@AGUPUBLICATIONS

1

3   Supporting Information for

4   **Onset and cause of increased seismic activity near Pecos, West Texas, USA from**
5   **observations at the Lajitas TXAR Seismic Array**

6   Cliff Frohlich[1,2]*, Chris Hayward[2], Julia Rosenblit[3], Chastity Aiken[4], Peter Hennings[5],
7   Alexandros Savvaidis[5], Casee Lemons[5], Elizabeth Horne[5], Jacob I. Walter[6], and Heather R.
8   DeShon[2]

9   [1]Institute for Geophysics, Jackson School of Geosciences, University of Texas at Austin, Austin,
10   TX 78758, USA.

11   [2]Huffington Department of Earth Sciences, Southern Methodist University, Dallas, TX 75275,
12   USA.

13   [3]Department of Geology, Portland State University, Portland, OR 97207, USA.

14   [4]Geosciences Marine - LAD, IFREMER, Plouzané, France.

15   [5]Bureau of Economic Geology, Jackson School of Geosciences, University of Texas at Austin,
16   TX 78758, USA.

17   [6]Oklahoma Geological Survey, University of Oklahoma, Norman, OK 73019, USA.

18   *Corresponding author. Cliff Frohlich (cliff@ig.utexas).

19

20   **Contents of this file**

21   Text S1

22   Figures S1 to S8

23   Caption for Dataset S1

24   Table S1

25   **Introduction**

26   The Supporting Information includes text describing the Supporting Information and the

27   computer-based event detection methodology (this file, Text S1), eight figures (Figures S1 to S8), a

28   table (Table S1) and a data set (Dataset S1). Figure S1 presents data concerning the fraction of

29   computer-identified candidate seismic events for which analysts could pick P and S arrivals. Figure S2

30   presents information used to define a magnitude scale for seismic events in Dataset S1. Figure S3

31   presents time-of-day vs year-of-occurrence information for seismic events in two locations delineated

32   in Figure 1 in the main article. Figures S4 to S8 present annual maps for the years 2000 to 2017 for

33   seismic events, wastewater disposal volumes, produced oil volumes, produced gas volumes, and

34   hydraulic fracture treatment fluid volumes. Dataset S1 is an Excel file presenting the TXAR seismic

35   catalog that resulted from the analysis in this investigation.

36   **Computer-based event detection methodology**

37   As noted in the Methods section, we applied a multi-stage F-K filtering process to identify

38   candidate events. We preprocessed the data by applying a bandpass filter with corners at 2 Hz and 10

39   Hz to the 40 sample/s data stream. We then downsampled the result with an anti-alias filter at 7.5 Hz

40   to 15 samples/s. The F-K spectra were calculated in a moving 8.5 second (128 sample) window

41   updated every 2.5 seconds. To detect candidate events, we applied three different detector

42   algorithms based upon the F-K spectra. Candidate events were signals that met the detection

43   thresholds for any one of the three detector algorithms.

44   Detector 1 triggered whenever the Fisher ratio (F-ratio), calculated from the F-K spectra, was above

45   1.4. Detector 1 overrode all other detectors.

46   Detector 2 also used the F-ratio as a metric, but a threshold of 1.05 rather than 1.4 for phase

47   slowness above 0.1 sec/km and back azimuths in the +/- 13 degree range. This resulted in slightly

48   higher sensitivity for events in the Delaware Basin at the expense of slightly more false alarms. In

49   practice, most of the events that trigger detector 2 without triggering the other detectors were too

50   small to be located using just TXAR.

51   Detector 3 is based upon the F-ratio normalized power calculated as:

52    $$Kmax_t = 20Log_{10}max_{ij}(abs(S_{ijt})) \hspace{3cm} \text{(S1)}$$

53   where S_{ijt} is the F-K spectra value for slowness *i, j* and time *t*. Recall that *t* is at the F-K spectra

54   sampling rate of 1/2.5 samples/second. The slowness grid was +/- 0.16 sec/km every 0.005 sec/km in

55   both east-west and north-south directions. Then a long term average median power was calculated

56   as:

57 $$Kmed_T = median(Kmax_t); \quad T - ltalen/2 < t < T + ltalen/2 \quad\quad\text{(S2)}$$

58 where T is the output time, sampled at 1/2.5 samples/second, and $ltalen$ is the long-term averaging

59 length of 480 seconds. A long term average variance $Kvar_T$ is calculated as the 480 second median

60 difference between $Kmax_T$ and $Kmed_T$, i.e:

61 $$Kvar_T = median(Kmax_T - Kmed_T). \quad\quad\text{(S3)}$$

62 The normalized power, $Knorm_T$ is

63 $$Knorm_T = (Kmax_T - Kmed_T)/Kvar_T \quad\quad\text{(S4)}$$

64 and the medium term average $Kmta_T$ is

65 $$Kmta_T = median(Kmax_t); \quad T - mtalen/2 < t < T + mtalen/2 \quad\quad\text{(S5)}$$

66 where the medium-term averaging length $mtalen$ is 20 seconds.

