

A Bayesian Change Point Model To Detect Changes In Event Occurrence Rates, With Application To Induced Seismicity

Abhineet Gupta

Graduate Student, Dept. of Civil and Environmental Engineering, Stanford University, Stanford, CA 94305, USA

Jack W. Baker

Associate Professor, Dept. of Civil and Environmental Engineering, Stanford University, Stanford, CA 94305, USA

ABSTRACT: A significant increase in earthquake occurrence rates has been observed in recent years in parts of Central and Eastern US. There is a possibility that this increased seismicity is anthropogenic and is referred to as induced seismicity. In this paper, a Bayesian change point model is implemented to evaluate whether temporal features of observed earthquakes support the hypothesis that a change in seismicity rates has occurred. This model is then used to estimate when the change is likely to have occurred. The magnitude of change is also quantified by estimating the distributions of seismicity rates before and after the change. These calculations are validated using a simulated data set with a known change point and event occurrence rates; and then applied to earthquake occurrence data for a site in Oklahoma.

1. INTRODUCTION

The level of seismicity in the Central and Eastern US (CEUS) has increased markedly since approximately 2009 (Ellsworth, 2013). For Oklahoma, the cumulative number of earthquakes with magnitude ≥ 3 since 1974 is shown in Fig. 1. The figure shows a marked increase in seismicity rate in Oklahoma after 2008. The magnitude 3 threshold was chosen since Coppersmith et al. (2012) described that there is catalog completeness in CEUS at this magnitude level. (All earthquake data from Oklahoma used in this paper has been obtained from catalogs developed and maintained by Oklahoma Geological Survey's Leonard Geophysical Observatory.)

There is a possibility that this increased seismicity is a result of underground wastewater injection (e.g., Ellsworth 2013; Keranen et al. 2013, 2014). Seismicity generated as a result of human activities is referred to as induced or triggered seismicity. Induced seismicity in a region can alter its seismic hazard and risk. One of the important components in calculating seismic hazard through Probabilistic Seismic Hazard Analysis (PSHA) is the activity

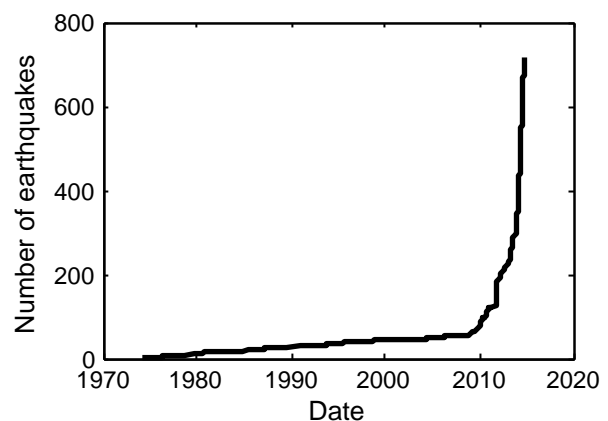


Figure 1: Cumulative number of earthquakes in Oklahoma with magnitude ≥ 3 from 1974 to Sept 2014

rate at a seismic source. For a given seismic source, we thus need to establish whether earthquake occurrence data indicates that a change in rates has occurred, or whether the earthquake activity is consistent with normal features of a process having a constant rate.

In this study, we describe a Bayesian change-point model that uses event occurrence data to indicate whether a change in event rates occurred. We assume that the event occurrences belong to a Poisson distribution. If a change in rates is detected, we also obtain the probability of change occurring at any given time, and the probability distributions of rates before and after the change. Along with the description of the model, an algorithmic implementation is provided. The model is validated through its application to a simulated data set with known properties.

The Bayesian model is then implemented on a region in Oklahoma to evaluate whether change in rates has occurred in this region. Finally, the distributions for seismicity rates before and after the change are calculated. This detection could be used to inform seismic hazard analysis and can serve as a decision support tool for operations that may be linked to induced seismicity.

