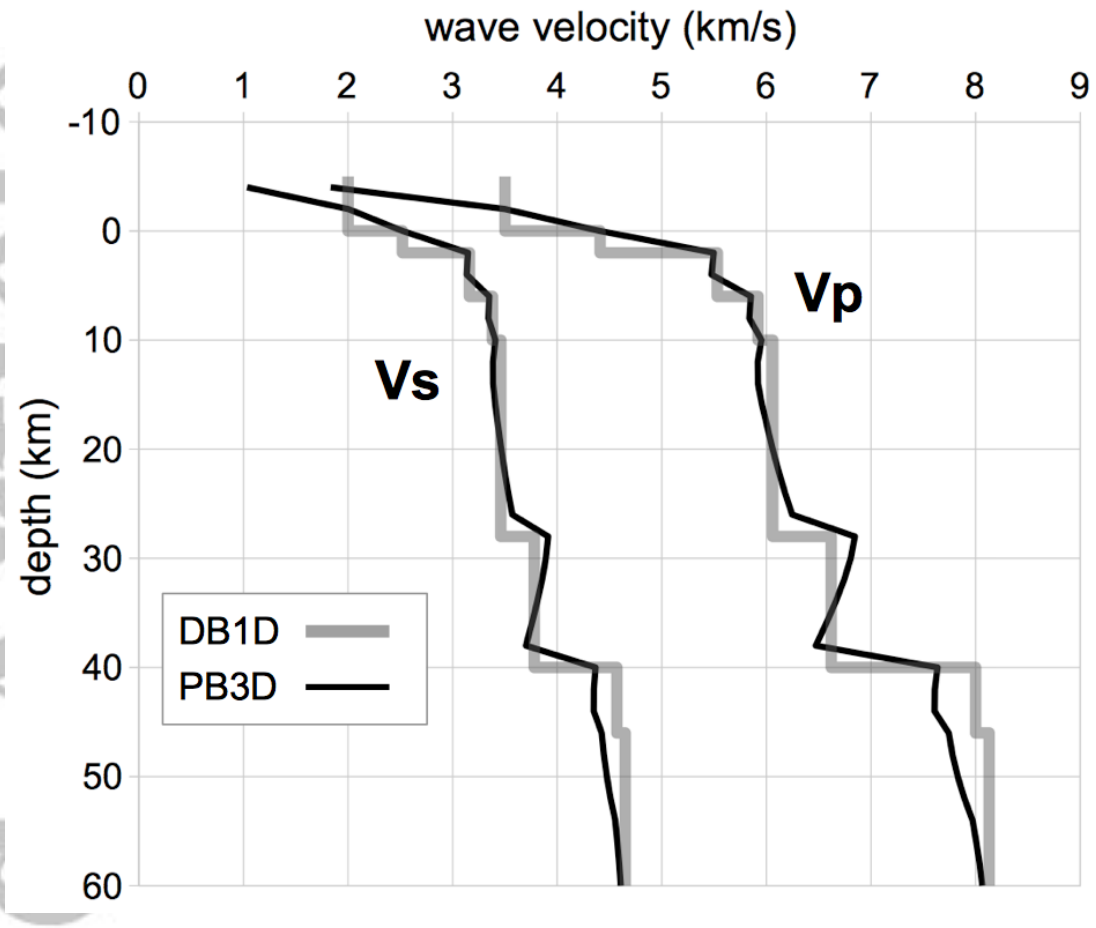

Figure 2. P and S velocity models DB1D and PB3D used in this study. The PB3D depth profile under station PB02 is plotted.

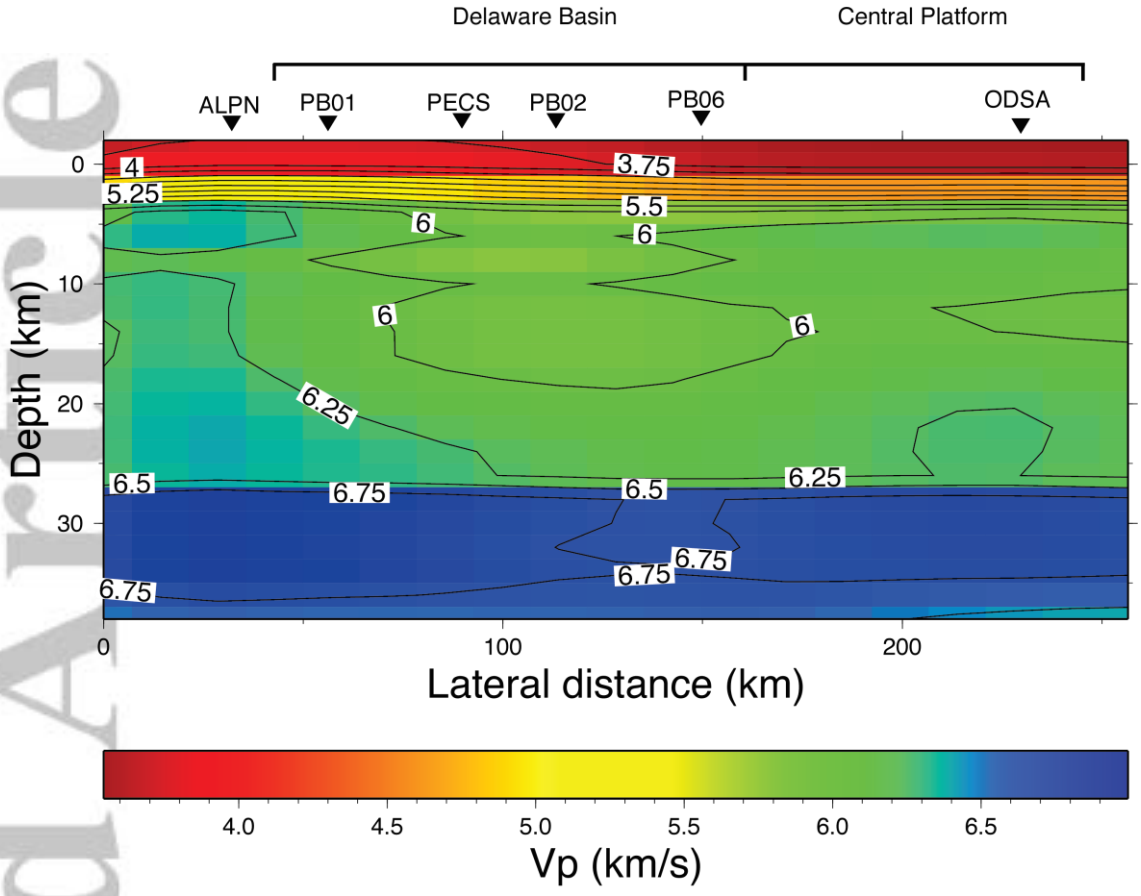

Figure 3. SW-NE cross section through PB02 for P velocity for the PB3D model used in this study. TexNet stations used in this study (triangles) are shown projected onto the section; contour velocities in km/s. See Figure 1 for location of section.

