Bulletin of the Seismological Society of America, Vol. 103, No. 5, pp. 2850–2861, October 2013, doi: 10.1785/0120130017

Modeling Earthquake Rate Changes in Oklahoma and Arkansas: Possible Signatures of Induced Seismicity

by Andrea L. Llenos and Andrew J. Michael

Abstract The rate of $M_L \geq 3$ earthquakes in the central and eastern United States increased beginning in 2009, particularly in Oklahoma and central Arkansas, where fluid injection has occurred. We find evidence that suggests these rate increases are man-made by examining the rate changes in a catalog of $M_L \geq 3$ earthquakes in Oklahoma, which had a low background seismicity rate before 2009, as well as rate changes in a catalog of $M_L \geq 2.2$ earthquakes in central Arkansas, which had a history of earthquake swarms prior to the start of injection in 2009. In both cases, stochastic epidemic-type aftershock sequence models and statistical tests demonstrate that the earthquake rate change is statistically significant, and both the background rate of independent earthquakes and the aftershock productivity must increase in 2009 to explain the observed increase in seismicity. This suggests that a significant change in the underlying triggering process occurred. Both parameters vary, even when comparing natural to potentially induced swarms in Arkansas, which suggests that changes in both the background rate and the aftershock productivity may provide a way to distinguish man-made from natural earthquake rate changes. In Arkansas we also compare earthquake and injection well locations, finding that earthquakes within 6 km of an active injection well tend to occur closer together than those that occur before, after, or far from active injection. Thus, like a change in productivity, a change in interevent distance distribution may also be an indicator of induced seismicity.

Introduction

The rate of $M_L \geq 3$ earthquakes in the central and eastern United States increased substantially beginning in 2009, particularly in regions such as Oklahoma, the New Mexico–Colorado border, and central Arkansas where fluid injection associated with oil and gas production occurs (Ellsworth *et al.*, 2012; Horton, 2012b). Are these earthquake rate changes in fact man-made, are they due to natural transient processes, or are they simply random fluctuations in rate? To investigate this question, we use statistical models and tests to determine whether these rate increases are statistically significant. We will rely on the epidemic-type aftershock sequence (ETAS) model (Ogata, 1988), a stochastic model based on empirical aftershock scaling laws such as the Omori law and the Gutenberg–Richter magnitude distribution, to establish whether the rate changes are due to increases in background seismicity rate, aftershock productivity, or some combination of these effects. The results may provide a way to distinguish between natural and anthropogenic earthquake rate changes.

This study focuses on Oklahoma and central Arkansas. From 1975 to 2008, a total of 40 $M_L \geq 3$ earthquakes occurred in Oklahoma, resulting in a low but relatively steady background seismicity rate of roughly one to two events per year. The seismicity rate began to increase in 2009 and continues to the present, but much of this increase is concentrated in the central part of the state (Fig. 1). During this period of heightened seismicity rate, an M_L 5.6 occurred in November 2011 east of Oklahoma City, the largest event in the state to occur in the last century. There has been some evidence to suggest that this earthquake was induced by wastewater injection activity at nearby fluid injection wells (Horton, 2012a; Keranen *et al.*, 2013).

We also examine seismicity rate changes in central Arkansas where, unlike Oklahoma, earthquake swarms have occurred prior to the start of fluid injection in 2009 (Fig. 2). Earthquake swarm activity began in January 1982 near the town of Enola (Johnston, 1982). Over the next three years, over 30,000 $M_L < 4.5$ earthquakes were recorded, potentially triggered by magma intrusion or natural fluid migration (Johnston, 1982; Chiu *et al.*, 1984). By 1987, seismicity had died down, with only 12 earthquakes being recorded in the area over a four-month period (Pujol *et al.*, 1989), and the low level of activity continued through the 1990s. In May 2001, another earthquake swarm began near Enola, consisting of 2500 $M_L < 4.4$ earthquakes over two months, which also may have been triggered by natural fluid migration

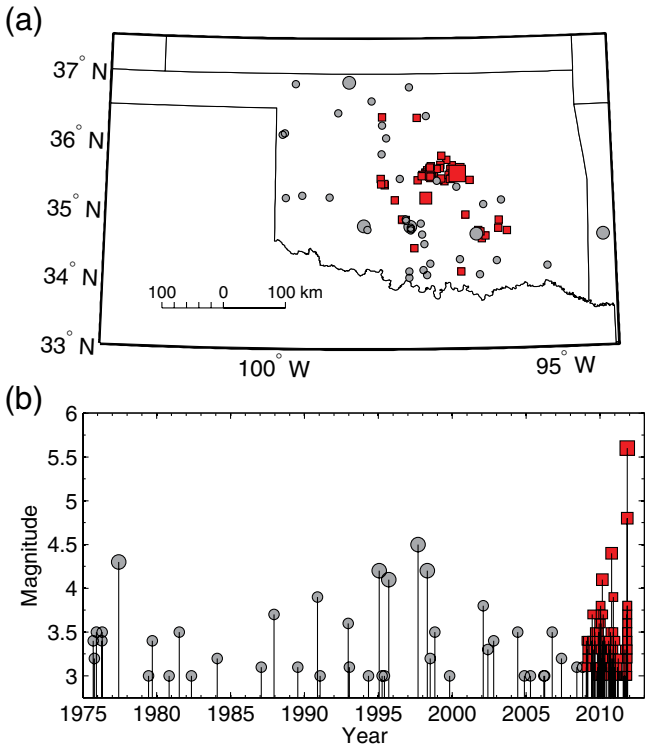

Figure 1. (a) Oklahoma $M_L \geq 3$ seismicity from the USGS PDE catalog. Events from 1975 to 2008 are shown by circles, events from 2009 to 2011 are shown by squares. Marker size indicates magnitude. (b) Magnitudes and occurrence times of seismicity data. The color version of this figure is available only in the electronic edition.

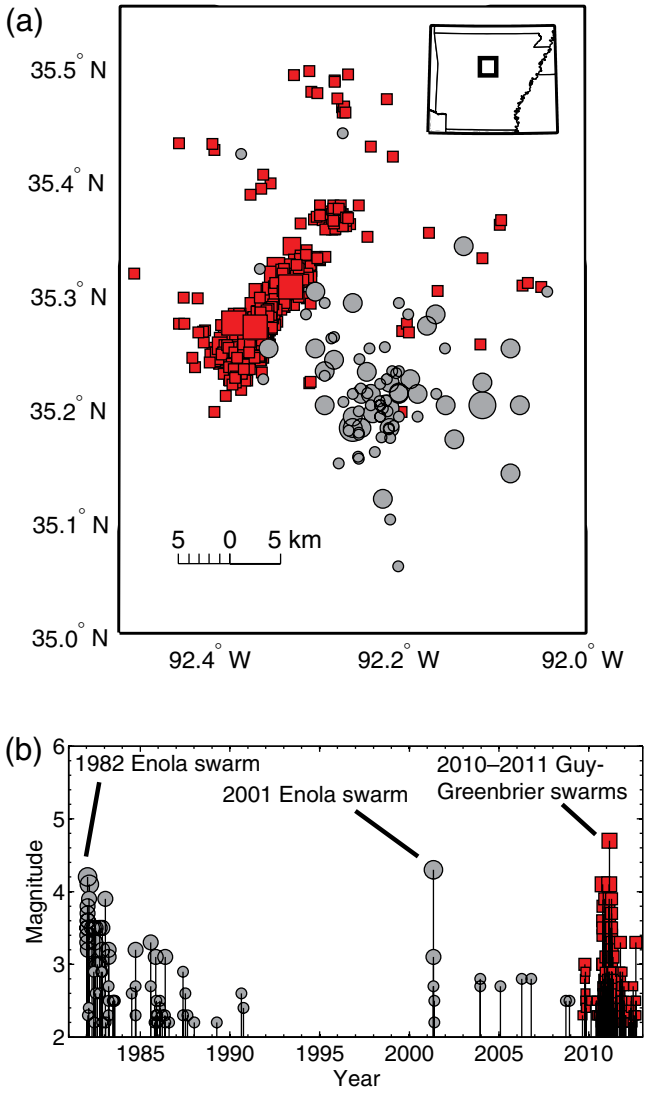

Figure 2. (a) Central Arkansas $M_L \geq 2.2$ seismicity from the Advanced National Seismic System (ANSS) catalog. Events from 1982 to 2008 (which occur prior to the start of wastewater injection) are shown by circles, events from 2009 to 2012 are shown by squares. Marker size indicates magnitude. (b) Magnitudes and occurrence times of seismicity data, with significant swarms labeled. The color version of this figure is available only in the electronic edition.

