Monte Carlo simulations for analysis and prediction of non-stationary magnitude-frequency distributions in probabilistic seismic hazard analysis

Mauricio Reyes Canales and Mirko van der Baan
Dept. of Physics, CCIS, University of Alberta, T6G 2E1, Canada

Summary

This abstract describes a first principles methodology to evaluate statistically the hazard related to non-stationary seismic sources like induced seismicity. For this approach, we use time-dependent Gutenberg-Richter parameters, which leads to a time-varying rate of earthquakes. This is achieved by deriving analytic expressions for occurrence rates which are verified using Monte Carlo simulations. We show as an example a recent case of induced seismicity, the Horn River Basin, Northeast British Columbia. In this example, the statistics from the Monte Carlo simulations agree with the analytical quantities. The results show that induced seismicity can change seismic hazard rates but that this greatly depends on both the duration and intensity of the non-stationary sequence as well as the chosen evaluation period. Further studies will include extensions to handle spatial source distributions as well as ground motion evaluation in order to generate a complete methodology for non-stationary seismic hazard analysis.

Introduction

Several studies (Atkinson et al., 2016; Ellsworth, 2013) have shown increased seismicity in geologically stable basins in North America, thought to be associated with hydraulic fracturing treatments and/or salt water disposal wells. Induced seismicity may change seismic hazard. Yet one of the main challenges in applying traditional probabilistic seismic hazard analysis (PSHA) is the assumption of temporal stationarity, which is unlikely true for induced seismicity (Atkinson et al., 2015). The Monte Carlo simulation method for PSHA is flexible enough to deal more easily with non-stationary seismicity. This method consists of two main steps: (1) generation of synthetic earthquake catalogs and (2) generation of ground motion catalogs by using ground motion prediction equations (GMPEs). In this study we develop a methodology based on the Monte Carlo simulation method for the generation of non-stationary earthquake catalogs using time-dependent Gutenberg-Richter parameters. We also derive analytical expressions for various occurrence likelihoods such as annual rates of earthquakes. We show that the Poisson model for occurrence statistics remains valid. This is important since it is a common assumption for traditional PSHA predictions. To exemplify the applicability of the developed methodology and evaluate the implications of non-stationary seismicity in the hazard analysis, we study a recent case of induced seismicity: The Horn River Basin, Northeast British Columbia.

Theory and Method

Monte Carlo simulation method: we assume that the magnitude-frequency distribution of the synthetic events is described by the Gutenberg-Richter (GR) distribution (Gutenberg and Richter, 1944). For non-stationary seismicity, the intercept *a* and slope *b* in the GR distribution become time-dependent:

$\log(N) = a(t) - b(t)m$, (1)

where *N* is the number of earthquakes with a magnitude greater than *m*. The *b(t)* -value indicates the ratio of small and large magnitude events and the *a(t)*-value is related to the number N_0 of earthquakes with a non-negative magnitude per time unit. The latter is given by $N_0 = 10^{a(t)}$. Eq. 1 can be used to compute the discrete cumulative probability density function (CDF) as (Baker, 2008):

$$F_M(m,t) = \frac{1-10^{-b(t)(m-M_{min})}}{1-10^{-b(t)(M_{max}-M_{min})}}, \quad (2)$$

where $F_M(m,t)$ denotes the Cumulative distribution function for *m* in time *t*, M_{max} is the maximum magnitude and M_{min} is the minimum magnitude considered, both kept constant. The discrete probability for a magnitude *m* to occur within the range [m_j, m_{j+1}], a magnitude bin, is given by (Baker, 2008):

$$P(m_j < m < m_{j+1}, t) = F_M(m_{j+1}, t) - F_M(m_j, t), \quad (3)$$

with m_j and m_{j+1} respectively the lower and upper magnitude and *j* is an integer index to create magnitude bins. The spacing between boundaries m_j and m_{j+1} refers to the magnitude bin size. To calculate the rate of earthquakes λ ($m_j < m < m_{j+1}, t$) per time unit for a magnitude bin, we have:

$$\lambda\,(m_j < m < m_{j+1}; t) = P(m_j < m < m_{j+1}; t) * N(M_{min} < m < M_{max}\,; t), \quad (4)$$

where $P(m_j < m < m_{j+1}, t)$ is the probability of occurrence for a magnitude bin in time *t*, eq. 3, and $N(M_{min} < m < M_{max}, t)$ is the expected number of earthquakes per time unit in the range *m* = [M_{min}, M_{max}], which is derived by the GR relation (Eq.1):

$$N(M_{min} < m < M_{max}, t) = 10^{(a(t)-b(t)M_{min})} - 10^{(a(t)-b(t)M_{max})}. \quad (5)$$