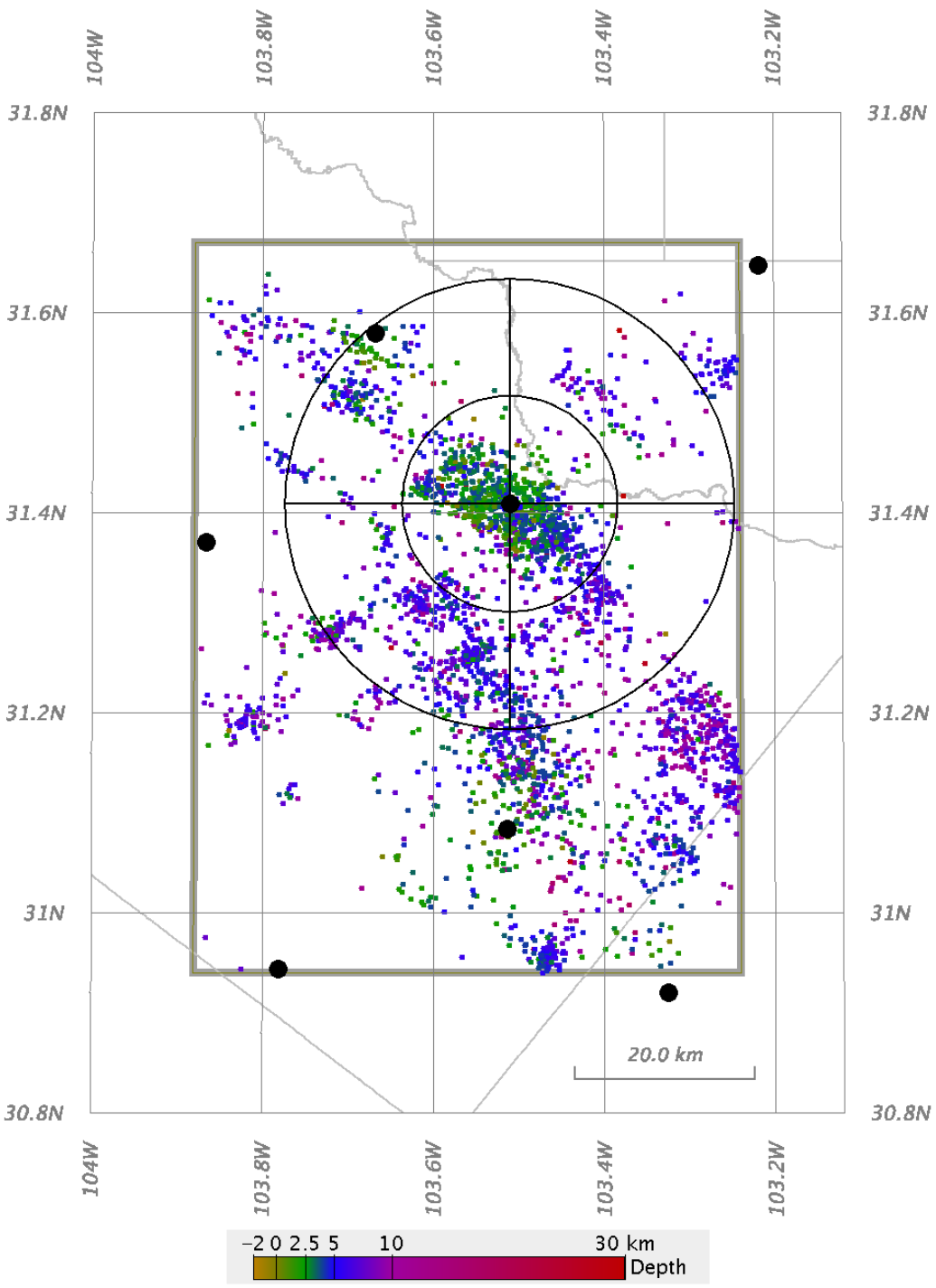

Figure 4. PB3D reference locations for Delaware Basin study area (dots colored by depth), Delaware Basin study area (rectangle), county lines (gray lines), and TexNet stations (black dots). Circles show 12- and 25-km distances from station PB02 defining PB02 event subsets.