(Rabak *et al.*, 2010). Following another period of relatively little seismicity, the earthquake rate once again increased significantly in 2010–2011, when another series of swarms occurred near the towns of Guy and Greenbrier. These swarms were likely induced by wastewater injection, which began in the area in April 2009 and was halted in 2011 because of the swarms (Horton, 2012b).

In this paper, we fit the ETAS model to earthquake catalogs from Oklahoma and Arkansas to investigate the significance and possible causes of the rate increases that are observed there beginning in 2009. To determine whether an earthquake rate increase is due to a change in background seismicity rate, aftershock productivity, or some combination of the two, we do the following: (1) fit the ETAS model parameters to the data, (2) convert origin times to transformed times (Ogata, 1992), and (3) use Runs (Wald and Wolfowitz, 1940), autocorrelation function (ACF), and Kolmogorov–Smirnov (KS) tests on interevent times to test the null hypothesis that the transformed times are drawn from a Poisson distribution with constant rate (as expected when no external processes trigger earthquakes besides a constant tectonic loading rate). We also test models that allow different sets of parameters to vary around a change point in 2009. Additionally, in Arkansas we combine earthquake and injection well data to explore spatiotemporal relationships between seismicity and fluid injection activity. These tests may provide a way to distinguish man-made from natural earthquake rate changes and provide insights into the physical mechanisms of induced seismicity.

Data and Methods

To investigate the observed rate change in Oklahoma, we analyze earthquake occurrence times and magnitudes from the U.S. Geological Survey (USGS) Preliminary Determination of Epicenters (PDE) catalog of $M_L \geq 3$ events occurring from January 1975 to November 2011 (Fig. 1; see Data and Resources). Using the maximum curvature (MAXC) algorithm (Wiemer *et al.*, 1998; Wiemer and Wyss, 2000), we

estimate the magnitude of completeness to be 2.7 ± 0.34 but choose a conservative threshold of $M_c = 3$. When the analysis is repeated with a slightly higher completeness threshold, the results are similar, suggesting that this does not bias the results. We will compare the results from Oklahoma to those obtained from an Advanced National Seismic System (ANSS) catalog of $M_L \geq 2.2 \geq$ events that occurred in central Arkansas from January 1982 to January 2012 (Fig. 2; see Data and Resources). The ANSS catalog in Arkansas has a lower magnitude of completeness, estimated with the MAXC algorithm to be 2.1 ± 0.06, which allows us to include more events in our statistical analysis.

The ETAS model is a point process model that describes the seismicity rate observed in a region as a summation of a background rate of independent events and the aftershocks triggered by each event (Ogata, 1988). Each earthquake can trigger its own aftershock sequence; therefore, given the occurrence time t_i and magnitude M_i of each event i that occurred prior to time t, and the magnitude of completeness M_c of the catalog, the earthquake occurrence rate λ at time t is

$$\lambda(t) = \mu + \sum_{\{i:t_i<t\}} \frac{K e^{\alpha(M_i - M_c)}}{(t - t_i + c)^p}, \quad (1)$$

in which μ is the background seismicity rate, K is the aftershock productivity, α determines how efficiently an earthquake of a given magnitude triggers aftershocks, and c and p are the Omori–Utsu parameters that describe the temporal decay of seismicity rate following a mainshock. The parameter c is a time offset between the mainshock and when the power law decay in rate is observed, and the parameter p governs how quickly the rate decays (Omori, 1894; Utsu, 1961). The parameters μ, K, α, c, and p are optimized to fit a catalog of events occurring from T_s, the start of the time span of interest, to T_e, the end of the time span of interest, by maximizing the point process likelihood L using the following function (Daley and Vere-Jones, 2002):

$$\ln L(\mu, K, c, \alpha, p) = \sum_{\{i:t_i<T_e\}} \ln \lambda(t_i) - \int_{T_s}^{T_e} \lambda(t) dt. \quad (2)$$

The fit of an optimized ETAS model to a catalog can be tested by first converting the occurrence times t_i to transformed times τ_i using the following theoretical cumulative function (Ogata, 1992):

$$\tau_i = \int_0^{t_i} \lambda(t) dt. \quad (3)$$

The transformed time τ_i thus represents the number of events that the optimized ETAS model predicts should have occurred in the time interval $[0, t_i]$. If the optimized ETAS model fits the seismicity well, then the transformed times behave as a Poisson process with constant unit rate. Therefore, a plot of the observed cumulative number of events versus transformed times should be linear with unit slope (i.e., $i = \tau_i$, or the ith event occurs exactly when the model predicts it should). Positive and negative deviations from the line indicate that the model is under- and overpredicting seismicity, respectively.

Moreover, if the transformed times are drawn from a Poisson distribution, then the transformed interevent times $\Delta\tau_i = \tau_{i+1} - \tau_i$ should be independent and identically drawn from an exponential distribution. To test this hypothesis, we use a suite of statistical tests, including the Runs test, ACF test, and KS test. The Runs test can be used to check the independence of the interevent times by comparing the distribution of runs (sets of sequential values that are either all above or all below the median) in the sequence, which should be approximately normal (Wald and Wolfowitz, 1940). The ACF test also checks for the presence of temporal correlation in the interevent times by computing the correlation of the series of interevent times with itself at different time lags and determining whether significant correlation exists at any lags. We use the KS test to check the goodness-of-fit of the interevent times to an exponential distribution, applying a Lilliefors correction based on Monte Carlo simulation to account for the mean of the distribution being unknown *a priori* and therefore estimated from the data (Lilliefors, 1969). Similar suites of tests have been used to evaluate the fit of a stationary ETAS model to fluid-triggered changes in earthquake rate (Lombardi *et al.*, 2010).

We can compare how well different models fit the same data set by computing the Akaike information criterion (AIC) for each model (Akaike, 1974). The AIC statistic is based on the likelihood L and accounts for models with different numbers of fitted parameters k:

$$\text{AIC} = -2 \ln L + 2k. \quad (4)$$

The model with the lower AIC is considered to be the best model, and differences of ~2 are considered to be significant at the 5% level (Hainzl and Ogata, 2005).

Results

We proceed by first fitting the ETAS model to the full-length catalogs in Oklahoma and Arkansas. The maximum likelihood estimates of the ETAS parameters fit to the 1975–2011 Oklahoma catalog are shown in Table 1. Visually, the ETAS model fits the catalog poorly (Fig. 3a,b). We use the Runs, ACF, and KS tests to quantitatively evaluate the fit of the ETAS model: if any of the three tests rejects the null hypothesis (i.e., p-value < 0.05), then we consider the model a poor fit to the data. In Oklahoma, the null hypothesis that the transformed interevent times are independent is rejected by the Runs test (p-value $= 0.0009$) and the ACF test, which indicates that the interevent times are temporally correlated. These results suggest that a single set of ETAS parameters cannot fit the entire 1975–2011 earthquake catalog, providing evidence that a significant rate change occurred

Table 1
Oklahoma ETAS Maximum Likelihood Estimation (MLE) Results

Data Set	μ (Events/Day)	K (Events/Day)	c (Days)	α	P	AIC
1973–2011	0.001 ± 0.009	0.026 ± 0.021	0.007 ± 0.06	1.46 ± 0.1	0.79 ± 0.02	1097.96
1973–2008	0.003 ± 0.004	$3 \times 10^{-17} \pm 0.03$	0.001 ± 0.86	2.21 ± 0.47	0.72 ± 0.09	553.61
μ estimated*	0.002 ± 0.005	0.018	0.02	1.88	0.96	571.99
2009–2011	0.046 ± 0.03	0.018 ± 0.03	0.02 ± 0.13	1.88 ± 0.1	0.96 ± 0.09	492.44
μ estimated†	0.095 ± 0.02	3×10^{-17}	0.001	2.21	0.72	672.08
K estimated†	0.003	0.015 ± 0.006	0.001	2.21	0.72	517.84
μ and K estimated†	0.04 ± 0.02	0.009 ± 0.008	0.001	2.21	0.72	491.84