2. BAYESIAN MODEL FOR CHANGE POINT DETECTION

Change point models are used to detect changes in occurrence rates of events. A Bayesian model for change point detection is implemented here to quantify changes in seismicity rates. The unknowns in our problem are the date of change, event rate before the change and rate after the change. Unlike parameter estimation techniques where a single value of a parameter is obtained, using a Bayesian model yields a probability distribution for the parameters. The probability distribution for rates before and after the change is helpful in seismic hazard calculation as it allows a more rigorous accounting of uncertainties.

For our model, it is assumed that occurrence of earthquakes is a Poisson process. Declustering an earthquake catalog (i.e. removing dependent aftershocks and foreshocks and only preserving independent mainshocks) accomplishes the requirements of this assumption as described by Gardner and Knopoff (1974) and van Stiphout et al. (2012). Furthermore, it is assumed that the seismicity rates before and after the change, as well as the date of change are mutually independent. Using Bayesian analysis to detect change point in a Poisson process

has been described by Raftery and Akman (1986).

2.1. Marginal posterior distributions of time of change and event rates

We define data in an observation period $[0, T]$ as a vector of inter-event times (i.e. the time between successive events) \mathbf{t} . The first event occurs at time 0 and the $n + 1^{th}$ event occurs at time T .

$$\mathbf{t} = \{t_1, t_2, \dots, t_n\} \quad s.t. \quad \sum_i t_i = T \quad (1)$$

The variables τ , λ_1 , and λ_2 define the date of change, event occurrence rate before the change, and occurrence rate after the change, respectively. Since the events belong to a Poisson process, the inter-event times are exponentially distributed.

$$f_{\lambda(s)}^T(t) = \lambda(s)e^{-\lambda(s)t} \quad (2)$$

such that

$$\lambda(s) = \begin{cases} \lambda_1, & 0 \leq s \leq \tau \\ \lambda_2, & \tau < s \leq T \end{cases} \quad (3)$$

A Gamma distribution with parameters k_j and θ_j is used to define the conjugate prior of λ_j .

$$\pi(\lambda_j) \propto \lambda_j^{k_j-1} e^{-\lambda_j/\theta_j} \quad (4)$$

The prior distribution for the time of change, τ is assumed to be uniformly distributed over the observation period. This implies that change is equally likely to occur at any time during the observation period.

$$\pi(\tau) = \frac{1}{T}, \quad 0 \leq \tau \leq T \quad (5)$$

The likelihood function for the unknown parameters can be written as

$$\begin{aligned} \mathcal{L}(\tau, \lambda_1, \lambda_2 | \mathbf{t}) &= \prod_{t=t_1}^{\tau} \lambda_1 e^{-\lambda_1 t} \prod_{t=t_{\tau+1}}^{t_n} \lambda_2 e^{-\lambda_2 t} \\ &= \lambda_1^{N(\tau)} e^{-\lambda_1 \tau} \lambda_2^{N(T)-N(\tau)} e^{-\lambda_2 (T-\tau)} \end{aligned} \quad (6)$$

where $N(t)$ represents the number of events between $[0, t]$. Using the fact that all parameters are

mutually independent, the posterior density is calculated as

$$\begin{aligned} \pi(\tau, \lambda_1, \lambda_2 | \mathbf{t}) \\ \propto \mathcal{L}(\tau, \lambda_1, \lambda_2 | \mathbf{t}) \pi(\lambda_1) \pi(\lambda_2) \pi(\tau) \end{aligned} \quad (7)$$

Then the marginal distributions for each of τ , λ_1 and λ_2 can be obtained by integrating the posterior density over the remaining two variables. The marginal posterior distribution of τ is calculated as

$$\begin{aligned} \pi(\tau | \mathbf{t}) &\propto \int_{-\infty}^{\infty} \int_0^{\infty} \pi(\lambda_1, \lambda_2, \tau | \mathbf{t}) d\lambda_1 d\lambda_2 \\ &= \pi(\tau) \\ &\quad \cdot \int_0^{\infty} \lambda_1^{N(\tau)+k_1-1} e^{-\lambda_1(\tau+\frac{1}{\theta_1})} d\lambda_1 \\ &\quad \cdot \int_0^{\infty} \lambda_2^{N(T)-N(\tau)+k_2-1} \\ &\quad \quad \cdot e^{-\lambda_2(T-\tau+\frac{1}{\theta_2})} d\lambda_2 \\ &= \frac{1}{T} \cdot \frac{\Gamma(r_1(\tau))\Gamma(r_2(\tau))}{S_1(\tau)r_1(\tau)S_2(\tau)r_2(\tau)} \end{aligned} \quad (8)$$

where

$$\begin{aligned} r_1(\tau) &= N(\tau) + k_1 \\ S_1(\tau) &= \tau + \frac{1}{\theta_1} \\ r_2(\tau) &= N(T) - N(\tau) + k_2 \\ S_2(\tau) &= T - \tau + \frac{1}{\theta_2} \end{aligned} \quad (9)$$