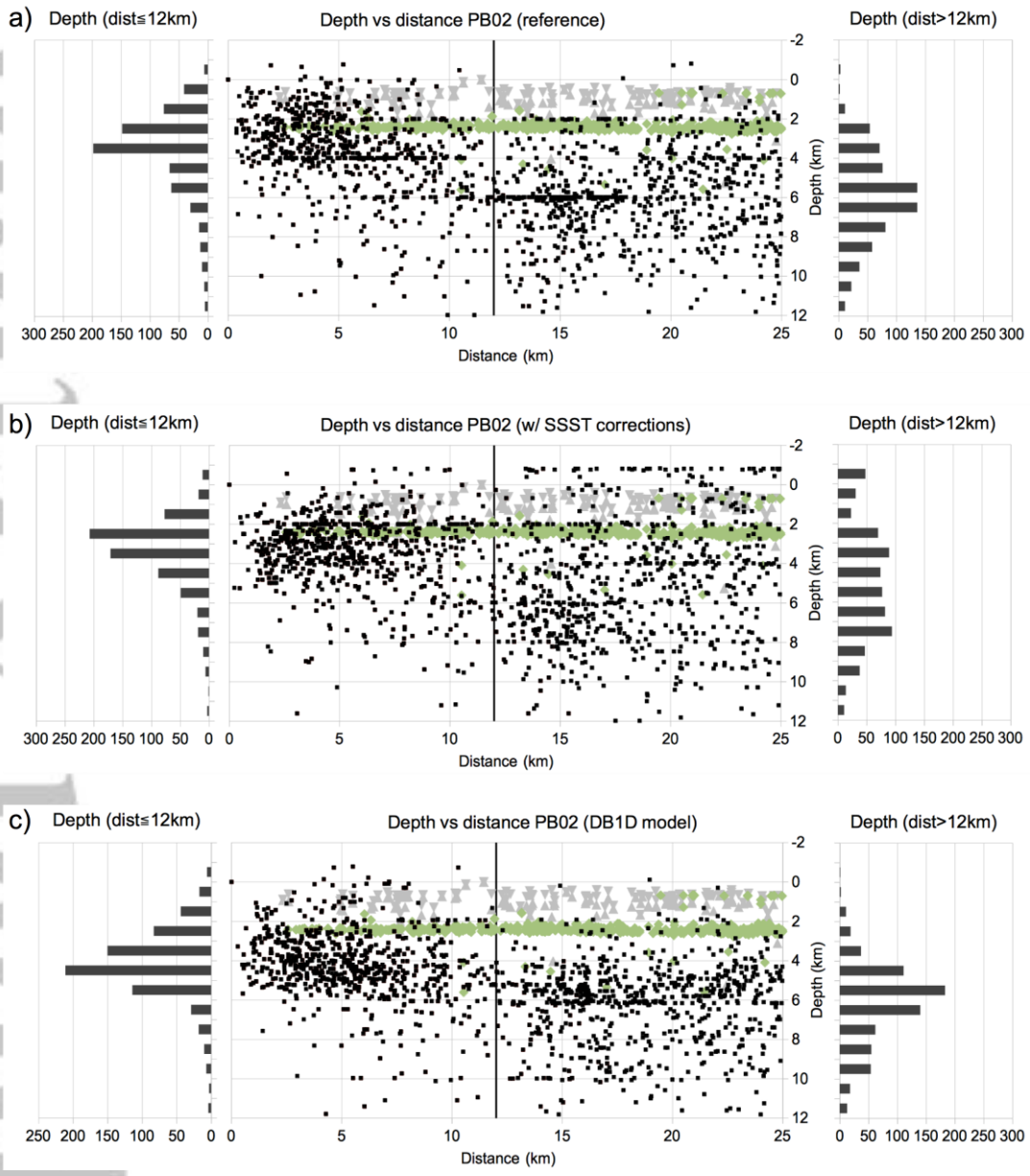

Figure 5. Section views showing (middle) hypocenter depth as a function of distance from station PB02 and corresponding depth histograms for events within 12-km distance from station PB02 (left) and 12-25km distance (right) for (a) PB3D model reference locations, and (b) PB3D relocations using SSST corrections, and (c) DB1D model locations. Also shown are heel and toe depths of HF wells within 25 km of station PB02 (green diamonds), upper and lower depths of SWD wells (Lemons et al., 2019) within 25 km of station PB02 (gray triangles), and 12-km distance from station PB02 (black vertical line). Apparent horizontal streaks in event hypocenters are due to location artifact caused by discontinuities in gradients of PB3D model velocities and station travel time fields at model-grid node depths.

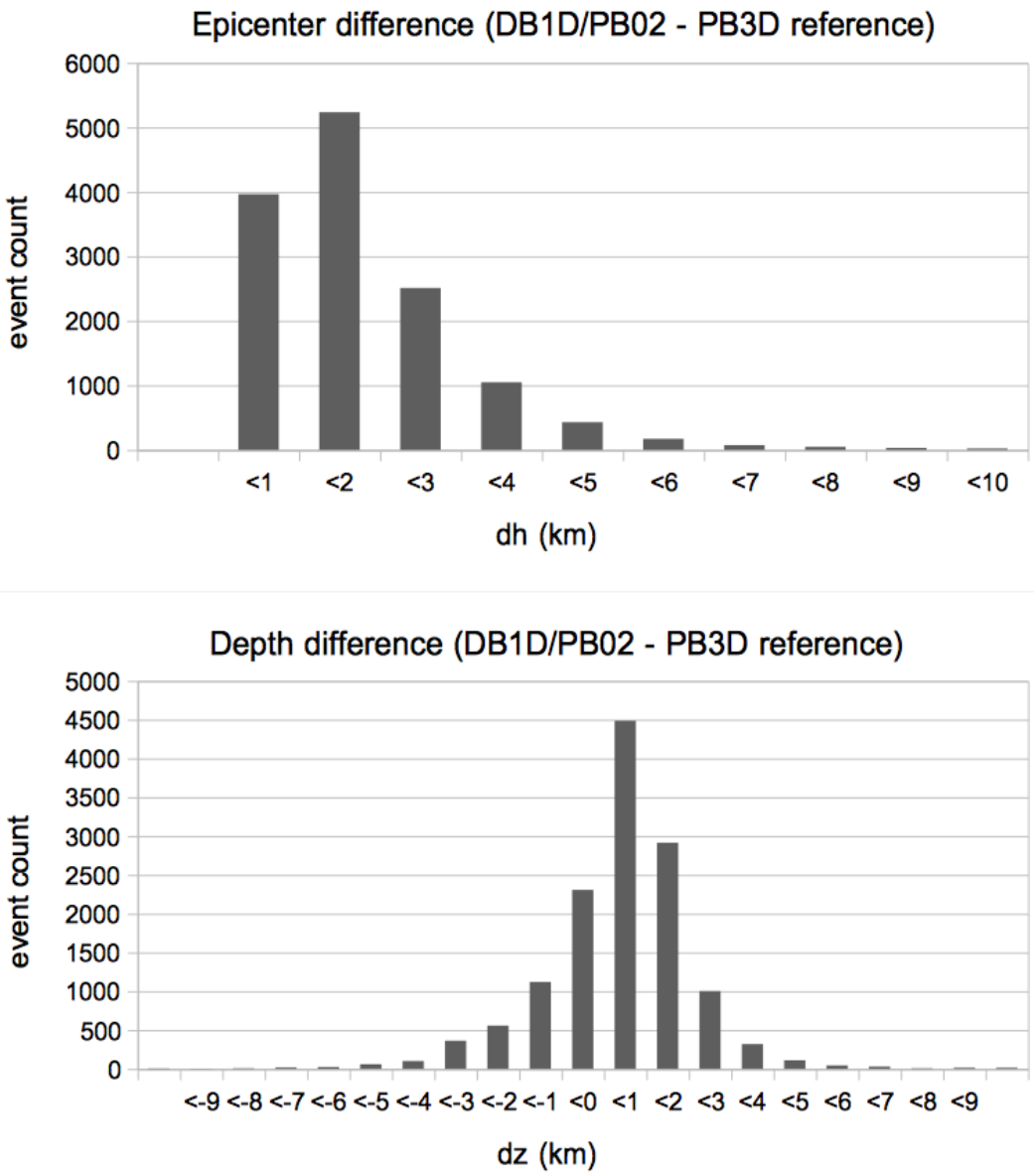

Figure 6. Histograms of differences between DB1D/PB02 12km, multiple, random-subset relocations and corresponding PB3D reference locations for (upper) epicentral distance (dh) and (lower) depth (dz).