67 At each station in the Lajitas array a noise compensation parameter $N_{i\underset{\wedge}{T}}$ is calculated based upon

68 envelope power as:

69 $$N_{i\underset{\wedge}{T}} = \sum_{t>\underset{\wedge}{T}-\frac{stalen}{2}}^{t<\underset{\wedge}{T}+\frac{stalen}{2}} abs(hilbert(x_{i\underset{\wedge}{t}}))^2/stalen \quad\quad\text{(S6)}$$

70 for station i and sample $\underset{\wedge}{t}$ of the time series station timeseries $x_{i\underset{\wedge}{t}}$, and where the short-term averaging

71 length $stalen$ is 5 sec. The T and $\underset{\wedge}{t}$ are then sampled at the 15 SPS time series rate. Then the

72 normalized envelope power $\boldsymbol{Np}_{i\underset{\wedge}{T}}$ is

73 $$\boldsymbol{Np}_{i\underset{\wedge}{T}} = N_{i\underset{\wedge}{T}} / \underline{N_{\underset{\wedge}{T}}} \quad\quad\text{(S7)}$$

74 where $\underline{N_{\underset{\wedge}{T}}}$ is the mean of $\boldsymbol{Np}_{i\underset{\wedge}{T}}$ over all array stations. The peak power difference $Npeak_{T\underset{\wedge}{T}}$ is

75 $$Npeak_{T\underset{\wedge}{T}} = max_i(\boldsymbol{Np}_{i\underset{\wedge}{T}}) - min_i(\boldsymbol{Np}_{i\underset{\wedge}{T}}) \quad\quad\text{(S8)}$$

76 and the LTA peak power difference $Nlta_{\underset{\wedge}{T}}$ is

77 $$Nlta_{T\underset{\wedge}{T}} = median(Npeak_{T\underset{\wedge}{T}}), \underset{\wedge}{T} - ltalen/2 < \underset{\wedge}{t} < \underset{\wedge}{T} + ltalen/2. \quad\quad\text{(S9)}$$

78 Finally the scaled noise adjustment $Nadj_{\underset{\wedge}{T}}$ is

79 $$Nadj_{T\underset{\wedge}{T}} = Nlta_{T\underset{\wedge}{T}}/ median(Nlta_{T\underset{\wedge}{T}}). \quad\quad\text{(S10)}$$

80   The detection metric D_T for detector 3 is then

81      $D_T = max_{ij}(F_{ijT}) \cdot Kmta_T / \sqrt{Nadj_T}$ (S11)

82   where F_{ijT} is the Fisher ratio calculated from the F-K spectra.

83   The third detector triggers whenever D_T exceeds 1.4. Detector 3 accounted for most of the

84   detections. Detectors 1 and 2 provide for detections that have high coherence even if the total power

85   may be too low to trigger detector 3.

86
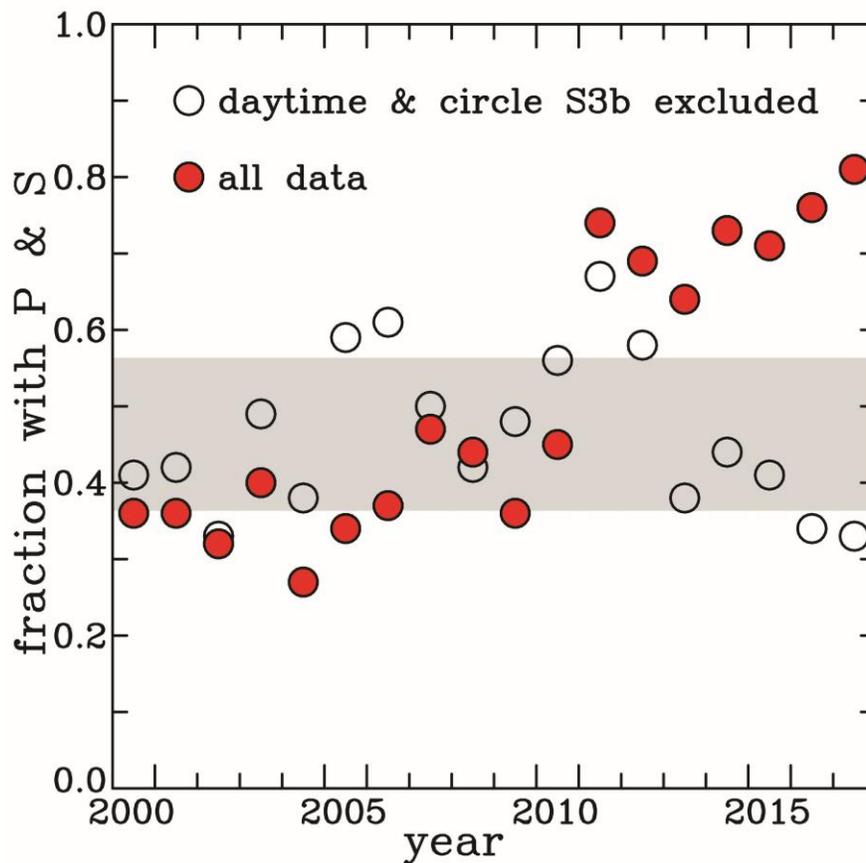
Figure S1. Annual fraction of seismic events detected by the computerized multi-stage filtering process for which analysts could pick both P and S arrivals. For all events having magnitude $M_{TXAR} > 1.0$ (red circles), the fraction detected averages 0.38 for years 2000-2010, but 0.73 for 2011-2017. However, if we exclude daytime events (13:00-01:00 UTC or 7 AM to 7PM CST) and events with computer-determined locations within the Pecos Cluster Circle S3b, the fraction picked averages 0.46±0.10 (gray band) for the entire period.
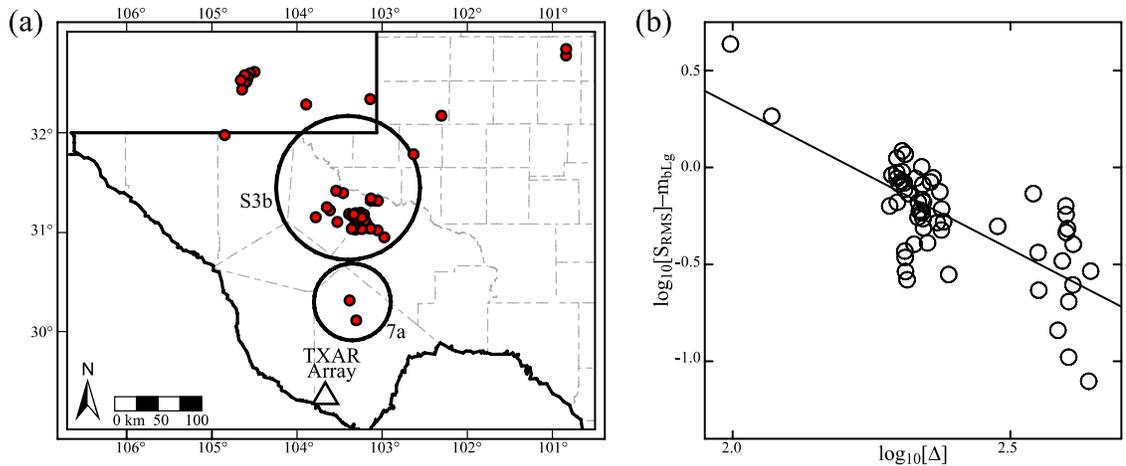
Figure S2. (a) Earthquakes in Dataset S1 with locations and m_{bLg} reported by the USGS. Red circles are 59 events used to determine a magnitude scale of the form M_{TXAR} = $\log_{10}S_{RMS}$ + $A\log_{10}\Delta$ + B that minimized the squared differences between M_{TXAR} and m_{bLg}. **(b)** For these earthquakes, plot of ($\log_{10}S_{RMS}$ - m_{bLg}) vs $\log_{10}\Delta$. Straight line is best-fitting relationship where A = 1.48 and B = -3.29.