The marginal posterior distribution of λ_1 is calculated as shown below. A closed form solution for this double integration does not exist. Hence, to evaluate the probability distribution, the time range is discretized on a per day basis, and summed up to approximate the marginal distribution.

$$\begin{aligned} \pi(\lambda_1 | \mathbf{t}) &\propto \int_0^T \int_0^{\infty} \pi(\lambda_1, \lambda_2, \tau | \mathbf{t}) d\lambda_2 d\tau \\ &= \int_0^T \int_0^{\infty} \lambda_2^{r_2(\tau)-1} e^{-\lambda_2 S_2(\tau)} d\lambda_2 \\ &\quad \cdot \pi(\tau) \lambda_1^{r_1(\tau)-1} e^{-\lambda_1 S_1(\tau)} d\tau \\ &\approx \sum_{\tau=0}^T \left[\frac{1}{T} \cdot \lambda_1^{r_1(\tau)-1} e^{-\lambda_1 S_1(\tau)} \right. \\ &\quad \left. \cdot \Gamma(r_2(\tau)) S_2(\tau)^{r_2(\tau)} \right] \end{aligned} \quad (10)$$

Similarly, the marginal posterior distribution of λ_2 can be calculated as

$$\begin{aligned} \pi(\lambda_2 | \mathbf{t}) &\propto \sum_{\tau=0}^T \left[\frac{1}{T} \cdot \lambda_2^{r_2(\tau)-1} e^{-\lambda_2 S_2(\tau)} \right. \\ &\quad \left. \cdot \Gamma(r_1(\tau)) S_1(\tau)^{r_1(\tau)} \right] \end{aligned} \quad (11)$$

Another quantity of interest is the ratio of pre-change event rate to post-change event rate (or vice-versa), defined as $\beta = \lambda_1/\lambda_2$. Lindley (1965) describes the following function of β conditional on τ to follow the *F-distribution* with d.o.f. p_1 and p_2 , where $p_k = 2r_k(\tau)$.

$$\frac{S_1(\tau)r_2(\tau)}{S_2(\tau)r_1(\tau)} \beta \sim F_{p_1, p_2} \quad (12)$$

The above equation can be used to compute the probability distribution of β conditional on τ .

$$\begin{aligned} p(\beta | \tau, \mathbf{t}) &= \frac{1}{B(r_1(\tau), r_2(\tau))} \left(\frac{r_1(\tau)}{r_2(\tau)} \right)^{r_1(\tau)} \\ &\quad \cdot \left(\frac{S_2(\tau) + S_1(\tau)\beta}{S_2(\tau)} \right)^{-(r_1(\tau)+r_2(\tau))} \\ &\quad \cdot \left(\frac{S_1(\tau)r_2(\tau)}{S_2(\tau)r_1(\tau)} \beta \right)^{r_1(\tau)-1} \\ &\quad \cdot \frac{S_1(\tau)r_2(\tau)}{S_2(\tau)r_1(\tau)} \end{aligned} \quad (13)$$

The marginal distribution of β is then calculated as

$$\begin{aligned} p(\beta | \mathbf{t}) &= \int_0^T p(\beta | \tau, \mathbf{t}) \pi(\tau | \mathbf{t}) d\tau \\ &\propto \int_0^T \pi(\tau) \beta^{r_1(\tau)-1} \\ &\quad \cdot (S_2(\tau) + S_1(\tau)\beta)^{-(r_1(\tau)+r_2(\tau))} d\tau \\ &\approx \sum_{\tau=0}^T \frac{1}{T} \cdot \beta^{r_1(\tau)-1} \\ &\quad \cdot (S_2(\tau) + S_1(\tau)\beta)^{-(r_1(\tau)+r_2(\tau))} \end{aligned} \quad (14)$$

It is noted after the evaluation of eq. 14 that there is some difference in the equation terms compared to those obtained by Raftery and Akman (1986). We are able to verify and validate our equation with the same data as Raftery and Akman (1986), and believe this formulation to be correct.