Uncertainties are 1 standard deviation.
*All other parameters held constant at 2009–2011 estimates to evaluate impact of parameter changes.
†All other parameters held constant at 1973–2008 estimates to evaluate impact of parameter changes.

during that time period. However, the null hypothesis that the transformed interevent times are drawn from an exponential distribution is not rejected by the KS test (p-value = 0.98). This suggests that, unlike the Runs and ACF tests, the KS test may not be sensitive to a change in rate, because the overall distribution is dominated by the majority of the events that occur in the later part of the catalog. When the ETAS model is fit to the 1975–2008 portion of the catalog (Table 1) and extrapolated to predict the seismicity in 2009–2011, the model predicts that three earthquakes should have occurred, but in reality 100 events occurred, an increase by over a factor of 30 (Fig. 3c). Simulations using the 1975–2008 parameters show that there is less than 0.1% chance of observing 15 events or more from 2009 to 2011. The transformed time plot also indicates that the 2009–2011 seismicity is vastly underpredicted by the 1975–2008 model (Fig. 3d). However, ETAS models fit to the 1975–2008 and 2009–2011 parts of the catalog separately explain the seismicity in their respective time periods well (Fig. 3e,f). For the 1975–2008 data, the null hypothesis is accepted by all three hypothesis tests with p-values ranging from 0.63 to 0.94. Similarly for the 2009–2011 data, the null hypothesis is accepted by all three tests with p-values of 0.15–0.93. This suggests that the basic triggering model applies throughout the catalog but the model parameters must change.

Similarly, the ETAS model poorly fits the 1982–2012 Arkansas earthquake catalog, both visually and quantitatively (Table 2; Fig. 4a,b). The null hypothesis is rejected by the Runs test (p-value = 0.015) and the ACF test and accepted by the KS test (p-value = 0.72), again due to the distribution of the later events in the catalog dominating the overall distribution. When the ETAS model is fit to the 1982–2008 part of the catalog and extrapolated through 2012, the seismicity is again underpredicted (243 expected events versus 529 observed; Fig. 4c,d). Simulations show that, when using the 1982–2008 parameters, there is less than 0.1% chance of observing 25 or more earthquakes during 2009–2012. However, a separate ETAS model fit to the 2009–2012 catalog fits the seismicity well (i.e., the null hypothesis is accepted with p-values ranging from 0.38 to 0.91; Fig. 4e,f).

The results of the preceding tests suggest that in both Oklahoma and Arkansas, at least one of the ETAS model parameters needs to change in 2009 in order to account for the increase in seismicity. To determine which parameters need to change the most to explain the seismicity increase in Oklahoma, we use the 1973–2008 ETAS parameters (Table 1) as a base model and evaluate fits to the 2009–2011 catalog for the following cases: (1) varying μ alone, (2) varying K alone, (3) varying both μ and K, and (4) varying all five model parameters. We compute the AIC for each case to evaluate the impact of fitting different sets of parameters to the 2009–2011 Oklahoma seismicity. The results are shown in Figure 5a,b and Table 1. The model in which both the background seismicity rate μ and aftershock productivity K vary is the best model, because the AIC is significantly decreased by estimating both parameters (491.844) rather than either parameter by itself (672.08 and 517.84 for the cases in which μ and K are each estimated, respectively). Adding the other three triggering parameters does not offer a significant improvement. This shows that changes in both the background seismicity rate and triggering properties of central Oklahoma are needed in order to fit the 2009–2011 seismicity data. While the background seismicity rate increases by over an order of magnitude, the aftershock productivity increases by several orders of magnitude. In fact, the aftershock productivity is so low when ETAS is fit to the 1973–2008 data that all of the seismicity prior to 2009 occurs as independent background events rather than as triggered aftershocks. In contrast, using the model where both μ and K vary to fit the 2009–2011 data, 47.5% of the earthquakes that occur from 2009 to 2011 have a greater than 50% chance of having been triggered by a previous event.

In Arkansas, we use the 1982–2008 ETAS parameters as a base model and again compute the AIC for different cases (Table 2). As in Oklahoma, varying the background rate and aftershock productivity is required in order to best fit the 2009–2012 data (Fig. 5c,d), which again indicates that the triggering properties in central Arkansas changed, causing much more triggered seismicity to occur after 2009. However, the best-fitting model also requires changes in the three other triggering parameters (c, α, and p). We therefore

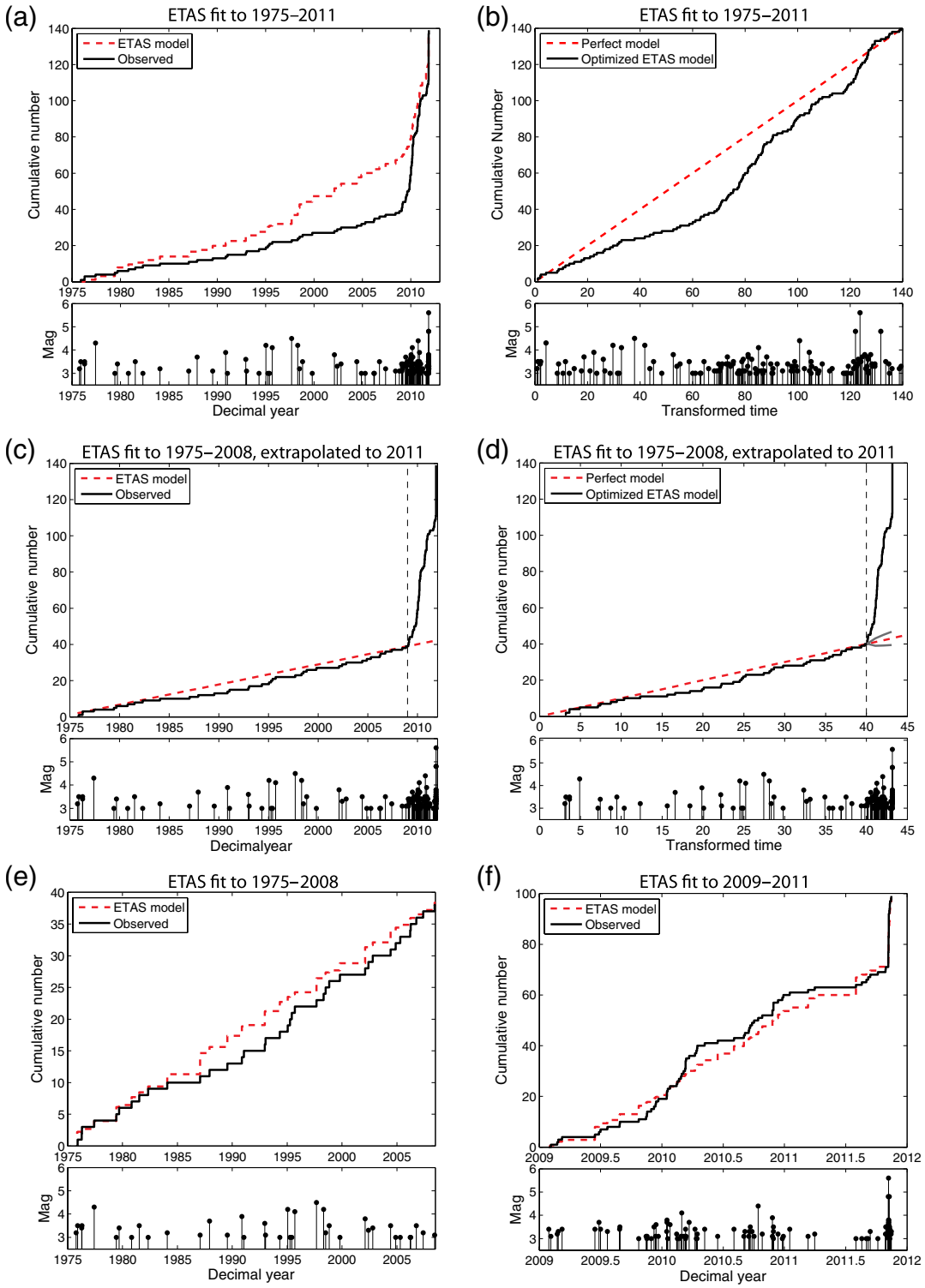

Figure 3. Oklahoma ETAS modeling results. (a) Cumulative number of events versus time observed (solid line) and predicted by ETAS fit from 1975 to 2011 (dashed line). ETAS fits poorly visually and quantitatively. (b) Transformed time plot. The dashed line indicates a perfect model, and the solid line is the ETAS model fit to the data. Less seismicity occurs than the model predicts. (c) Cumulative number of events versus time observed (solid line) and predicted by ETAS fit from 1975 to 2008 and extrapolated through 2011 (dashed). The vertical dashed line indicates the start of extrapolation in 2009. The model underpredicts 2009–2011 seismicity by a factor of 33. (d) Transformed time plot for ETAS fit from 1975 to 2008 and extrapolated through 2011. The vertical dashed line indicates the start of extrapolation; gray lines indicate 2σ bounds. More seismicity occurs than the model predicts. (e) Cumulative number of events versus time observed (solid line) and predicted by ETAS (dashed line) fit to the 1973–2008 catalog. ETAS provides a good fit (all three tests accept the null hypothesis). (f) Cumulative number of events versus time observed (solid line) and predicted by ETAS (dashed line) fit to the 2009–2011 catalog. ETAS provides a good fit (all three tests accept the null hypothesis). The color version of this figure is available only in the electronic edition.