(a)

(b)
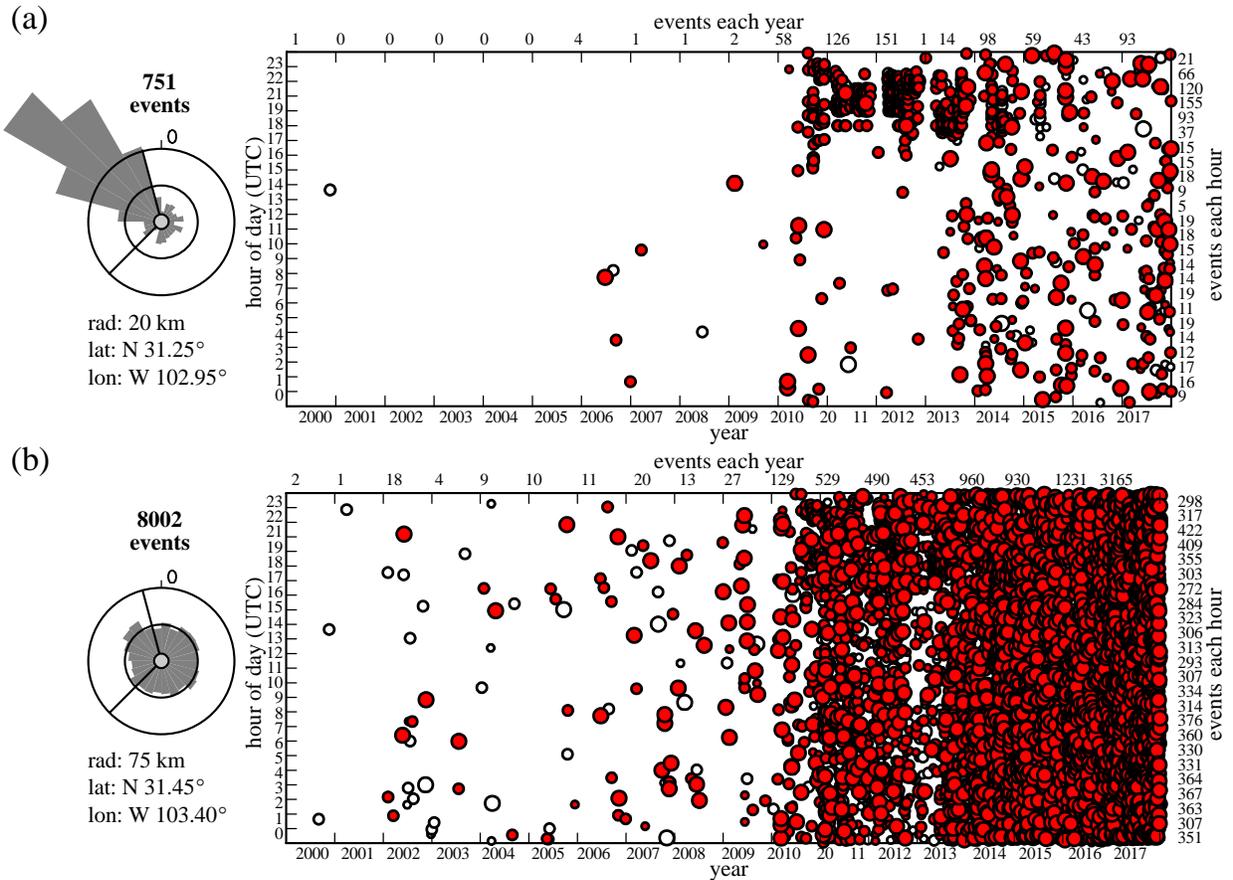
101
102   **Figure S3.** Time-of-day vs year-of-occurrence for selected seismic events. See caption
103   to Figure 7 for plot description. **(a)** Events within the 20-km-radius circle labeled S3a in
104   Figure 1a. The majority of events occur between 1 PM and 5 PM local time suggesting