2.2. Bayes factor

The change-point model described in the previous section does not indicate whether the data supports the presence of a change in the observed time range. When the model is applied to observed data, it assumes that there is a change and calculates the probability of change on any given date. To indicate whether data favors a change-point model, the Bayes factor is used.

The Bayes factor is a Bayesian alternative to hypothesis testing. In this case, it is defined as the ratio of the likelihood function for a constant rate model (H_0) to that for a change model (H_1). Hence, it is used to compare which model better describes the data.

$$B_{01}(\mathbf{t}) = \frac{\mathcal{L}(H_0 | \mathbf{t})}{\mathcal{L}(H_1 | \mathbf{t})} \quad (15)$$

The constant rate model is characterized by a single unknown parameter - the rate of occurrence, λ_0 . A gamma distribution with parameters k_0 and θ_0 , similar to eq. 4, is used as its prior distribution. Then

$$\begin{aligned} \mathcal{L}(H_0 | \mathbf{t}) &= \int_0^\infty \mathcal{L}(\lambda_0 | \mathbf{t}) \pi(\lambda_0) d\lambda_0 \\ &= \frac{\Gamma(N(T) + k_0)}{\Gamma(k_0)} \left(\frac{1}{\theta_0}\right)^{k_0} \\ &\quad \cdot \left(\frac{1}{\theta_0} + T\right)^{-(N(T)+k_0)} \end{aligned} \quad (16)$$

and

$$\begin{aligned} \mathcal{L}(H_1 | \mathbf{t}) &= \int_0^T \int_0^\infty \int_0^\infty \mathcal{L}(\tau, \lambda_1, \lambda_2 | \mathbf{t}) \\ &\quad \cdot \pi(\lambda_1) \pi(\lambda_2) \pi(\tau) d\lambda_1 d\lambda_2 d\tau \\ &\approx \left(\frac{1}{\theta_1}\right)^{k_1} \left(\frac{1}{\theta_2}\right)^{k_2} \frac{1}{\Gamma(k_1)\Gamma(k_2)} \\ &\quad \cdot \sum_{\tau=0}^T \left[\frac{\pi(\tau)\Gamma(r_1(\tau))\Gamma(r_2(\tau))}{S_1(\tau)^{r_1(\tau)}S_2(\tau)^{r_2(\tau)}} \right] \end{aligned} \quad (17)$$

If the value of parameters for gamma conjugate priors are $k_j = 0.5$ and $\theta_j \rightarrow \infty$ for $j = 0, 1, 2$, then it is shown by Raftery and Akman (1986) that the

equation for Bayes factor can be simplified to

$$\begin{aligned} B_{01}(\mathbf{t}) &= 4\sqrt{\pi} T^{-n} \Gamma(n + 1/2) \\ &\quad \cdot \left[\sum_{\tau=0}^T \Gamma(r_1(\tau))\Gamma(r_2(\tau)) \right. \\ &\quad \left. \cdot S_1(\tau)^{-r_1(\tau)} S_2(\tau)^{-r_2(\tau)} \right]^{-1} \end{aligned} \quad (18)$$

Smaller values of Bayes factor (less than 1) imply that the change model is more strongly supported by the data. Bayes factor values of less than 0.01 (or greater than 100) are typically used to indicate decisive preference for one model or the other (Kass and Raftery, 1995). In this study, a Bayes factor of smaller than 1×10^{-3} is used to indicate that data favors the change model compared to the constant rate model (i.e if the Bayes factor for an observed period is calculated to be less than 1×10^{-3} , it indicates that a change has occurred within this observed period).