Table 2
Arkansas ETAS Maximum Likelihood Estimation (MLE) Results

Data Set	μ (Events/Day)	K (Events/Day)	c (Days)	α	p	AIC
1982–2008	0.001 ± 0.01	0.028 ± 0.028	0.004 ± 0.003	0.92 ± 0.14	0.96 ± 0.02	680.33
μ estimated*	0.001 ± 0.007	0.082	0.046	1.09	1.27	749.66
2009–2012	0.016 ± 0.025	0.082 ± 0.007	0.046 ± 0.019	1.09 ± 0.03	1.27 ± 0.02	37.17
μ estimated†	0.032 ± 0.026	0.028	0.004	0.92	0.96	342.46
K estimated†	0.001	0.062 ± 0.005	0.004	0.92	0.96	141.57
μ and K estimated†	0.007 ± 0.026	0.061 ± 0.006	0.004	0.92	0.96	117.95
μ, K, α estimated†	0.007 ± 0.026	0.052 ± 0.012	0.004	1.24 ± 0.06	0.96	115.66
μ, K, c estimated†	0.007 ± 0.024	0.062 ± 0.007	0.004 ± 0.007	0.92	0.96	119.65
μ, K, p estimated †	0.011 ± 0.015	0.061 ± 0.006	0.004	0.92	1.05 ± 0.04	89.24
μ, K, α, p estimated †	0.011 ± 0.026	0.055 ± 0.010	0.004	1.12 ± 0.07	1.04 ± 0.01	89.49
μ, K, c, p estimated †	0.016 ± 0.025	0.089 ± 0.012	0.046 ± 0.032	0.92	1.27 ± 0.02	36.50
μ, K, c, α estimated †	0.007 ± 0.026	0.052 ± 0.011	0.005 ± 0.009	1.25 ± 0.06	0.96	117.16

Uncertainties are 1 standard deviation.
*All other parameters held constant at 2009–2012 estimates to evaluate impact of parameter changes.
†All other parameters held constant at 1982–2008 estimates to evaluate impact of parameter changes.

computed the AIC also for cases where three and four parameters were allowed to vary and found that changes in c and p improved the fit the most (Table 2). However, we will focus on changes in μ and K, because changes in these two parameters are required to best fit the data in both Oklahoma and Arkansas, although the relative importance of the parameters differs in the two regions. In Arkansas, changing K is more necessary to fit the data than changing μ, unlike in Oklahoma where changes in both μ and K are important (Fig. 5). This may be because the rate is relatively high in Arkansas because of the earlier Enola swarms, while in Oklahoma there was little seismicity and thus a low μ prior to 2009.

The two standard deviation confidence intervals of the K parameters, between the two time periods, overlap for both Oklahoma and Arkansas (Tables 1 and 2). This is partially due to trade-offs between parameters of the ETAS model within a single data set. Our conclusion that both μ and K changed between the two time periods is based on the reduction in AIC values and is supported by the simulations of the number of events in the later time period when using the parameters from the earlier time period.

We can check the robustness of these results by fitting the ETAS model to the latter part of the catalogs, then applying these fitted triggering parameters (K, c, α, and p) to the earlier parts of the catalog and estimating the background seismicity rate μ. Because of the larger amount of data available post-2009, the model parameters during this time period are better constrained. If applying the post-2009 triggering parameters and allowing μ to vary results in a good fit to the pre-2009 data, this would imply that the triggering parameters are simply poorly constrained by the data available prior to 2009. For both Oklahoma and Arkansas, applying the post-2009 triggering parameters to the pre-2009 catalogs results in estimates of μ that are higher than the pre-2009 estimates but still provide worse fits to the data (Tables 1 and 2). The higher aftershock productivity post-2009 causes the pre-2009 seismic activity to be overestimated. This gives us some confidence that the change in triggering parameters is real and not an artifact of changes in data quantity. We also simulated the number of earthquakes during the earlier time period using the parameters from post-2009 and there is $<0.1\%$ chance of those parameters producing so few events during the earlier time period.

It is possible that the apparent change in aftershock productivity in 2009 is due to biased parameter estimates caused by incorrectly assuming a constant μ after 2009 when in fact it might be time varying due to injection. Hainzl *et al.* (2013) showed that this assumption can lead to an overestimation of K and an underestimation of α. While it is true that this assumption may cause our parameter estimates to be somewhat biased, Figure 5 demonstrates that these parameters still need to change significantly in order to explain the observed rate change. The rate change cannot be explained simply by varying μ in time alone. Moreover, available monthly injection data for the wells nearest the Prague, Oklahoma, sequence do not show a significant change in the time variation of injection rate when comparing before and after 2009 (Keranen *et al.*, 2013). Therefore, there does not seem to be much evidence to support a suddenly time-varying μ following 2009, at least on the monthly scale.

It is also possible that the apparent change in aftershock productivity is caused by a change in the spatial distribution of events. If events are occurring spatially closer together, they could play more of a role in triggering one another. Therefore, changes in spatial distribution might also be diagnostic of induced earthquake rate changes. We can test this by comparing the interevent distance distributions of earthquakes that occur before, during, and after active injection in Arkansas. In the study area, wastewater fluid injection began in April 2009 but stopped following the 2011 swarms (Horton, 2012b). Figure 6a compares the locations of earthquakes and Underground Injection Control (UIC) class II wells obtained from the Arkansas Oil and Gas Commission (see Data and Resources). We calculate the distance between each