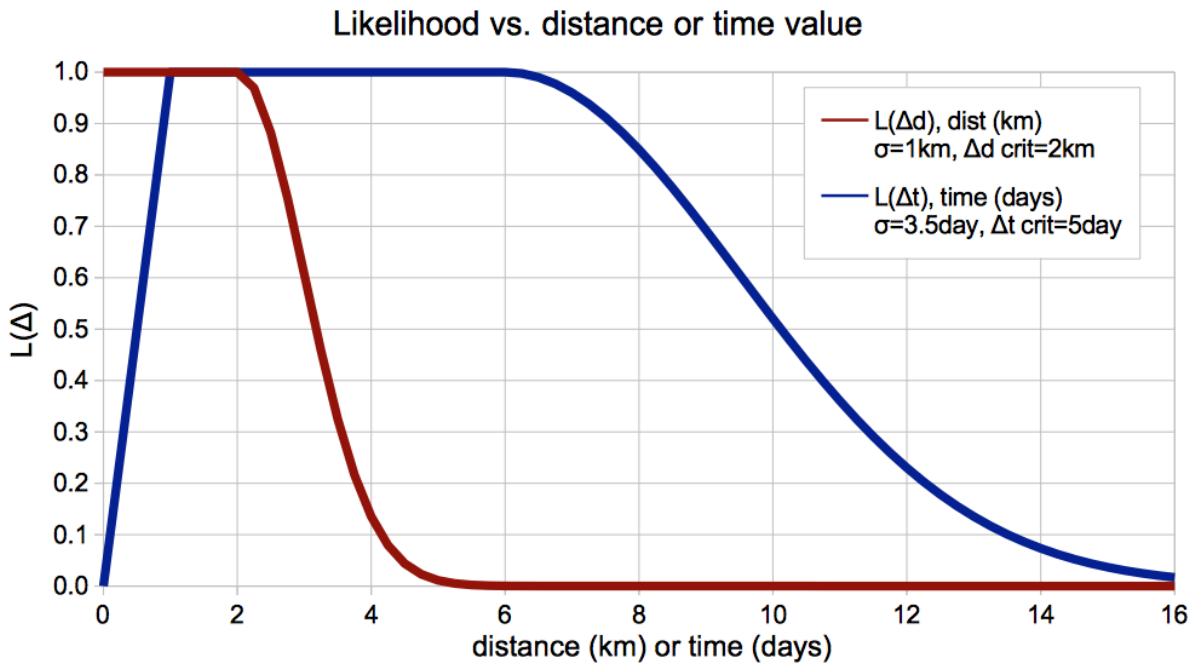

Figure 7. Likelihoods $L(\Delta t)$, where Δt is the difference between the event origin time and the real or estimated well activity stop date, and $L(\Delta d)$, where Δd is the difference in horizontal distance between an event and the center point between the toe and heel coordinates for a HF job or well location for SWD. $L(\Delta t)$ and $L(\Delta d)$ are shown for the optimal, rounded mean association parameters in Table 1.

20190625_FRAC/runs/match_event_frac_3001_SWD_all_wells/md2.0_sd1.0_st3.5_pc0.05_mfd120.0_ftp24.0_ftpw1.0_fto5.0/match_event_frac_publication.html Figure_timeline_clustersABCD.odg

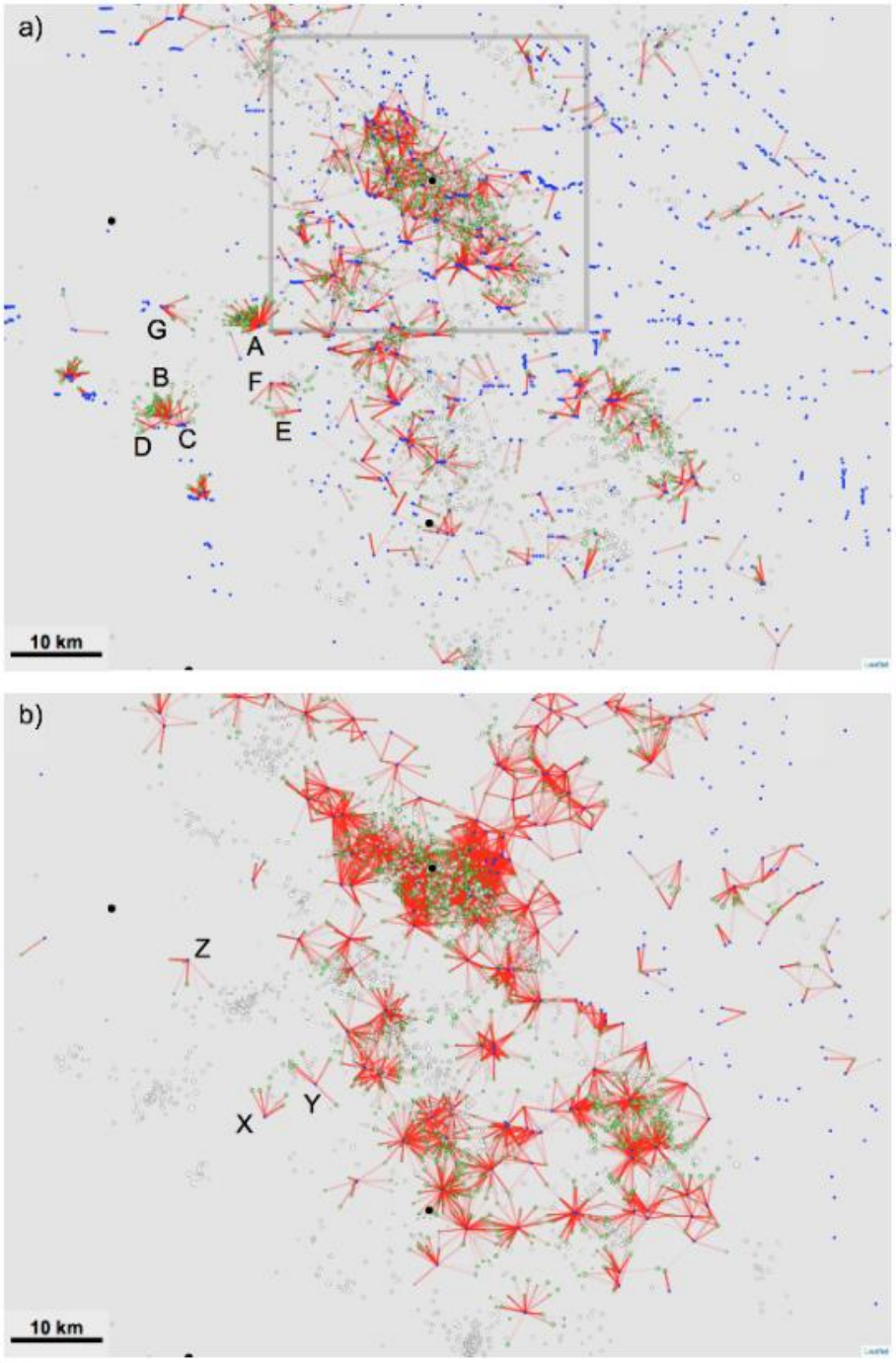

Figure 8. Probabilistic association of reference events with (a) HF jobs and (b) SWD wells. Shown are: TexNet stations (black dots); mean HF toe/heel location or SWD well locations (blue dots); reference events that are associated (green dots) or unassociated (white dots); and event-well activity associations (red lines, with opacity set in proportion to association probability L_A). Letters identify isolated well-event clusters used to examine association of seismicity with HF and SWD activity. Gray box in (a) delimits area shown in Figure 10.