3. ALGORITHMIC IMPLEMENTATION OF THE MODEL

Since all the unknown variables in the model described above (τ , λ_1 , and λ_2) are continuous and closed form solutions are not possible for marginal distributions of λ_1 , and λ_2 , some approximations are required to implement the algorithm for the change-point model. Additionally, overflow conditions are encountered in the algorithm, for instance when computing $\Gamma(x)$ function for big values. The approximations and inputs required for the implementation of our algorithm are described below. Later, the verification and validation processes for the algorithm are described.

3.1. Approximations and prior parameter inputs for the algorithm

Under the first approximation, the continuous variables are discretized. Since it is not computationally possible to calculate the marginal distribution of τ over a continuous range, the algorithm is implemented on a per-day basis. Hence, probability of change happening at time τ is calculated for every day of the observation period. The proportionality in eq. 8 is converted to a probability by dividing the values for each day by the sum of values for all

days in the observation period, such that the sum of probabilities equals one.

To compute the marginal distributions for λ_1 , λ_2 , and β , the integral over τ is discretized to a summation by again dividing the observation period on a per-day basis. Then the marginal distributions are approximated as shown in the last step of eqs. 10, 11 and 14. The results are calculated at discrete values of λ_1 , λ_2 , and β over a certain range $[X_{low}, X_{high}]$, and the proportionality is converted to probability again by dividing each result by the sum of results over the complete range. It is ensured that the selected range $[X_{low}, X_{high}]$ for each of the variables is wide enough such that probabilities at the extreme points are essentially zero.

Secondly, to address the overflow problem, the results are calculated in the *log* domain and then converted back to the original domain. One of the reasons for the overflow problem is computation of $\Gamma(x)$ which grows very rapidly with increasing values of x . In the *log* domain, algorithms are available to calculate $\log[\Gamma(x)]$ directly without first needing to calculate $\Gamma(x)$. Another overflow problem is encountered when the results obtained in *log* domain at each discrete value are converted back to original space and summed together. Their sum could be very large and outside the floating point range of the computer. Hence, to calculate probability at each discrete value, the *log* results are scaled such that the sum is within the range of computation.

As an example of the implementation of approximations and modifications described above, the algorithm for calculating the probability of change happening on any given day is written as:

(a) Discretize the observation period on a per-day basis, $\boldsymbol{\tau} = \{\tau_i\}_{i=1}^T$.

(b) At each τ_i , calculate

$$prob_i = -\log(T) + \log[\Gamma(r_1(\tau_i))] \\ + \log[\Gamma(r_2(\tau_i))] - r_1(\tau_i) \log[S_1(\tau_i)] \\ - r_2(\tau_i) \log[S_2(\tau_i)].$$

(c) Find a *scale* such that

$$\sum_i e^{prob_i - scale} \leq realMax$$

while preserving as many smallest values as possible. Here *realMax* is the largest finite floating-point number in IEEE double precision. Update $prob_i = prob_i - scale$.

(d) Normalize to obtain probability,

$$prob_i = \frac{e^{prob_i}}{\sum_i e^{prob_i}}.$$

The final step in the implementation of the algorithm is selecting the values for the prior parameters (k_j and θ_j) for the gamma distribution. The values are selected as the same ones that are used in the development of the formulation for Bayes factor in eq. 18 (i.e. $k_j = 0.5$ and $\theta_j \rightarrow \infty$ for $j = 0, 1, 2$). Brief analysis was performed to assess the sensitivity of the prior parameter values on results, and it was found that marginal distributions of τ , λ_1 , λ_2 , and β did not vary significantly with different prior parameters.

3.2. Verification and validation

The algorithm developed above for the change-point model is verified by comparing the results with those presented in Raftery and Akman (1986). (The data used for comparison is the coal-mining disasters data described by Jarrett (1979) and extended per Raftery and Akman (1986)). The results are in very good agreement. The algorithms are further validated through application on simulated data.

To obtain the simulated data, inter-event times are randomly generated from an exponential distribution with different rates and a known change point. The change-point model is then implemented on this data and the results for τ , λ_1 , and λ_2 are compared with the inputs. Here, we describe one simulation case for validation.