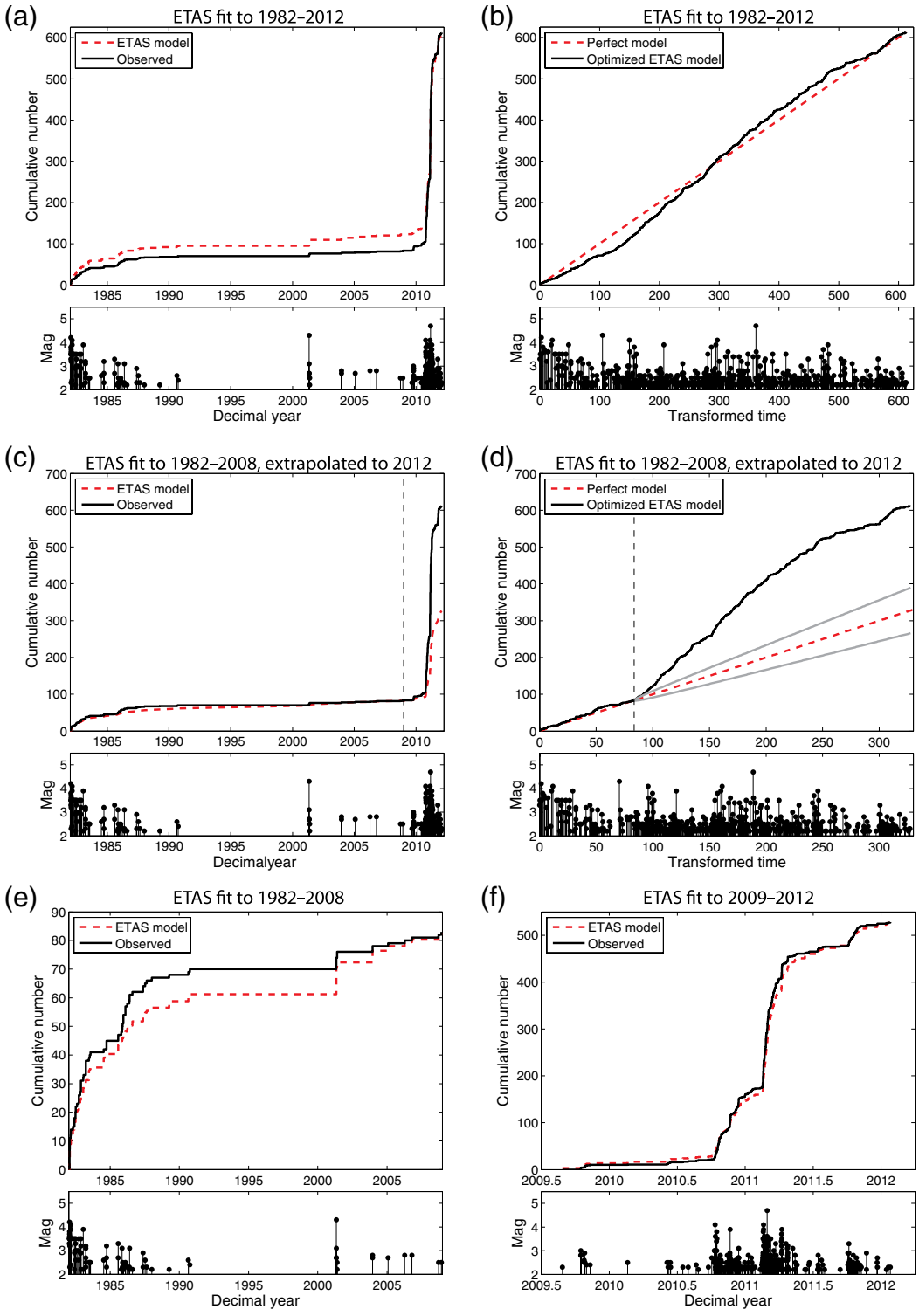

Figure 4. Arkansas ETAS modeling results. (a) Cumulative number of events versus time observed (solid line) and predicted by ETAS fit from 1982 to 2012 (dashed line). ETAS poorly fits, both visually and quantitatively. (b) Transformed time plot. The dashed line indicates a perfect model; the solid line is the ETAS model fit to the data. The model first overpredicts then underpredicts the data. (c) Cumulative number of events versus time observed (solid line) and predicted by ETAS fit from 1982 to 2008 and extrapolated through 2012 (dashed line). The model underpredicts the 2009–2012 seismicity. (d) Transformed time plot for ETAS fit from 1982 to 2008 and extrapolated through 2012. The vertical dashed line indicates the start of extrapolation; gray lines indicate 2σ bounds. More seismicity occurs than the model predicts. (e) Cumulative number of events versus time observed (solid) and predicted by ETAS (dashed) fit to the 1982–2008 catalog. ETAS provides a good fit (all three tests accept the null hypothesis). (f) Cumulative number of events versus time observed (solid line) and predicted by ETAS (dashed line) fit to the 2009–2012 catalog. ETAS provides a good fit (all three tests accept the null hypothesis). The color version of this figure is available only in the electronic edition.

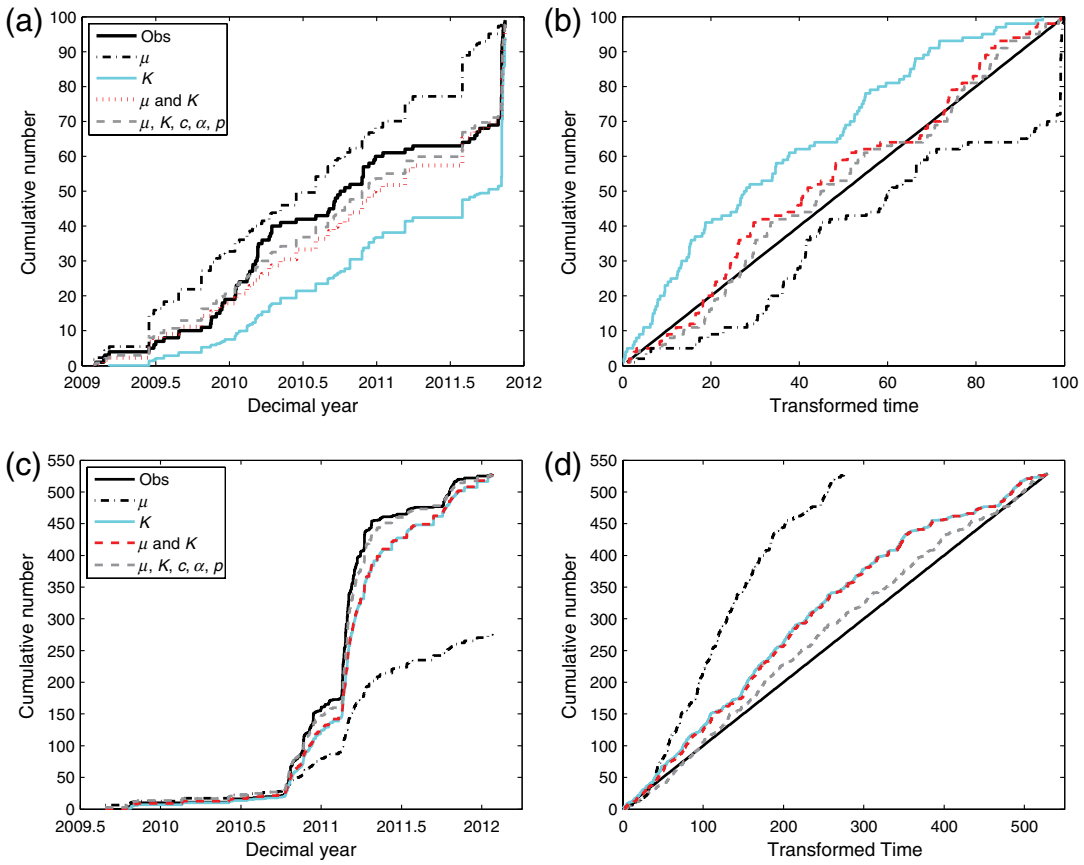

Figure 5. Comparison of how well changing different parameters from their pre-2009 values can explain post-2009 seismicity. (a) Cumulative number of events observed (solid black line) and predicted for models with varying different parameters (see legend for line styles) in Oklahoma. While the five-parameter model most closely matches the observations, its AIC shows it does not offer significant improvement over the $\mu + K$ model (Table 1). (b) Transformed time plots for the different models in Oklahoma. The solid black line indicates a perfect fitting model. (c) Cumulative number of events observed (solid black line) and predicted for models varying different parameters (see legend for line styles) in Arkansas. The five-parameter model most closely matches the data and offers significant improvement over the $\mu + K$ model. (d) Transformed time plots for the different models in Arkansas. The color version of this figure is available only in the electronic edition.

earthquake and the nearest well and compare the distribution of these distances for earthquakes that occur before injection (1982–2008), during injection (2009–2011), and after injection stopped (2011–2012). A significant change in this distribution occurs at an earthquake–well distance of 6 km for the earthquakes that happened during the years of active injection, which is not seen for the earthquakes that occurred in the time before or after (Fig. 6b). Roughly 85% of the earthquakes during 2009–2011 occurred within 6 km of an active injection well. Following shut-in of the last of the wells in late 2011, the seismicity rate has decreased, and the spatial distribution is similar to the earthquakes that occurred prior to 2009.

We can then compare the distribution of interevent distances (i.e., distance between consecutive earthquakes) for the earthquakes that occurred during 2009–2011 near an active well to those that occurred before, after, and far from active injection (Fig. 6c). Two-sample KS tests show that the interevent distance distribution of the earthquakes that occur within 6 km of an active well is significantly different than the other distributions. Around 85% of these earthquakes happen within 5 km of the previous earthquake, indicating that earthquakes that occur near active wells tend to occur closer to one another in space. The seismicity that happens after injection stops has a similar interevent distance distribution as the seismicity from before injection began. Thus, like a change in aftershock productivity, a significant change in interevent distance distribution may also be an indicator of induced seismicity.