In order to generate a synthetic event catalog, we apply Monte Carlo sampling using a uniform distribution, where we iterate over both time and all magnitude bins. For each time step, we draw a uniformly distributed random number *rand* between [0, 1]. If *rand* is lower than the rate of earthquakes λ ($m_j < m < m_{j+1}, t$) for the considered magnitude bin *m* = [m_j, m_{j+1}], then an event occurs, that is:

If *rand* < λ ($m_j < m < m_{j+1}, t$), occurrence= 1; If *rand*> λ ($m_j < m < m_{j+1}, t$), occurrence=0. (6)

This assumes that the occurrence rate is always less than 1. If the rate of earthquakes per time unit for a magnitude bin *m* = [m_j,m_{j+1}] exceeds 1, then we reduce the considered time units or we reduce the size of the magnitude bins. Every random generation represents a specific time sample, say one year, depending on the type of rate used. This process is repeated N_t times, generating a single realization for a specific magnitude bin (where N_t is the number of time samples). One realization is a possible scenario of the future seismicity for a given magnitude bin. If we generate N_r different realizations for a magnitude bin, we will obtain a matrix with dimensions $N_r \times N_t$. By creating N_r realizations for each magnitude bin, we will generate multiple matrices $N_r \times N_t$. These $N_r \times N_t$ matrices per magnitude bin are combined into a larger array, containing all magnitude-occurrence information. We can extract statistical results from this simulated array. For instance, the rate of earthquakes for a magnitude bin is obtained by counting the number of events with same magnitude *m* in time *t*, and dividing by the number of realizations.

Poisson distributions and derived statistical quantities: The Poisson distribution has been traditionally used to describe the number of events within a certain time interval for stationary earthquake rates (Cornell, 1968). It is easily modified to represent non-stationary rates as well, as long as the earthquake catalogs are generated using uniformly distributed numbers *rand* in eq.6. The non-stationary Poisson model has a rate of occurrence that varies with time; in this case, we use the mean $m_\lambda(t_a; t_b)$ of the time-dependent rate instead of a constant rate of occurrence (Sigman, 2013):

$$P[N = n; t_a, t_b] = \frac{m_\lambda{}^n(t_a, t_b)(t_b - t_a)^n e^{-m_\lambda(t_a, t_b)(t_b - t_a)}}{n!}, \quad (7)$$

where $m_\lambda(t_a; t_b)$ is the mean of the time-varying rate of occurrence $\lambda\,(t)$ in the time interval $t = [t_a, t_b]$, defined as (Sigman, 2013):

$$m_\lambda(t_a; t_b) = \frac{\int_{t_a}^{t_b} \lambda(s)\,ds}{t_b - t_a}. \quad (8)$$

$\lambda(t)$ could be the rate of earthquakes $\lambda\,(M_{min} < m < M_{max}; t)$ per time unit for the range [M_{min},M_{max}]:

$$\lambda\,(M_{min} < m < M_{max}; t) = P(M_{min} < m < M_{max}; t) * N(M_{min} < m < M_{max}\,; t), \quad (9)$$

where $N(M_{min} < m < M_{max}\,; t)$ is given by eq. 5 and $P(M_{min} < m < M_{max}; t)$ is a modification of eq.3, where we use the entire range *m* = [M_{min} , M_{max}] instead of a magnitude bin range. Replacing $\lambda(t)$ by $\lambda\,(M_{min} < m < M_{max}; t)$ in eq. 8, $m_\lambda(t_a; t_b)$ results in the mean rate of earthquakes $m_\lambda(M_{min} < m < M_{max}; t)$ for the entire range *m* = [M_{min} , M_{max}] in the time interval $t = [t_a, t_b]$. The probability $P[N = n; t_a, t_b]$, eq. 7, is obtained from the simulation arrays by counting the number of realizations with 0, 1, 2...*n* occurrences in a magnitude range, and dividing by the number of realizations.