105   they are caused by human activities such as quarrying. The highest activity rates are in
106   the daytime between 2010 and 2014. **(b)** Events within the 75-km-radius circle labeled
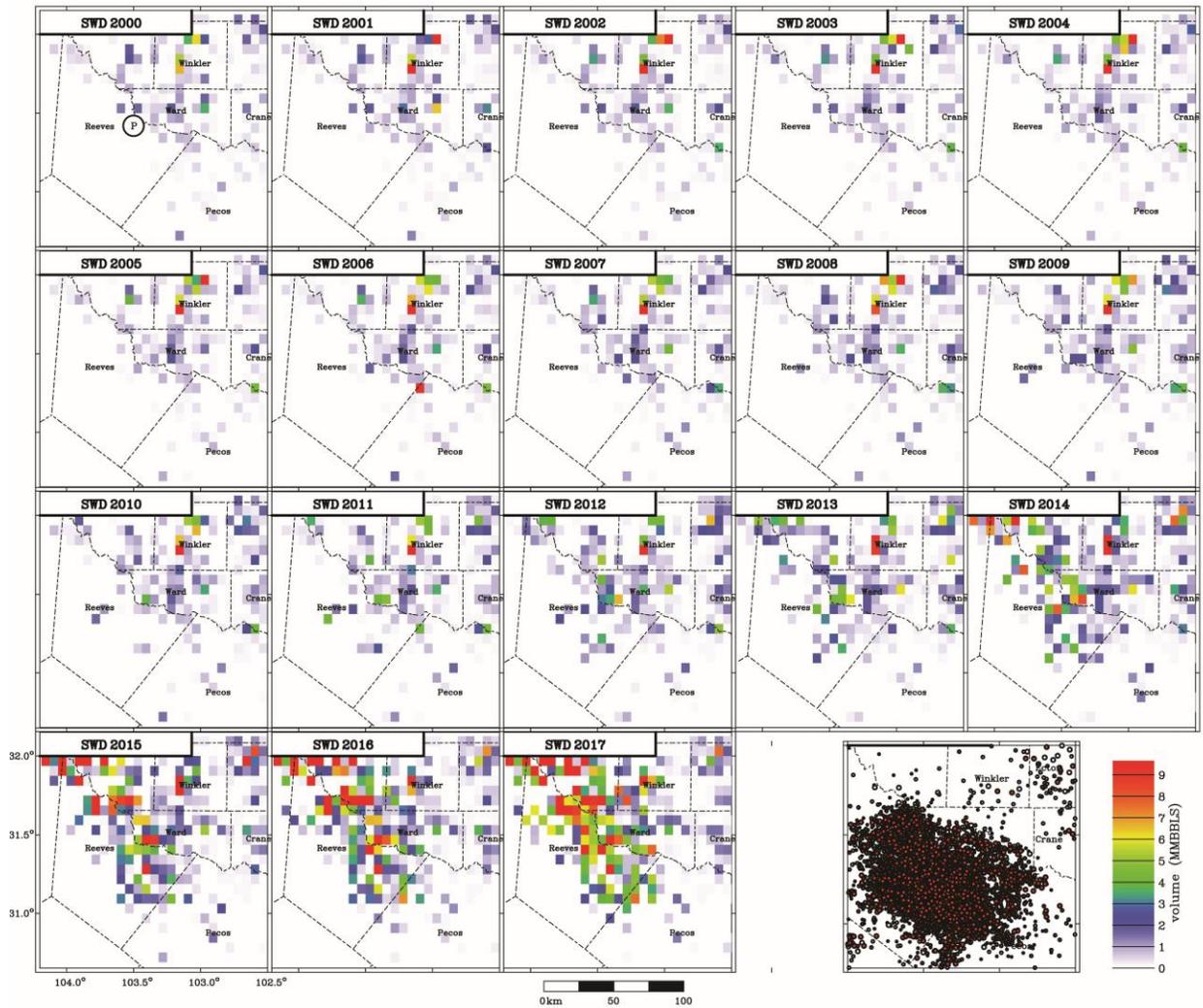107   S3b in Figure 1a. This circle also includes the events in both circles labeled 7c and S3a.
108