Discussion

We have found that in both Oklahoma and Arkansas, a change in the background seismicity rate and triggering parameters used in the ETAS model must occur in order to account for the large increase in seismicity following 2009. Natural earthquake swarms have typically been modeled as time-dependent background seismicity rates in the ETAS model, caused, for example, by changes in pore fluid pressure due to fluid migration (e.g., Hainzl and Ogata, 2005; Lombardi *et al.*, 2010; Daniel *et al.*, 2011) or changes in

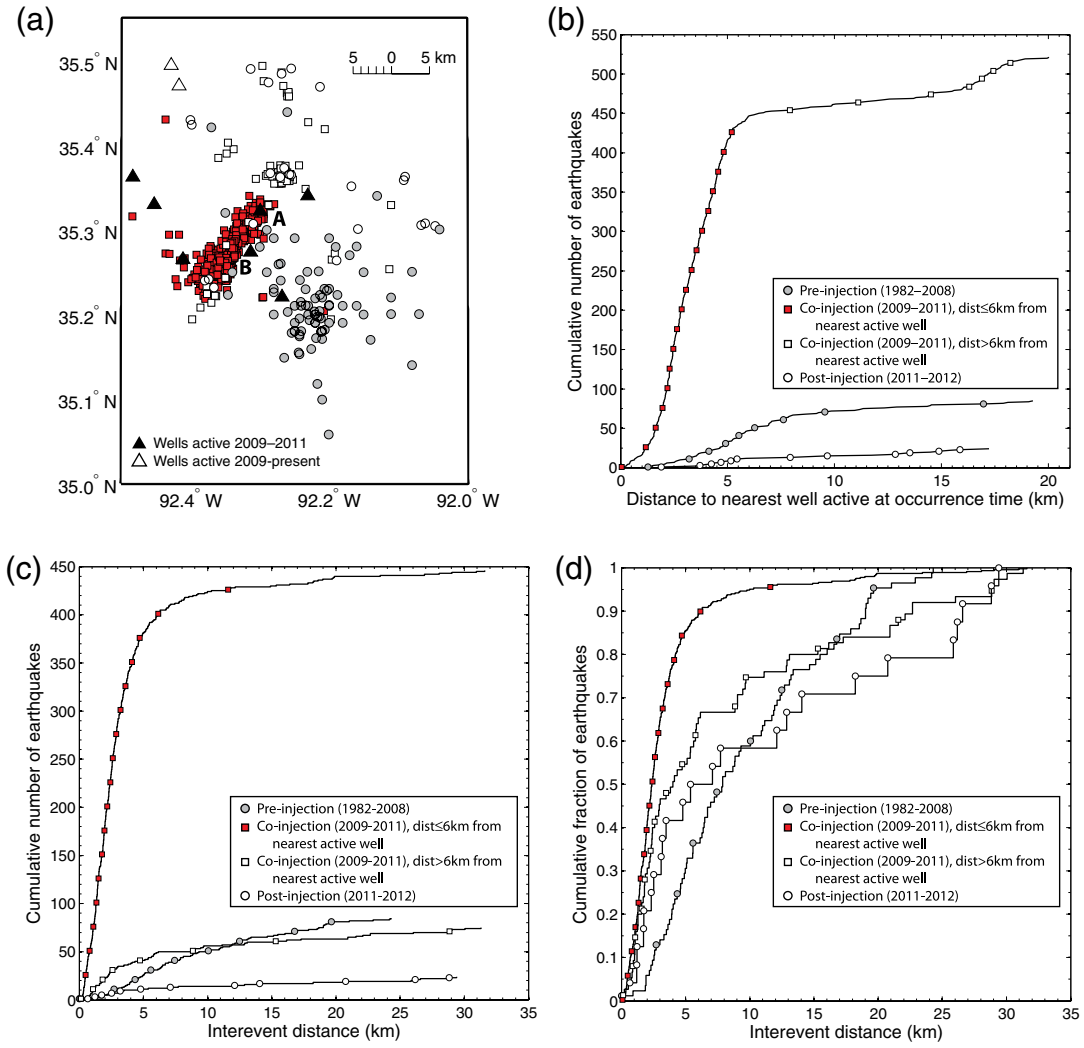

Figure 6. Comparing earthquake and injection well locations in Arkansas. (a) Arkansas 1982–2012 seismicity (circles and squares) and injection well locations (triangles) from the Arkansas Oil and Gas Commission. Injection wells have been active since 2009: black triangles were shut down in 2011, and white triangles are still operational. Seismicity is color coded by time and distance from an active well, with gray circles occurring before injection began in 2009, filled squares occurring in the years of active injection (2009–2011) within 6 km of an active well, white squares occurring in 2009–2011 farther than 6 km from an active well, and white circles occurring after the shutdown of injection in 2011 (see legend in [b]). (b) Cumulative distribution of the distance between each earthquake and the nearest active well. For the earthquakes that occur during injection (squares), a distinct change in the distribution happens at an earthquake–well distance of 6 km. (c) Cumulative distributions of interevent distances for each of the different populations of events shown. Earthquakes that occur close to an active well are a much larger population than the others. (d) Cumulative fraction plot comparing the shapes of the distributions shown in (c). Earthquakes that occur during injection within 6 km of an active well have a significantly different interevent distance distribution than the other earthquakes. The color version of this figure is available only in the electronic edition.

stress rate due to aseismic creep events (e.g., Llenos *et al.*, 2009; Llenos and McGuire, 2011; Okutani and Ide, 2011). To our knowledge, variations in aftershock productivity have not yet been associated with naturally occurring earthquake swarms, so changes in productivity may be a way to distinguish between natural and man-made earthquake rate changes.

The seismicity of Arkansas offers a way to explore how swarms triggered by natural processes may differ from swarms that have occurred following fluid injection. We can directly compare the 1980s Enola swarms that were likely triggered by natural fluid migration (Johnston, 1982; Chiu *et al.*, 1984; Rabak *et al.*, 2010) to the 2010–2011 Guy–Greenbrier swarms that were likely induced (Horton, 2012b). We find that the two sets of swarms cannot be explained by the same set of ETAS parameters (Table 3). Relative to the Enola swarms, the Guy–Greenbrier swarms show both higher background seismicity rate and aftershock productivity. The interevent distance distribution is also significantly different for the two swarms, which can essentially be seen in Figure 6d, because the Enola swarms comprise the majority of the events in the 1982–2008 time period and the Guy–Greenbrier swarms make up the majority of the events during the 2009–2011 time period. These results suggest that there is something fundamentally different about how the

Table 3
Arkansas Swarm Comparison

Data Set	μ (Events/Day)	K (Events/Day)	c (Days)	α	p	AIC
Enola, 1982–1987	0.007 ± 0.025	0.030 ± 0.034	0.002 ± 0.018	0.71 ± 0.19	0.95 ± 0.03	470.77
μ estimated*	0.010 ± 0.014	0.082	0.061	1.08	1.39	462.77
Guy–Greenbrier, 2010–2012	0.095 ± 0.618	0.082 ± 0.079	0.061 ± 0.284	1.08 ± 0.67	1.39 ± 0.3	−133.98
μ estimated†	0.204 ± 0.038	0.030	0.002	0.71	0.95	142.79
K estimated†	0.007	0.064 ± 0.006	0.002	0.71	0.95	−18.06
μ and K estimated †	0.013 ± 0.051	0.063 ± 0.006	0.002	0.71	0.95	−16.52
μ, K, α estimated †	0.015 ± 0.053	0.049 ± 0.016	0.002	1.22 ± 0.07	0.95	−24.03
μ, K, c estimated †	0.011 ± 0.049	0.067 ± 0.006	0.003 ± 0.008	0.71	0.95	−21.996
μ, K, p estimated †	0.026 ± 0.061	0.062 ± 0.006	0.002	0.71	1.01 ± 0.01	−33.81
μ, K, α, p estimated †	0.028 ± 0.098	0.051 ± 0.024	0.002	1.13 ± 0.09	1.01 ± 0.03	−38.50
μ, K, c, p estimated †	0.089 ± 0.055	0.096 ± 0.011	0.059 ± 0.007	0.71	1.39 ± 0.02	−130.67
μ, K, c, α estimated †	0.012 ± 0.048	0.052 ± 0.006	0.004 ± 0.004	1.24 ± 0.03	0.9457	−30.46

Uncertainties are 1 standard deviation.
*All other parameters held constant at 2010–2012 estimates to evaluate impact of parameter changes.
†All other parameters held constant at 1982–1987 estimates to evaluate impact of parameter changes.

earthquake swarms in the 1980s–1990s were triggered and how the earthquake swarms of 2010–2011 were triggered, which is consistent with the hypothesis that the Guy–Greenbrier swarms were induced.