Horn River Basin case example

We study the impact of induced seismicity on hazard analysis, using the recent activity in the Horn River Basin, Northeast B.C. as an example. Several studies have been made in the Horn River Basin due to the significant increase of seismicity related to the hydraulic fracturing activities conducted between Dec. 2006 and Dec.2011 (BC oil and gas commission, 2012), particularly in the Etsho area. The detected seismicity in the area was very low prior 2006, but with an important increase since Dec. 2006, particularly between 2010 and 2011 in line with the amount of human activity. Due to the lack of recorded seismicity at the Horn River Basin, we assume that the GR parameters before Dec. 2006, for natural seismicity, are based on the GR parameters described by the 2015 National seismic-hazard model of Canada (Halchuk et al., 2014). They use a 3 sets of GR parameters, that were normalized because the natural and induced seismicity cover different areas (IS, figure 1 (A)). These results in: (1) *a*=0.85, *b*=0.8685; (2) *a*=1.44, *b*=1.090; (3) *a*=0.26, *b*=0.64, with weights of 0.68, 0.16 and 0.16 respectively.

For the period between Dec. 2006 and Dec. 2011, we assume that the GR parameters are based on the catalog from Farahbod et al., (2015 b) which contains induced earthquakes in the Horn River Basin. This catalog consists of 338 events recorded between Dec. 2006 and Dec. 2011, with magnitudes ranging between *m*= [1.0, 3.6]. From the catalog, it is possible to distinguish 2 periods with clearly different recurrence statistics: a 1st period with lower earthquake rates between Dec. 2006 - Dec.2009, and a 2nd period with higher rates between Dec.2009 - Dec. 2011 (Figure 1(B)). After Dec. 2011 we assume that the seismicity rates return to their natural state. For both periods of induced seismicity, we estimate the following sets of GR parameters: *a*=4.52 and *b*=1.45 for the 1st period, and *a*=4.72 and *b*=1.21, for the 2nd period (Figures 1(C) and 2(A)). We use the Maximum likelihood (Aki, 1965) and the maximum curvature methods (Wiemer and Wyss, 2000) for fitting and magnitude of completeness, respectively.

We compute 5 independent simulations each comprising N_r=100,000 realizations with a time duration of N_t =120 months (Dec.2004-Dec. 2014). We consider a magnitude bin size = 0.01, and M_{max} =5.0 and M_{min}=2.5. Figure 2 (B) shows the predictions for the annual rate of earthquakes $\lambda\ (m_j < m < m_{j+1}, t)$, eq. 4, for the periods before and after the induced events, as well as the 1st and 2nd period of the induced seismicity. Both the theoretical and the quantities derived from the five Monte Carlo simulations are plotted. Using different time intervals $t = [t_a, t_b]$, we compute the probability $P[N = n; t_a, t_b]$ of *n* occurrences, eq. 7, within *m*= [2.5, 3.1] and *m*= [2.5,3.6], in order to study the impact of the induced seismicity in the Horn River Basin. Figure 2 (C) shows the probability of *n* occurrences for the intervals before and after the induced seismicity, and for the 1st period and 2nd period of induced seismicity. These results confirm the change in seismicity during the period of induced seismicity. The simulated and analytical results also agree with the statistics given by the catalog. For instance, the most likely number of events within *m*= [2.5,3.6], is 95 events. This is very close to the actual number of earthquakes *m*= [2.5,3.6] recorded between Dec. 2009 - Dec. 2011, which was 99 events.

Conclusions

A method to compute synthetic earthquake catalogs and associated occurrence earthquake statistics is developed for non-stationary seismicity, using Monte Carlo simulations. We demonstrate that the Poisson model remains valid for non-stationary induced seismicity. However, non-stationary GR parameters must be included in order to properly assess the hazard for this type of seismicity. Tests showed excellent agreements between analytical predictions and numerical results. In the simulated forecasts, we assume that the GR induced parameters are known. Next steps will include relationships between earthquake parameters and injection volumes, and extensions to handle spatial source distributions as well as GMPEs in order to generate a complete methodology for non-stationary PSHA.

Acknowledgements

The authors thank the sponsors of the Microseismic Industry Consortium for financial support, and Honn Kao for providing an updated event catalog for the Horn River area.