109
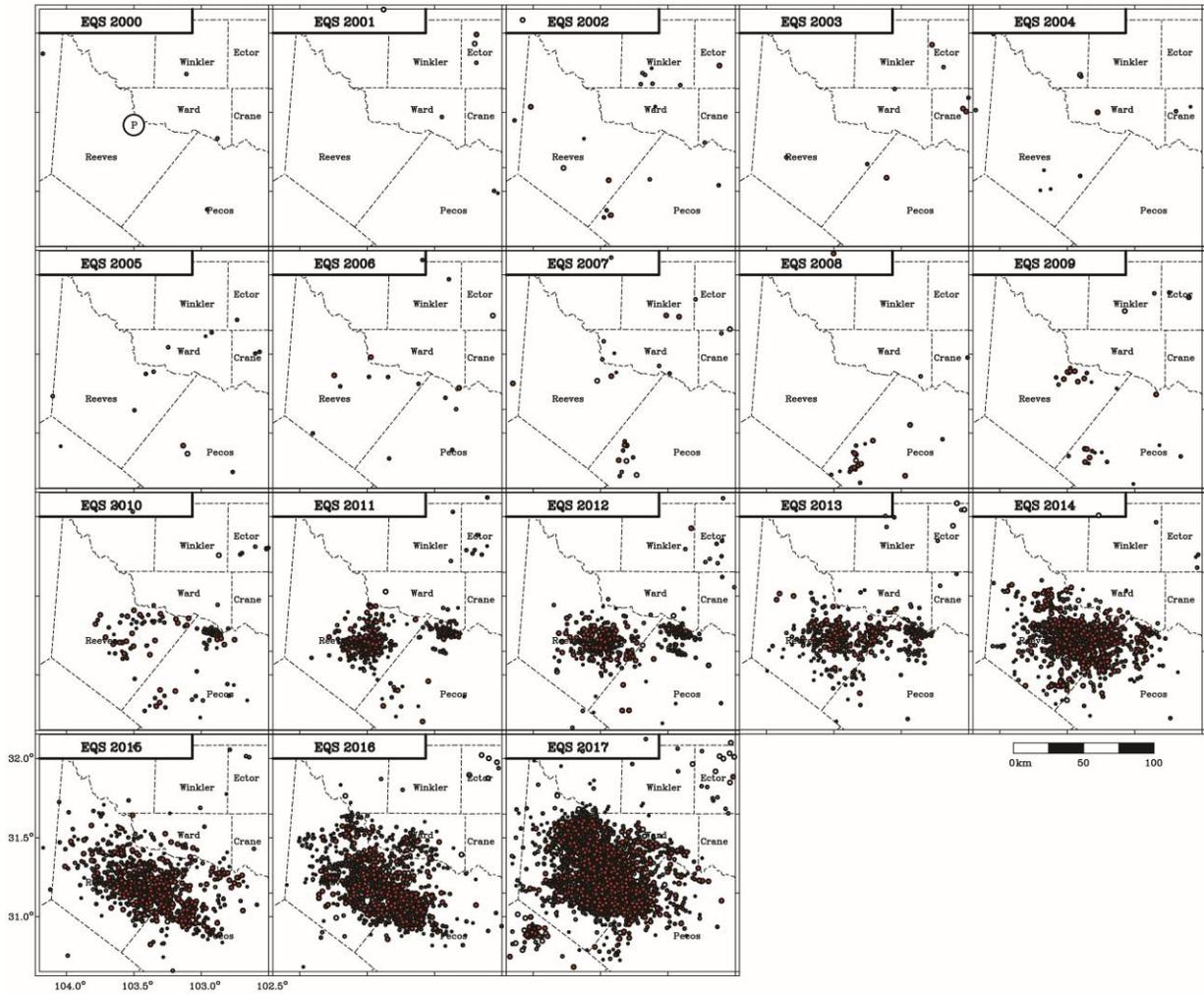
110
111   **Figure S4.** Maps of TXAR-determined epicenters (EQS) near the town of Pecos, 2000-
112   2017. The circle labeled "P" on the 2000 map shows the location of Pecos.
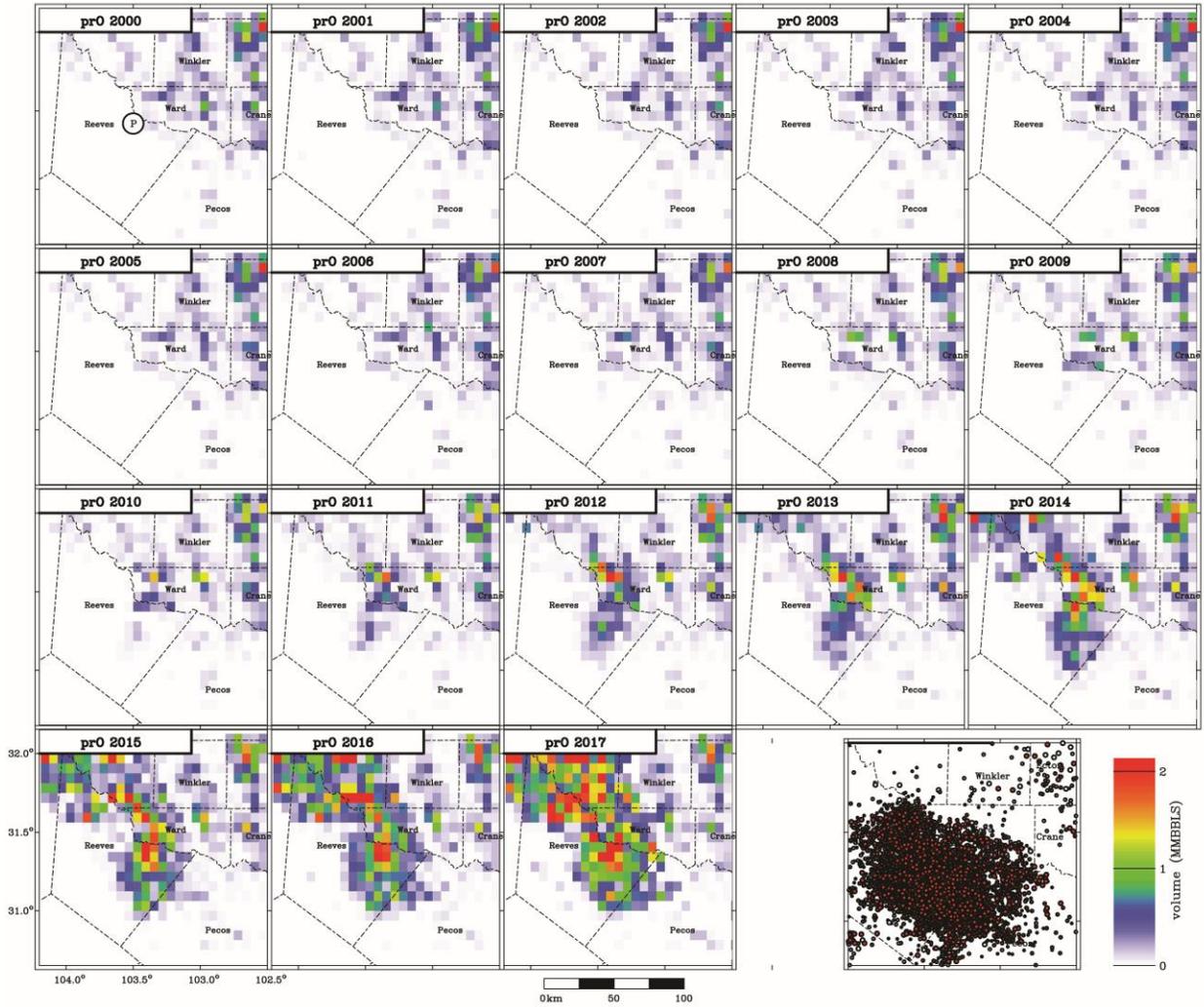
113