For this validation case, 100 inter-event duration are generated at a rate of 0.005 events/day and another 50 are generated at a rate of 0.015 events/day. A plot of cumulative number of events is shown in Fig. 2. The first event is assumed to occur on 2000-01-01. Since the event rate reduces starting from the 101st event, a change occurs between the 100th and the 101st event. Without any additional

information, we cannot determine a specific date of change. Hence, it is assumed that change occurred the day before the 101st event which is known from the data as 2052-06-01.

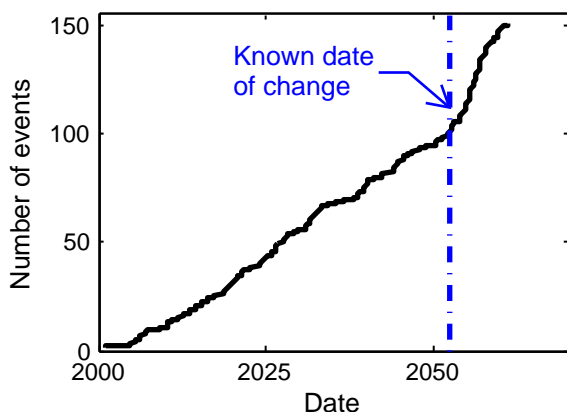


Figure 2: Cumulative number of events for simulated data with known change on 2052-06-02

The change-point model yields a Bayes factor of 6.3×10^{-7} for this data, implying that data strongly supports the change-point model. The probability of change happening on a given date is plotted on Fig. 3. A 95% credible interval for dates of change is determined between 2050-03-07 and 2054-05-28. Hence, the actual date of change is within the 95% credible interval. Additionally, the date of change with highest associated probability is found to be 2052-06-01. This date exactly matches our input date of change. Thus, the Bayesian change point model correctly estimates the date of change on simulated data.

The rates of event occurrence are also estimated using the model. The results of probability distributions of rates before and after the change are shown in Fig. 4. The modes of posterior distributions for rates are 0.0054 and 0.0157 respectively, which are in good agreement with the input rates.

4. CHANGE POINT MODEL IMPLEMENTATION IN OKLAHOMA

The change-point model described in previous sections, is used to quantify the change in seismicity rates in a local region of Oklahoma. The local region that is used in this paper is a 25 km radius

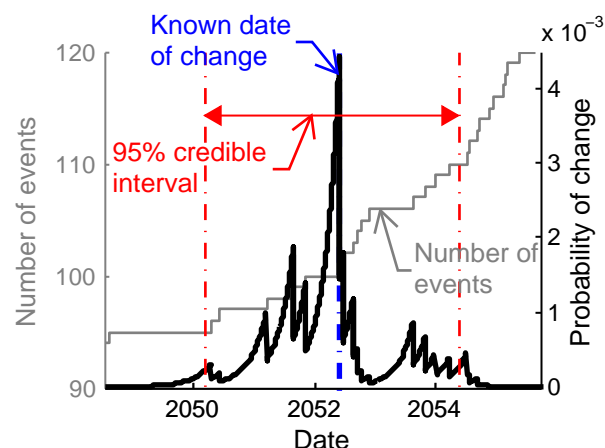


Figure 3: Probability of change on simulated data represented by the thick solid line, with its 95% credible interval between 2050-03-07 and 2054-05-28

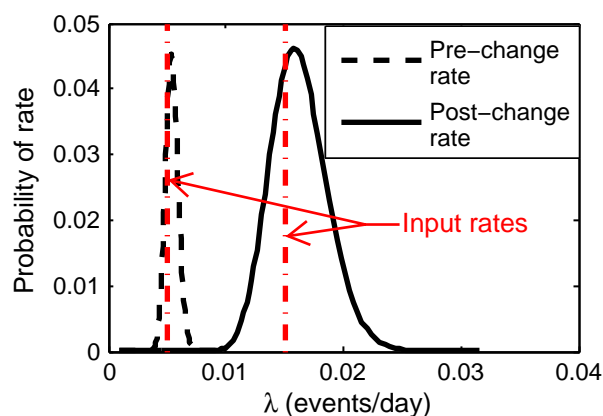


Figure 4: Probability of pre-change and post-change rates

area centered around the location of "Well 1" described in Keranen et al. (2013). The center of this region is located 10 km north-west of Prague, OK at (35.56°N, 96.75°W).