Figure_timeline_clustersABCD.odg

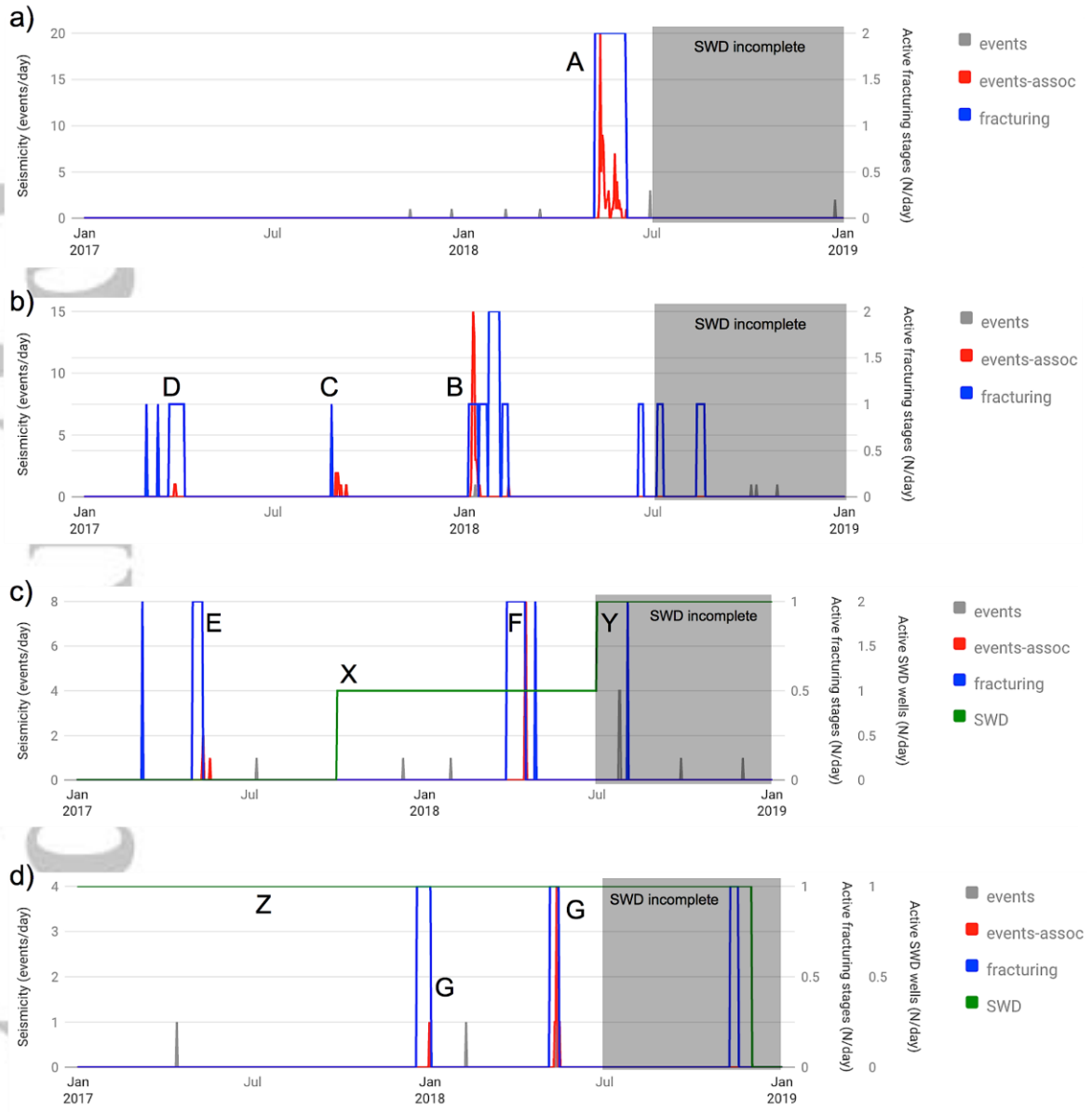

Figure 9. Timeline of daily counts of seismic event and well activity for wells associated with or within (a) cluster A, (b) clusters B, C and D, (c) clusters E, F and Y, and (d) clusters G and Z. Events associated to HF wells are shown in red; events that could be associated by distance only to any of the wells are shown in gray. Gray areas shows period after July 2018 when the SWD activity dataset may be incomplete.