114
115 **Figure S5.** Maps of wastewater disposal volumes (SWD) near Pecos, 2000-2017. The
116 circle labeled "P" on the 2000 map shows the location of Pecos. For comparison, map
117 at lower right presents TXAR catalog seismicity 2000-2017.
118
119

120

121
122 **Figure S6.** Maps of produced oil volumes (prO) near Pecos, 2000-2017. For
123 comparison, map at lower right presents TXAR catalog seismicity 2000-2017. The circle
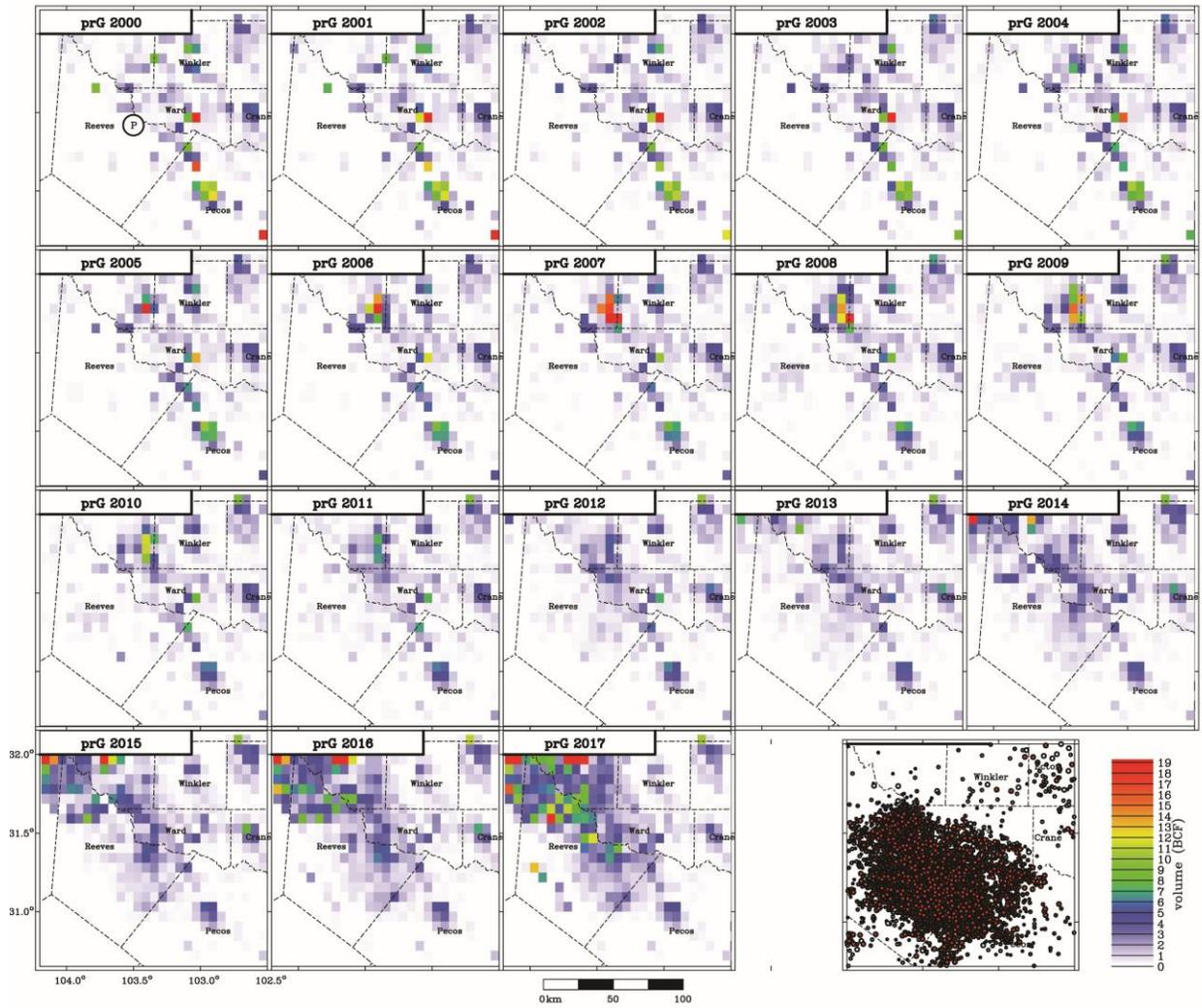124 labeled "P" on the 2000 map shows the location of Pecos.
125
126
127