The cumulative number of events within the chosen region from the declustered catalog are shown with the dashed line in Fig. 5. No events are observed between 1974 and 2009 in this region. A Bayes factor of 9.6×10^{-28} is calculated for this data, implying that the data almost certainly supports the change-point model over a constant rate model. The probability of change occurring on any given day is then calculated and is shown in Fig. 5.

The highest probability for change is observed on 2009-06-13. A 95% credible interval is calculated to be between 2008-12-11 and 2010-02-24. This implies that there is a high chance that change in seismicity rates in this local region occurs between late 2008 and early 2010.

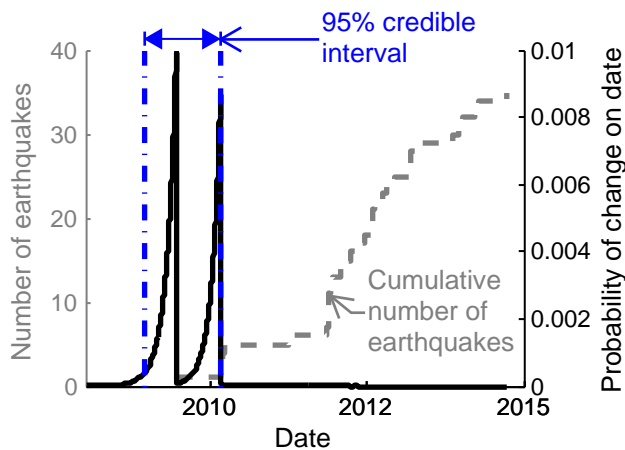


Figure 5: Probability of change on any given day for the local region, represented by the solid line

The probabilities of seismicity rates before and after the change are also calculated and are shown in Fig. 6. Since the pre-change rate is governed by the period of no seismicity between 1974 and 2009, its distribution is left-tailed. The modes for pre-change and post-change rates are 6.6×10^{-5} and 1.8×10^{-2} , respectively.

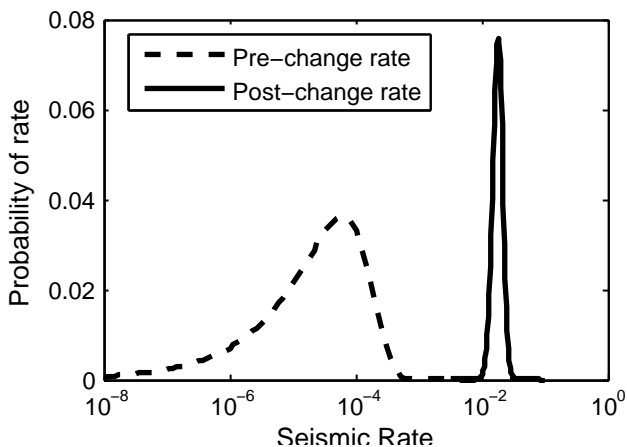


Figure 6: Probability of activity rates for earthquakes in the local region

The ratio of pre-change to post-change rate, β is shown in Fig. 7. The highest probability is calculated at a ratio of 3.4×10^{-3} . This implies that the post-change rate is about 300 times greater than the pre-change rate.

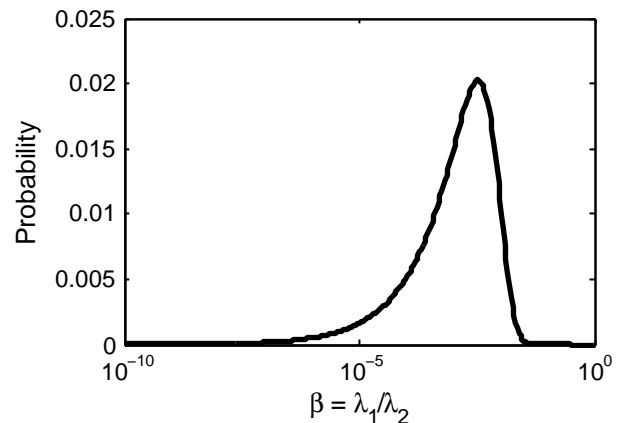


Figure 7: Probability of ratio of pre-change to post-change rates in the local region

5. CONCLUSIONS

An algorithm for change-point detection in event rates using Bayesian statistics was developed in this paper. The Bayes factor was defined in the model to determine whether event occurrence data supported that a change had occurred in event rates. If the data supported a change model, the probability of change happening on a given day could be computed. The model could also be used to estimate the probability distributions of event rates before and after the change. The model was validated through its application on a simulated data set with known properties.