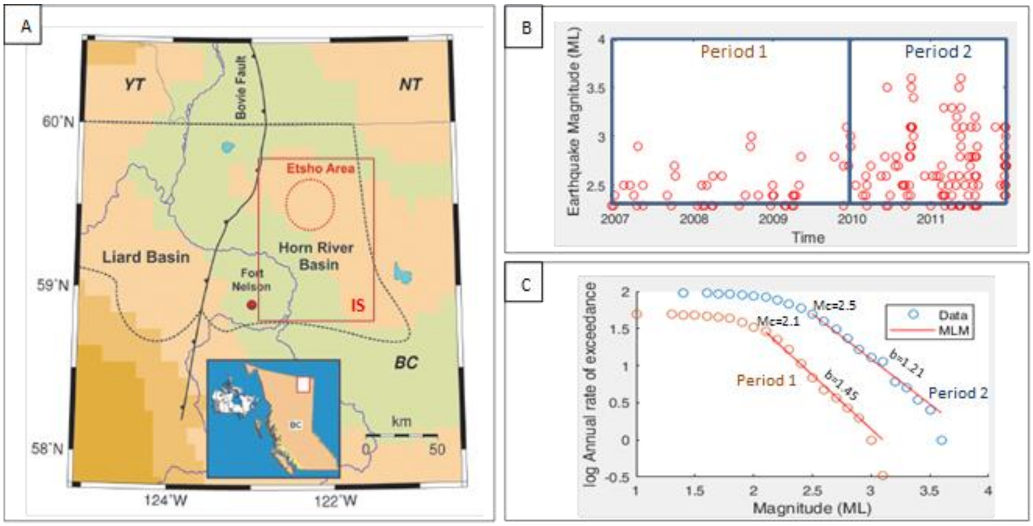

Figure 1: (A), Location of the Horn River Basin, Northeast British Columbia. The red square shows the area of the induced source (Modified from Farahbod et al.,2015a). (B) Earthquake magnitudes as a function of time. (C) Magnitude-frequency distribution plot for the two periods of induced seismicity.