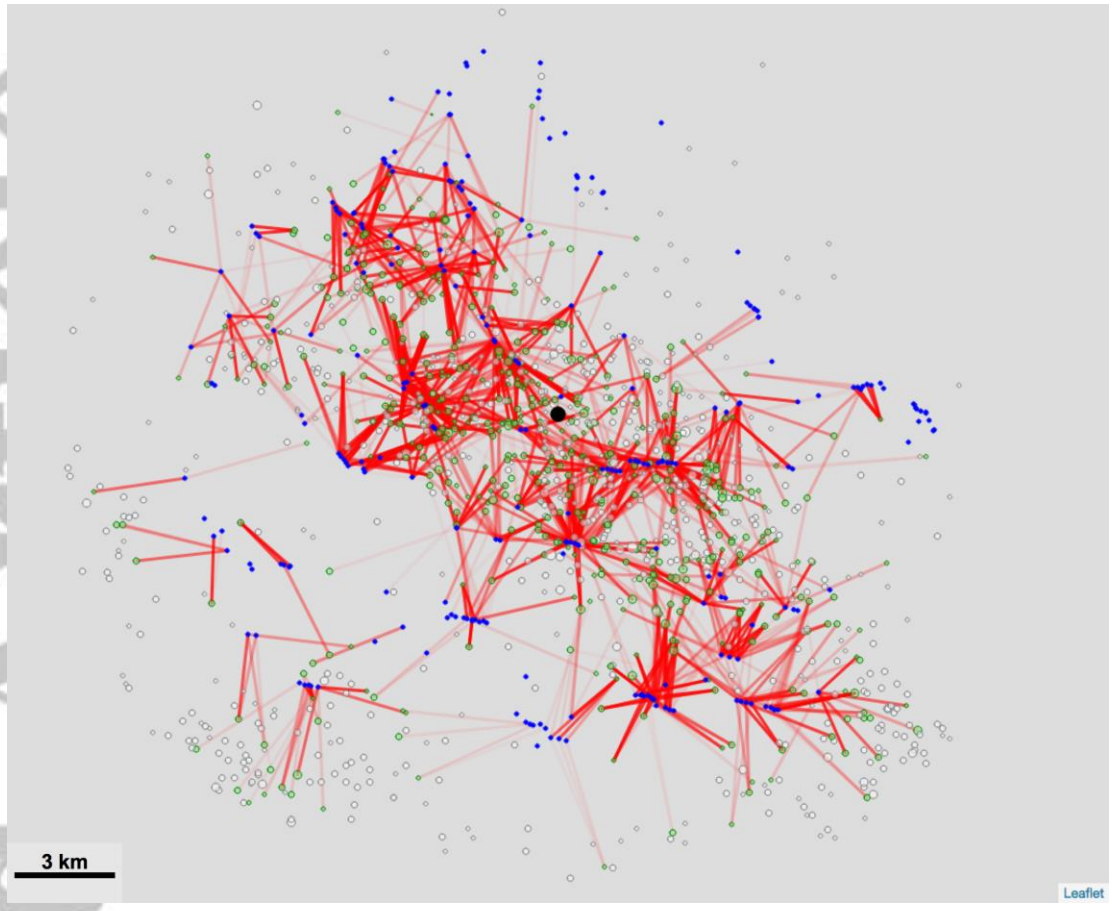

Figure 10. Probabilistic association of reference events with HF jobs (zoom; see Figure 8a for zoom location). Shown are: station PB02 (black dot); HF job mean toe/heel location (blue dots) for wells with mean toe/heel location within 12km of station PB02; reference events that are associated (green dots) or unassociated but within association distance ($L(\Delta d)\geq 0.05$, white dots); and event-HF associations (red lines) with opacity set in proportion to association probability L_A.

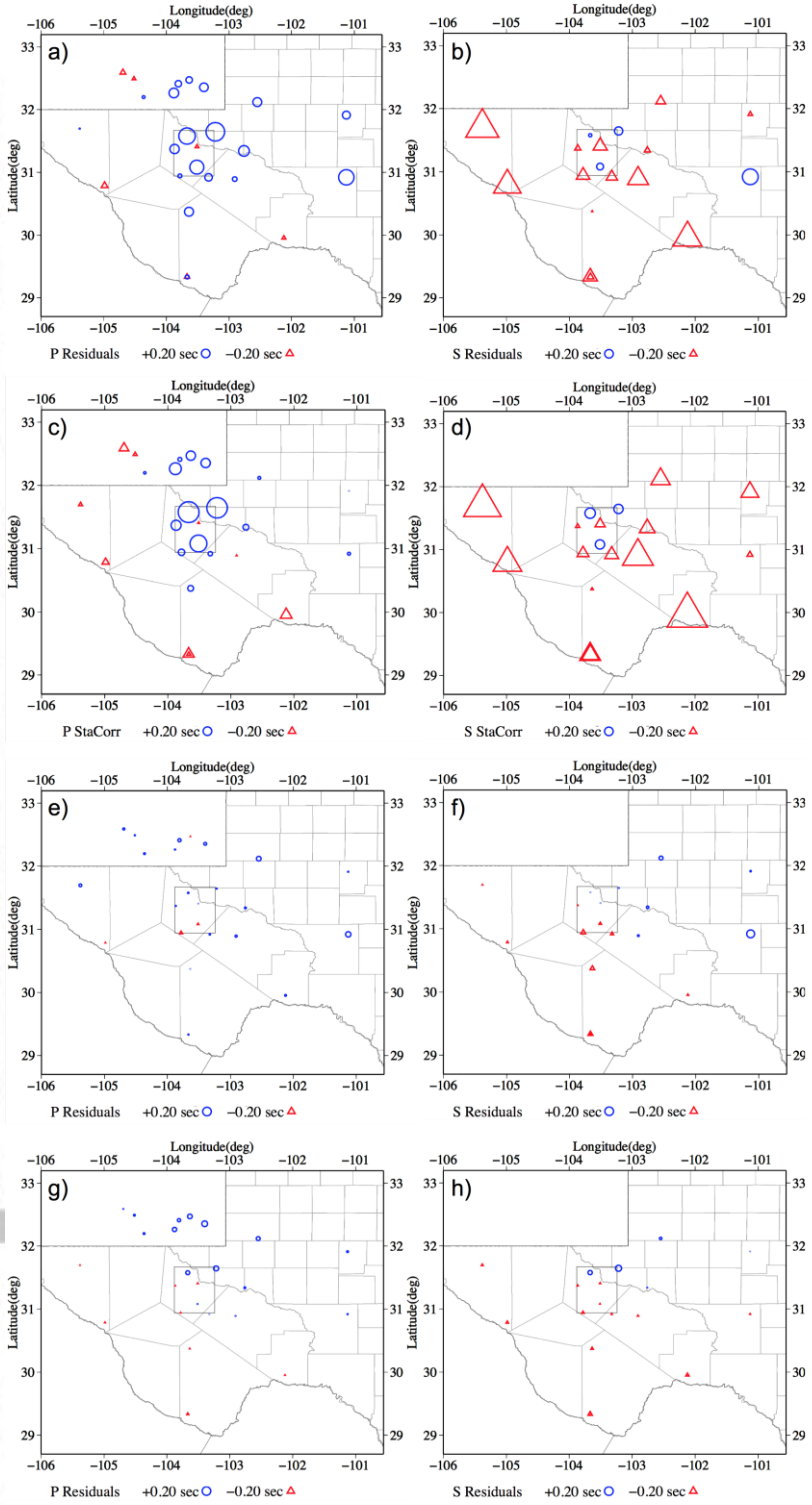

Figure 11. Mean, station (a) P and (b) S residuals (SSST corrections, Table S3) for reference PB02 12km event set locations. Station PPGT (c) P and (d) S corrections (Table S2) with event locations constrained by event-fracturing job associations. Mean, station residuals for PB02 12-km event set events after relocation, with PPGT corrections for (e) P and (f) S, and with SSST corrections for (g) P and (h) S.