One well-known physical mechanism for induced seismicity as well as natural fluid-driven swarms is pore pressure diffusion. Increases in pore fluid pressure act to reduce fault strength, bringing pre-existing fractures closer to failure according to the Mohr–Coulomb failure criterion. The amount of seismicity induced therefore depends on the ambient tectonic stress, as well as local geological and hydraulic conditions. The initiation of fluid injection in a region would lead to substantial increases in pore fluid pressure, which build up over time and diffuse outward from a well. Thus, induced seismicity can continue even after injection has ceased, as was the case at the Rocky Mountain Arsenal near Denver, Colorado, where three $m_b \sim 4.5$ earthquakes occurred the year after waste fluid injection ceased (Healy *et al.*, 1968; Hermann *et al.*, 1981; Hsieh and Bredehoeft, 1981). In Arkansas, the wastewater injection wells nearest to the Guy–Greenbrier swarms were shut down by the end of July 2011, but increased seismicity rates could be expected to continue while the pore pressure buildup from injection dissipates (Horton, 2012b). Indeed, another small swarm occurred in October 2011, although the seismicity rate since then seems to have decreased, generally following the Omori law. This is similar to what was observed at Basel, Switzerland, following the shut-down of an injection well used for an Enhanced Geothermal System project (Bachmann *et al.*, 2011). Thus, an increase in background seismicity rate may reflect events triggered by stress changes due to a pore pressure front, as seen in natural fluid-driven swarms (e.g., Hainzl and Ogata, 2005).

We can begin to compare the migration of seismicity in space and time from an injection well using simple models of pore pressure diffusion that assume a point source pressure perturbation occurs in an infinite homogeneous isotropic poroelastic saturated medium (Shapiro *et al.*, 1997). Figure 7 shows the seismicity migration with respect to two UIC wells in the study area, A (Permit 43266) and B (Permit 36380). Both wells inject into the Ozark aquifer and were shut down by 4 March 2011, shortly after the occurrence of an M_L 4.7, the largest event in the Guy–Greenbrier swarm (Horton, 2012b). Little is known about the hydrogeological conditions in the study area, but the migration of the earthquakes closest to these wells are consistent with hydraulic diffusivities ranging from 0.01 to 0.1 m^2/s. From previous studies of reservoir-induced seismicity, diffusivities in the crust tend to range from 0.5 to 50 m^2/s (e.g., Talwani and Acree, 1985). At the Rocky Mountain Arsenal, where induced earthquakes occurred in a similar depth range in fractured Precambrian crystalline basement as the Guy–Greenbrier events, the diffusivity was on the order of 1 m^2/s (Hsieh and Bredehoeft, 1981; Horton, 2012b). Lower diffusivities, which indicate a time lag between injection and triggered earthquakes, may be possible in situations where fluid is injected into reservoirs that are bound by faults that act as barriers to fluid flow, which may have been the case for the November 2011 Oklahoma earthquake sequence (Keranen *et al.*, 2013).

The physical meaning of an increase in aftershock productivity that may be associated with the commencement of fluid injection in a region is more elusive and beyond the scope of this paper. Changes in injection volume, rate, or pressure may cause changes in pore pressure that are sufficient to bring the entire system closer to a critical state of global failure. Then the increase in productivity may reflect earthquakes triggered by coseismic stress changes on faults that were closer to failure than they had been prior to the start of injection. Lei *et al.* (2008) suggest a similar mechanism governing an earthquake sequence induced by water injection at the Rongchang gas field in China, where the events are triggered by stress changes from both pore pressure diffusion and seismic slip. Recent laboratory experiments have

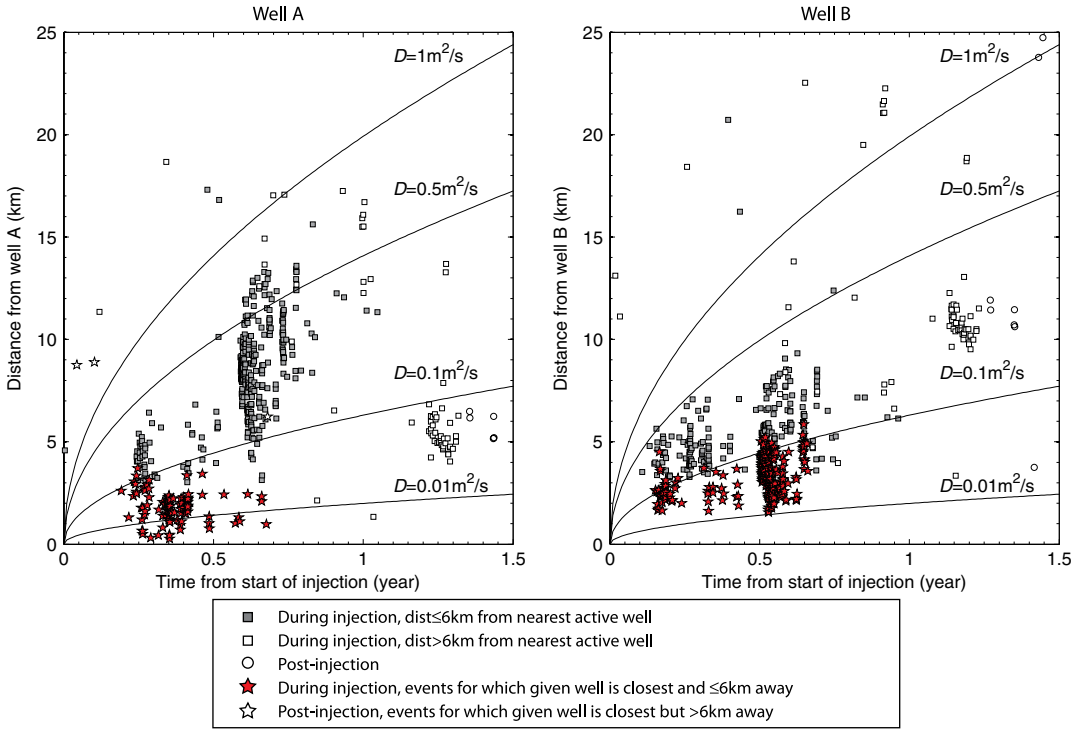

Figure 7. Pore pressure diffusion modeling for wells A and B, indicated in the map in Figure 6a. Black lines are theoretical pore pressure diffusion curves for different values of hydraulic diffusivity *D*. Seismicity from 2009 to 2012 are plotted by distance and time from start of injection at the given well. Gray squares are events that occur from 2009 to 2011 within 6 km of any active well, white squares occur from 2009 to 2011 farther than 6 km from any active well, white circles occur after injection stops in 2011. Stars indicate the events that occur during the years of injection (2009–2011) that are closest to either well A (left) or well B (right). Filled stars occur within 6 km of the well and white stars occur farther than 6 km from the well. Assuming these starred events are most likely to be triggered by the given well, their distributions are fit by relatively low diffusivities. The color version of this figure is available only in the electronic edition.

also shown that samples under excess pore pressure show lower shear strengths than samples deformed at the same stress conditions at normal pore pressure (Ougier-Simonin and Zhu, 2013). Therefore an increase in aftershock productivity may be expected in regions that have had significant pore pressure increases due to the start of fluid injection.

Conclusion

The results of this study may seem to suggest that absolute values of ETAS parameters themselves may be a way to distinguish between natural and man-made seismicity, particularly due to wastewater injection. For example, it has been shown that natural swarms in a number of different regions tend to have low values of α (e.g., Hainzl and Ogata, 2005; Lei *et al.*, 2008). Perhaps induced seismicity in general tends to have higher values of both μ and K. Testing this hypothesis however is complicated by the fact that ETAS parameters can vary widely from region to region, and μ and K are particularly shown to be location dependent (Ogata, 2004). Our results demonstrate that within one region, it is the time-dependent changes in these parameters that may be more indicative of whether or not seismicity has been triggered by a different process.