After validating the model on simulated data, it was implemented on a region in Oklahoma. The model detected that a change in seismicity rates occurred sometime between late 2008 and mid 2010. This period agreed well with the time of change expected through a visual inspection of the data. The post-change rate was estimated to be approximately 300 times the pre-change rate. This high increase in activity rates can substantially increase the seismic hazard. Assuming that this local region

is a seismic source and that the magnitude distribution is held constant from pre-change to post-change, this would imply an increase of about 300 times in rate of exceeding any ground motion at this site from this source. However, more research needs to be carried out before calculating hazard from this increased seismicity. The magnitude distribution of induced earthquakes could be different than tectonic earthquakes due to difference in b -values (in a Gutenberg-Richter relation) or due to an upper bound on earthquake magnitudes. Additionally, there could be a difference in ground motions from induced earthquakes compared to natural earthquakes (Hough, 2014).

The change point model does not make any association with the causes of rate change. Hence, after determining that a change has occurred, it should be linked with a physical phenomenon that might have caused this change. In the case of induced seismicity in Oklahoma, some of these physical phenomena could be change in number of injection wells, cumulative injection volume or basement pore pressure. The information about dates of change provided by the change-point model can be used to identify the causes of induced seismicity. This identification can assist in decision-making for operations potentially linked with induced seismicity and can thus be used as a tool for risk mitigation.

Although the change-point model described here was applied for the case of induced seismicity, this is a versatile model with other potential applications. Some of the other applications of this model could be to detect change in storm occurrence rates to inform decisions about climate change, or to detect change in population migration rates to inform decisions about developing urban infrastructure. Thus the change-point model described in this paper can serve as a decision-support tool for a variety of applications involving occurrences of potentially non-stationary events.

6. REFERENCES

Coppersmith, K. J., Salomone, L. A., Fuller, C. W., Glaser, L. L., Hanson, K. L., Hartleb, R. D., Lettis, W. R., Lindvall, S. C., McDuffie, S. M., McGuire, R. K., and others (2012). "Central and Eastern United States (CEUS) seismic source characterization (SSC)

for nuclear facilities project." *Report no.*, Electric Power Research Institute (EPRI).

- Ellsworth, W. L. (2013). "Injection-induced earthquakes." *Science*, 341(6142), 1225-942.
- Gardner, J. K. and Knopoff, L. (1974). "Is the sequence of earthquakes in southern California, with aftershocks removed, Poissonian?." *Bulletin of the Seismological Society of America*, 64(5), 1363-1367.
- Hough, S. E. (2014). "Shaking from injection-induced earthquakes in the Central and Eastern United States." *Bulletin of the Seismological Society of America*, 104(5), 2619-2626.
- Jarrett, R. G. (1979). "A note on the intervals between coal-mining disasters." *Biometrika*, 66(1), 191-193.
- Kass, R. E. and Raftery, A. E. (1995). "Bayes factors." *Journal of the American Statistical Association*, 90(430), 773-795.
- Keranen, K. M., Savage, H. M., Abers, G. A., and 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.
- Keranen, K. M., Weingarten, M., Abers, G. A., Bekins, B. A., and Ge, S. (2014). "Sharp increase in central Oklahoma seismicity since 2008 induced by massive wastewater injection." *Science*, 345(6195), 448-451.
- Lindley, D. V. (1965). *Introduction to probability and statistics from Bayesian viewpoint. part 2 inference*, Vol. 2. Cambridge University Press Archive.
- Raftery, A. and Akman, V. (1986). "Bayesian analysis of a Poisson process with a change-point." *Biometrika*, 73(1), 85-89.
- van Stiphout, T., Zhuang, J., and Marsan, D. (2012). "Seismicity declustering." *Community Online Resource for Statistical Seismicity Analysis*, 10.