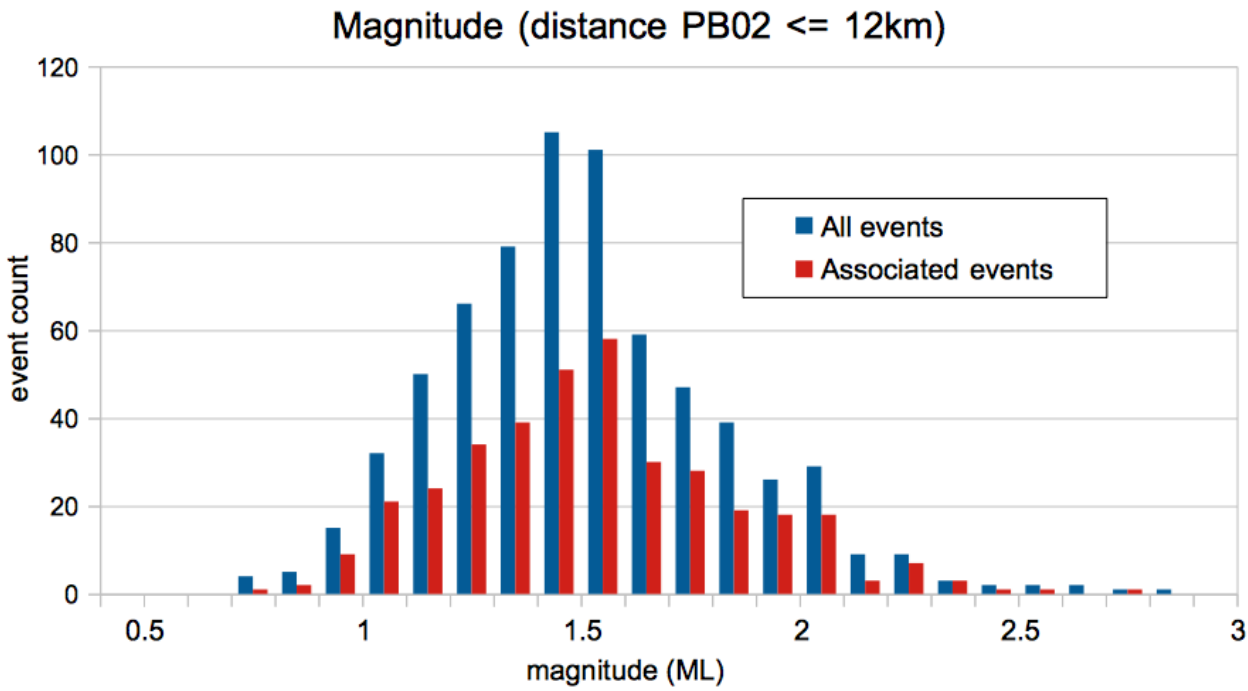

Figure 12. Histogram of magnitudes for associated and all events with reference location epicenters within 12-km of station PB02.

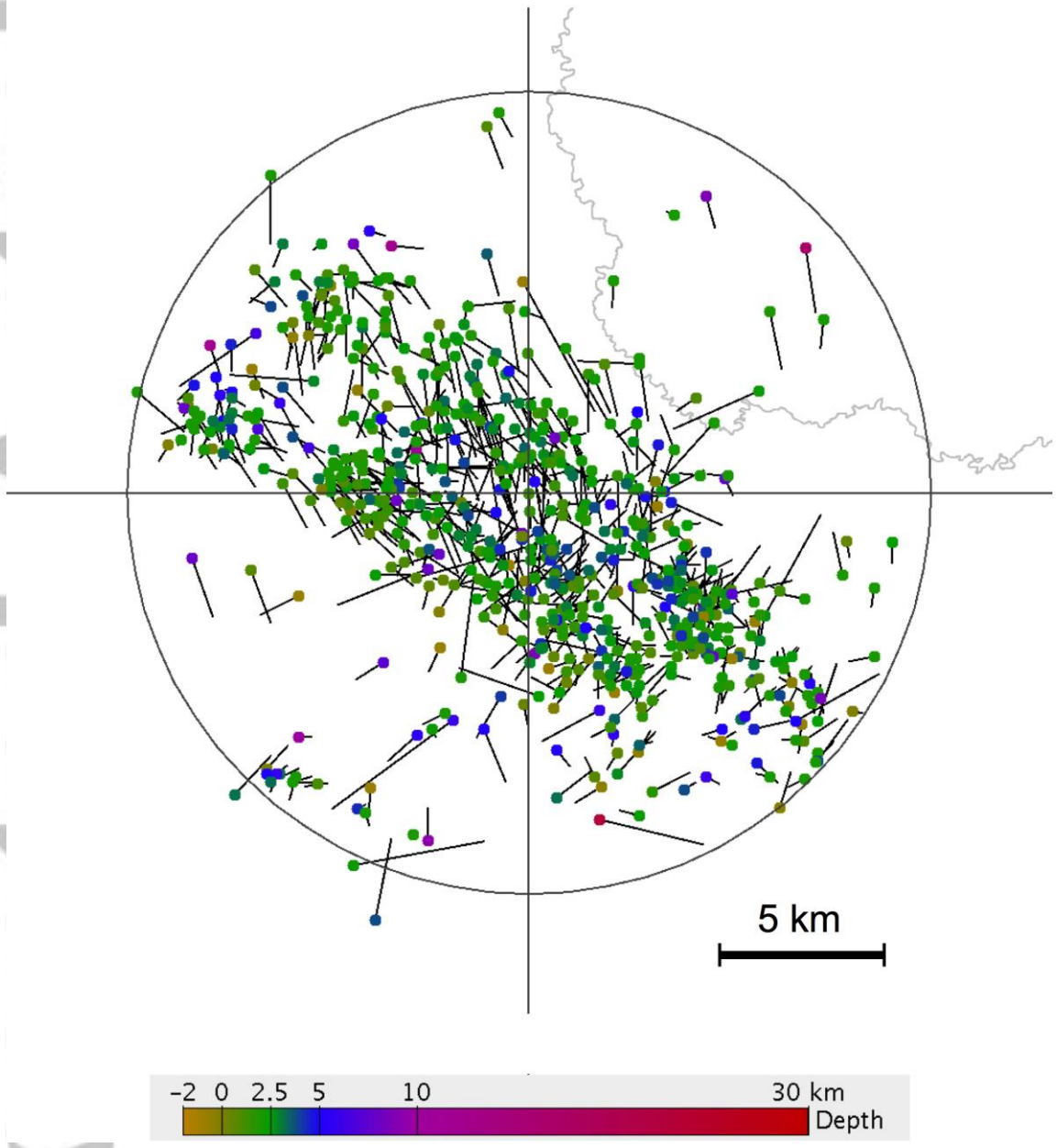

Figure 13. PB02 12-km events relocated in the PB3D model using PPGT corrections (dots colored by depth) and their epicenter shift from reference PB02 12-km events (black line segments). Circle shows 12-km distance centered on station PB02 defining PB02 reference event subset. Note that PB02 is not a reference station, and a moving spatial window would be used in implementation of PPGT corrections, exactly as with established SSST procedures.