128
129

130
131 **Figure S7.** Maps of produced gas volumes (prG) near Pecos, 2000-2017. For
132 comparison, map at lower right presents TXAR catalog seismicity 2000-2017. The circle
133 labeled "P" on the 2000 map shows the location of Pecos.
134
135

136
137

Figure S8. Maps of hydraulic fracturing treatment fluid (TrF) volumes near Pecos, 2000-2017. For comparison, map at lower right presents TXAR catalog seismicity 2000-2017. The circle label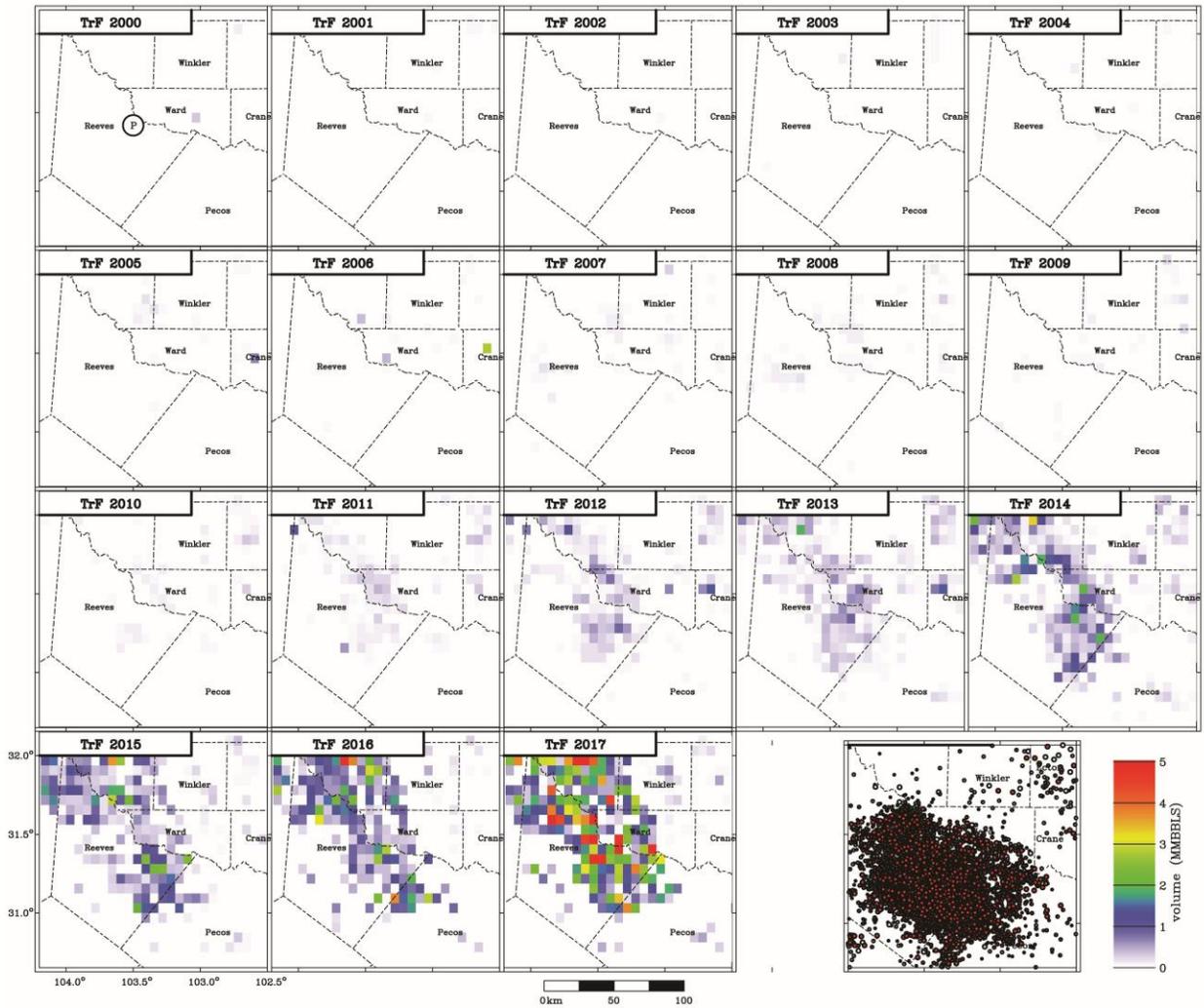ed "P" on the 2000 map shows the location of Pecos. Mapped volumes are as reported by FracFocus and may be incomplete, especially in the earlier years mapped.