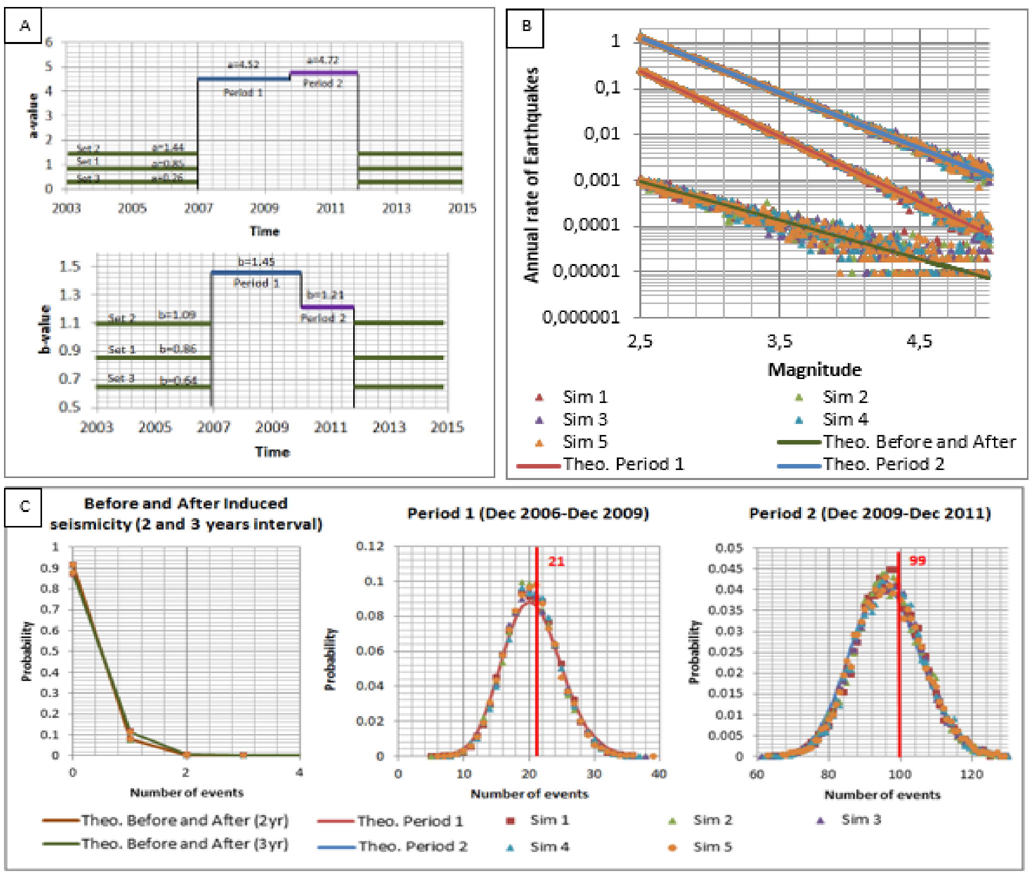

Figure 2: (A) Temporal evolution of the *a*- and *b*-values in the Horn River Basin. (B), Annual rate of earthquakes as a function of magnitude for the periods before and after the induced seismicity, as well as the 1st and 2nd period of induced seismicity. (C) Probability as a function of the number of occurrences within *m*=[2.5,3.6] for the periods before and after induced seismicity (Left, using 2 and 3 years time length) and the 2nd period of induced seismicity (Right, using 2 years time length). Probability as a function of the number of occurrences within *m*=[2.5,3.1] for the 1st period of induced seismicity (Centre, using 3 years time length). The actual number of events in the 1st and 2nd period of induced seismcity, are indicated with a red line. Theo=Theoretical values. Sim=simulation number, each using 100,000 realizations.

References

Aki, K., 1965, Maximum likelihood estimate of b in the formula logN= a-bM and its confidence limits: Bull. Earthq. Res. Inst, 43, 237-239.

Atkinson,G., Ghofrani, H., Assatourians, K., 2015, Impact of induced seismicity on the evaluation of seismic hazard: Some preliminary considerations: Seismological. Research Letters, 86, 3, 1009-1021.

Atkinson, G., Eaton, D., Ghofrani, H., Walker, D., Cheadle, B.,Schultz, R., Shcherbakov, R., Tiampo, K., Gu, J., Harrington, R., Liu, Y., van der Baan, M., Kao, H., 2016, Hydraulic fracturing drives induced seismicity in the western Canada sedimentary basin: Seismological Research Letters, 87, 3, 631-647.

Baker, W., 2008, An introduction to probabilistic seismic hazard analysis (PSHA). Version 1.3, doi: https://web.stanford.edu/˜bakerjw/ Publications/Baker_(2008)_Intro_to_PSHA_v1_3.pdf.

B.C oil and gas commission, 2012,Investigation of observed seismicity in the Horn River Basin, technical report, doi:www.bcogc.ca/node/8046/download?documentID=1270.

Cornell, C., 1968, Engineering seismic risk analysis: Bulletin of the Seismological Society of America,58,1583-1606.

Ellsworth, W., 2013, Injection induced earthquakes: Science, 341, 142-145.

Farahbod, A., Kao, H., Cassidy, J., Walker, D., 2015a,How did hydraulic-fracturing operations in the Horn River Basin change seismicity patterns in the northeastern British Columbia, Canada?: The leading edge, 34, 658-663.

Farahbod, A., Kao, H., Cassidy, J., Walker, D., 2015b, Investigation of regional seismicity before and after hydraulic fracturing in the Horn River Basin, northeast British Columbia: Canada J. Earth Sciences, 52, 112-122.

Gutenberg, R., Richter, C.F., 1944, Frequency of earthquakes in California: Bulletin of the Seismological Society of America, 34 , 185-188.

Halchuk, S., T. I. Allen, J. Adams, and G. C. Rogers, 2014,Fifth generation seismic hazard model input files as proposed to produce values for the 2015 National Building Code of Canada: Geol. Surv. Canada, Open- File Rept. 7576, doi: 10.4095/293907.

Sigman, K. 2013, Non-stationary Poisson processes and Compound (batch) Poisson processes.: Columbia Edu. doi:http://www.columbia.edu/˜ks20/4404-Sigman/4404-Notes-NSP.pdf.

Wiemer, S. and Wyss, M., 2000, Minimum magnitude of completeness in earthquake catalogs: examples from Alaska, the western United States, and Japan: Bulletin of the Seismological Society of America,90, no.4.859-869.