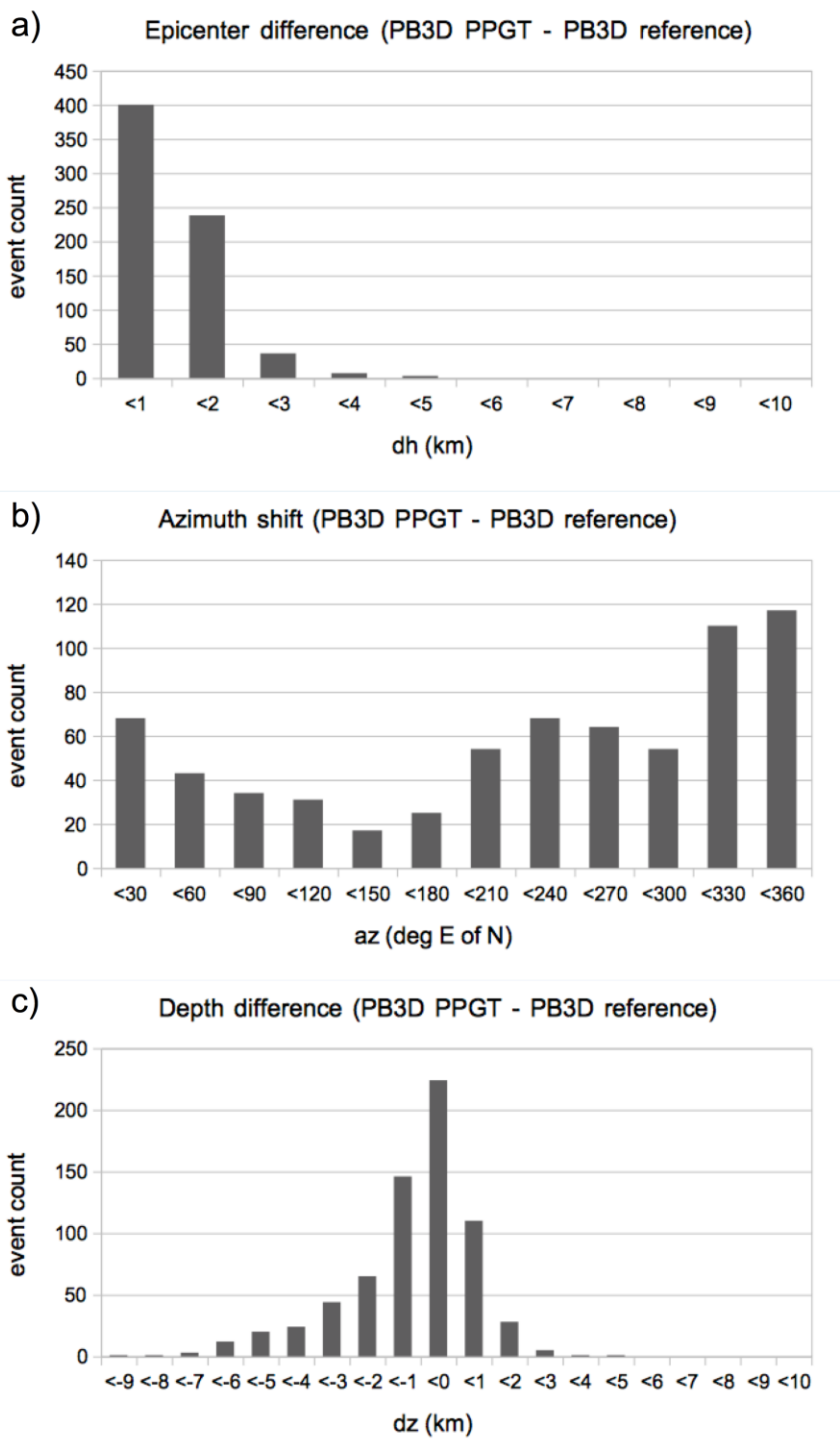

Figure 14. Histograms of differences between PB02 12-km events relocated using PPGT corrections and reference locations for (a) epicentral distance (dh), (b) azimuth of epicenter shift (az), and (c) depth (dz).

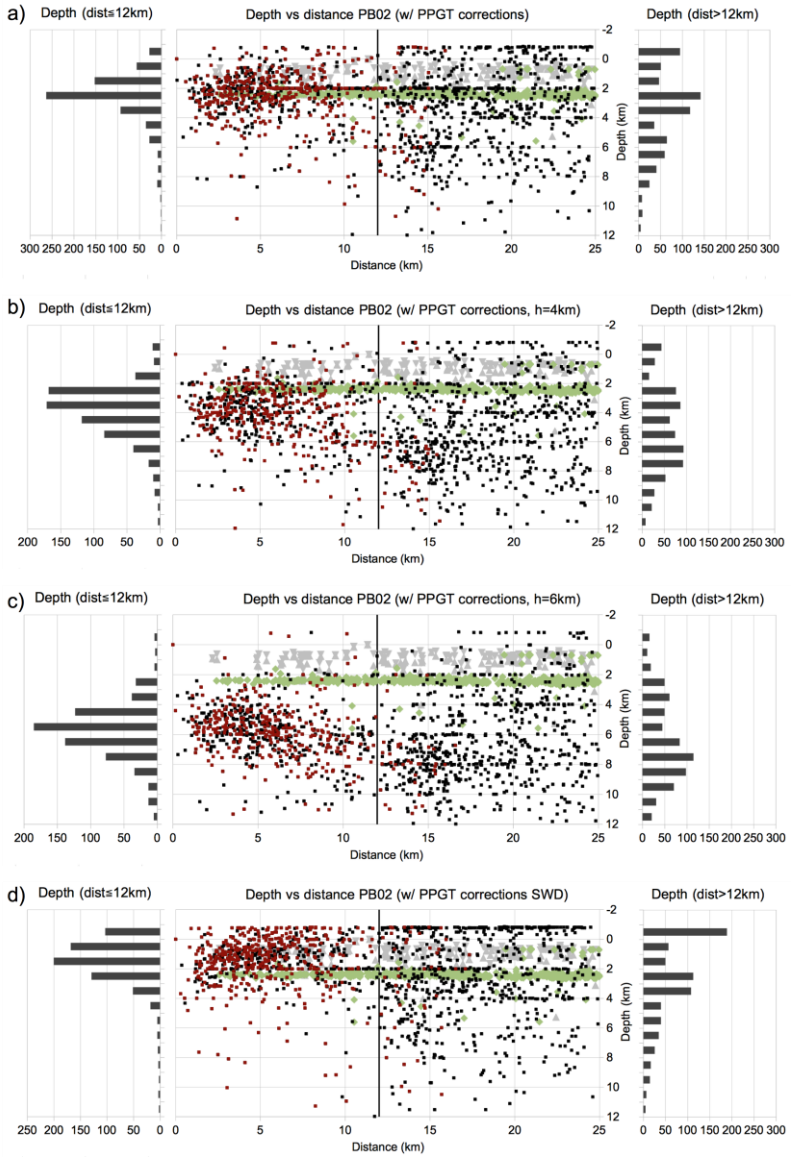

Figure 15. Section views showing (middle) depth as a function of distance from station PB02 and corresponding depth histograms within 12-km distance from station PB02 (left) and 12-25km distance (right) for relocations using (a) PPGT corrections , (b) PPGT corrections with GT depths fixed at 4km, (c) PPGT corrections with GT depths fixed at 6km, and (d) PPGT corrections using SWD wells as GT. All locations use the PB3D model. Sections show PB02 25-km events associated for each relocation case (red squares); other PB02 25km events (black squares); histograms show ssall events. Also shown are heel and toe depths of fracturing wells within 25 km of station PB02 (green diamonds), upper and lower depths of SWD wells (Lemons et al., 2019) within 25 km of station PB02 (gray triangles), and 12-km distance from station PB02 (black vertical line) used in generating PPGT corrections (Figure 10) and for analyses in Figures 11 and 13. Apparent horizontal streaks in event hypocenters are due to location artifact caused by discontinuities in gradients of PB3D model velocities and station travel time fields at model-grid node depths.