145
146   **Dataset S1. TXAR seismic catalog.** This Excel file has 10,753 seismic events in West
147 Texas and New Mexico with locations determined only using data from the TXAR
148 seismic array in Lajitas, TX. Data entry columns are:
149 A. OriginTXAR – origin time
150 B. latTXAR – latitude (degrees)
151 C. lonTXAR – longitude (degrees)
152 D. magTXAR – magnitude (see eq. (2) in text)
153 E. back azimuthTXAR – back azimuth from station TX31 (degrees)
154 F. distTXAR – distance from station TX31 (km)
155 G. PonsetTXAR – P time picked by computer algorithm
156 H. Panalyst – P time picked by analyst (clock time, seconds only)
157 I. diffP – difference between analyst and machine P picks (sec)
158 J. Sanalyst – S time picked by analyst (clock time, seconds only)
159 K. SmPanalyst – (S-P) interval (sec); difference between Sanalyst and Panalyst
160 L. QP – quality factor for P pick, assigned by analyst
161 M. QS – quality factor for S pick, assigned by analyst
162 N. Srms – root mean square velocity amplitude (nanometers/s) for S, assigned by
163   computer
164 O. Speak – peak amplitude (nanometers/sec) for S, assigned by computer
165 P. other location key – 1=double-difference location determined by C. Aiken using
166   TexNet data; 2=location determined by J. Walter using EarthScope data;
167   3=USGS location; 4=TexNet regular location
168 Q. magOther – magnitude assigned by other locating organization
169 R. latOther- latitude assigned by other locating organization
170 S. lonOther - longitude assigned by other locating organization
171 T. originOther – origin time assigned by other locating organization
172 U. IDOther – identification code assigned by locating agency
173
174

175   Table S1. Annual statistics for TXAR events detected by computerized multi-stage filtering

176   process, and events with P and S picked by analysts. Columns labeled 'all' are numbers of

177   events detected; number with $M_{TXAR} \geq 0.8$, and number with $M_{TXAR} \geq 1.0$. Columns labeled

178   'good' are numbers with analyst-selected P & S picks and specified magnitudes. Columns

179   labeled 'fraction' are fractions of all events having analyst picks and specified magnitudes.

180

All data							
year	all	all M≥0.8	all M≥1.0	good M≥0.8	good M≥1.0	fraction M≥0.8	fraction M≥1.0
2000	499	234	168	76	61	0.32	0.36
2001	932	414	298	137	107	0.33	0.36
2002	929	427	294	111	95	0.26	0.32
2003	928	401	283	126	114	0.31	0.40
2004	797	438	346	110	93	0.25	0.27
2005	933	492	390	156	131	0.32	0.34
2006	818	385	273	117	101	0.30	0.37
2007	874	339	225	127	106	0.37	0.47
2008	961	424	289	175	127	0.41	0.44
2009	846	420	308	132	112	0.31	0.36
2010	1269	638	492	253	223	0.40	0.45
2011	2101	1007	748	699	553	0.69	0.74
2012	1704	836	613	539	425	0.64	0.69
2013	1694	816	582	493	375	0.60	0.64
2014	2719	1338	916	941	671	0.70	0.73
2015	2708	1329	934	934	661	0.70	0.71
2016	3166	1555	1064	1168	806	0.75	0.76
2017	7235	3469	2222	2728	1801	0.79	0.81
Night events only (01:00-13:00 UTC), and Pecos Cluster Circle S3b events <u>excluded</u>							
year	all	all M≥0.8	all M≥1.0	good M≥0.8	good M≥1.0	fraction M≥0.8	fraction M≥1.0
2000	214	67	44	24	18	0.36	0.41
2001	463	141	86	42	36	0.30	0.42
2002	439	134	82	32	27	0.24	0.33
2003	418	111	73	39	36	0.35	0.49
2004	332	135	106	49	40	0.36	0.38
2005	402	144	109	76	64	0.53	0.59
2006	238	63	46	29	28	0.45	0.61
2007	277	71	42	25	21	0.35	0.50
2008	281	101	62	38	26	0.38	0.42

2009	332	123	84	49	40	0.40	0.48
2010	379	120	87	52	49	0.43	0.56
2011	413	197	155	116	104	0.59	0.67
2012	272	70	38	28	22	0.40	0.58
2013	305	98	63	31	24	0.32	0.38
2014	370	116	71	38	31	0.33	0.44
2015	384	118	74	37	30	0.31	0.41
2016	394	113	76	35	26	0.31	0.34
2017	623	204	125	50	41	0.25	0.33

Night events only (01:00-13:00 UTC),
and including <u>only</u> events within Pecos Cluster Circle S3b

year	all	all $M \geq 0.8$	all $M \geq 1.0$	good $M \geq 0.8$	good $M \geq 1.0$	fraction $M \geq 0.8$	fraction $M \geq 1.0$
2000	29	8	5	4	3	0.50	0.60
2001	91	24	13	9	5	0.38	0.38
2002	114	32	16	11	9	0.34	0.56
2003	72	30	20	13	11	0.43	0.55
2004	73	32	21	13	11	0.41	0.52
2005	100	54	38	29	22	0.54	0.58
2006	68	30	21	14	11	0.47	0.52
2007	64	20	12	9	7	0.45	0.58
2008	59	20	15	11	8	0.55	0.53
2009	75	26	18	14	10	0.54	0.56
2010	153	64	53	47	38	0.73	0.72
2011	559	181	110	162	98	0.90	0.89
2012	488	165	102	136	88	0.82	0.86
2013	525	190	110	157	94	0.83	0.85
2014	1056	440	276	385	245	0.88	0.89
2015	998	439	285	385	248	0.88	0.87
2016	1359	589	385	530	345	0.90	0.90
2017	3659	1518	934	1369	853	0.90	0.91

181