To summarize, the rate of small-to-moderate earthquakes in Oklahoma and Arkansas changed significantly in 2009, and these changes do not seem to be due to random fluctuations in natural seismicity rates. Our results suggest that significant changes in both the background rate of independent events and the triggering properties need to occur in order to explain the increases in seismicity, in contrast to what is typically observed when modeling natural earthquake swarms. Moreover, in Arkansas, earthquakes that occur near active wastewater injection wells tend to occur closer together than the earthquakes that occur before, after, or far from active injection. Therefore, a change in aftershock productivity or a change in the interevent distance distribution may be a way to distinguish natural earthquake rate changes from rate changes associated with wastewater injection. Improved seismic monitoring near injection wells will be necessary to resolve whether changes in productivity or changes in spatial distribution are the dominant factor in the increased earthquake rates seen at these locations.

Data and Resources

Seismicity data were obtained from the Advanced National Seismic System catalog at www.quake.geo.berkeley.edu/anss/catalog-search.html (last accessed November

2012) and the U.S. Geological Survey-National Earthquake Information Center Preliminary Determination of Epicenters catalog at earthquake.usgs.gov/earthquakes/eqarchives/epic/ (last accessed November 2011). Arkansas injection well data were obtained from the Arkansas Oil and Gas Commission at www.aogc.state.ar.us/JDesignerPro/JDPArkansas/default.htm (last accessed November 2012).

Acknowledgments

We thank Bill Ellsworth, Art McGarr, and David Lockner for insightful discussions and Jeanne Hardebeck and Justin Rubinstein for their helpful reviews. We also thank Sebastian Hainzl and Corinne Bachmann for their constructive comments.

References

Akaike, H. (1974). A new look at the statistical model identification, *IEEE Trans. Automat. Contr.* **AC-19,** 716–723.

Bachmann, C. E., S. Wiemer, J. Woessner, and S. Hainzl (2011). Statistical analysis of the induced Basel 2006 earthquake sequence: Introducing a probability-based monitoring approach for Enhanced Geothermal Systems, *Geophys. J. Int.* **186,** 793–807, doi: 10.1111/j.1365-246X.2011.05068.x.

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

Daley, D. J., and D. Vere-Jones (2002). *An Introduction to the Theory of Point Processes*, Vol. 1, Second Ed., Springer, New York.

Daniel, G., E. Prono, F. Renard, F. Thouvenot, S. Hainzl, D. Marsan, A. Helmstetter, P. Traversa, J. L. Got, L. Jenatton, and R. Guiguet (2011). Changes in effective stress during the 2003–2004 Ubaye seismic swarm, France, *J. Geophys. Res.* **116,** doi: 10.1029/2010JB007551.

Ellsworth, W. L., S. H. Hickman, A. L. Llenos, A. McGarr, A. J. Michael, and J. L. Rubinstein (2012). Are seismicity rate changes in the midcontinent natural or manmade? *Seismol. Res. Lett.* **83** (abstract), 403.

Hainzl, S., and Y. Ogata (2005). Detecting fluid signals in seismicity data through statistical earthquake modeling, *J. Geophys. Res.* **110,** doi: 10.1029/2004JB003247.

Hainzl, S., O. Zakharova, and D. Marsan (2013). Impact of aseismic transients on the estimation of aftershock productivity parameters, *Bull. Seismol. Soc. Am.* **103,** doi: 10.1785/0120120247.

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

Hermann, R. B., S.-K. Park, and C.-Y. Wang (1981). The Denver earthquakes of 1967–1968, *Bull. Seismol. Soc. Am.* **71,** 731–745.

Horton, S. (2012a). Deep fluid injection near the *M* 5.6 Oklahoma earthquake of November, 2011, *Seismol. Res. Lett.* **83** (abstract), 420.

Horton, S. (2012b). Disposal of hydrofracking waste fluid by injection into subsurface aquifers triggers earthquake swarm in central Arkansas with potential for damaging earthquake, *Seismol. Res. Lett.* **83,** 250–260.

Hsieh, P. A., and J. D. Bredehoeft (1981). A reservoir analysis of the Denver earthquakes: A case of induced seismicity, *J. Geophys. Res.* **86,** 903–920.

Johnston, A. C. (1982). Arkansas' earthquake laboratory, *Eos Trans. AGU* **63,** 1209–1210.

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

Lei, X., G. Yu, S. Ma, X. Wen, and Q. Wang (2008). Earthquakes induced by water injection at ~3 km depth within the Rongchang gas field, Chongqing, China, *J. Geophys. Res.* **113,** doi: 10.1029/2008JB005604.

Lilliefors, H. W. (1969). On the Kolmogorov-Smirnov test for the exponential distribution with mean unknown, *J. Am. Stat. Assoc.* **64,** 387–389.

Llenos, A. L., and J. J. McGuire (2011). Detecting aseismic strain transients from seismicity data, *J. Geophys. Res.* **116,** doi: 10.1029/2010JB007537.

Llenos, A. L., J. J. McGuire, and Y. Ogata (2009). Modeling seismic swarms triggered by aseismic transients, *Earth Planet. Sci. Lett.* **281,** 59–69.

Lombardi, A. M., M. Cocco, and W. Marzocchi (2010). On the increase of background seismicity rate during the 1997–1998 Umbria–Marche, central Italy, sequence: Apparent variation or fluid-driven triggering? *Bull. Seismol. Soc. Am.* **100,** doi: 10.1785/0120090077.

Ogata, Y. (1988). Statistical models for earthquake occurrences and residual analysis for point processes, *J. Am. Stat. Assoc.* **83,** 9–27.

Ogata, Y. (1992). Detection of precursory relative quiescence before great earthquakes through a statistical model, *J. Geophys. Res.* **97,** 19,845–19,871.

Ogata, Y. (2004). Space-time model for regional seismicity and detection of crustal stress changes, *J. Geophys. Res.* **109,** no. B03308, doi: 10.1029/2003JB002621.

Okutani, T., and S. Ide (2011). Statistic analysis of swarm activities around the Boso peninsula, Japan: Slow slip events beneath Tokyo Bay? *Earth Planets Space* **63,** 419–426.

Omori, F. (1894). On the aftershocks of earthquakes, *J. Coll. Sci. Imp. Univ. Tokyo* **7,** 111–200.

Ougier-Simonin, A., and W. Zhu (2013). Effects of pore fluid pressure on slip behaviors: An experimental study, *Geophys. Res. Lett.* **40,** doi: 10.1002/grl.50543.

Pujol, J., J. M. Chiu, A. C. Johnston, and B. H. Chin (1989). On the relocation of earthquake clusters. A case history: The Arkansas swarm, *Bull. Seismol. Soc. Am.* **79,** 1846–1862.

Rabak, I., C. Langston, P. Bodin, S. Horton, M. Withers, and C. Powell (2010). The Enola, Arkansas, intraplate swarm of 2001, *Seismol. Res. Lett.* **81,** 549–559.

Shapiro, S. A., E. Huenges, and G. Borm (1997). Estimating the crust permeability from fluid-injection-induced seismic emission at the KTB site, *Geophys. J. Int.* **131,** F15–F18.

Talwani, P., and S. Acree (1985), Pore pressure diffusion and the mechanism of reservoir-induced seismicity, *Pure Appl. Geophys.* **122,** 947–965.

Utsu, T. (1961). A statistical study on the occurrence of aftershocks, *Geophys. Mag.* **30,** 521–605.

Wald, A., and J. Wolfowitz (1940). On a test whether two samples are from the same population, *Ann. Math. Statist.* **11,** 147–162.

Wiemer, S., and M. Wyss (2000). Minimum magnitude of completeness in earthquake catalogs: Examples from Alaska, the western United States, and Japan, *Bull. Seismol. Soc. Am.* **90,** 859–869.

Wiemer, S., S. R. McNutt, and M. Wyss (1998). Temporal and three-dimensional spatial analyses of the frequency-magnitude distribution near Long Valley caldera, California, *Geophys. J. Int.* **134,** 409–421.

U.S. Geological Survey
MS 977
345 Middlefield Road
Menlo Park, California 94025
allenos@usgs.gov

Manuscript received 22 January 2013