

Uncertainty Propagation through Probabilistic Seismic Hazard Analysis

Tenzin Yangkey
Duke University

December 15, 2019

Abstract

Probabilistic Seismic Hazard Analysis (PSHA) has been used widely over the last 50 years by seismologist and engineers to quantify seismic hazard level and develop building code requirements to build safer buildings for the future. Other applications include determining earthquake insurance rates, deciding safety criteria for nuclear power plants and making official national hazards maps for future safety. Aftermath the recent 2015 Gorkha earthquake in Nepal(a highly seismic prone country), two different PSHA studies (McNamara et al., 2017; Rahman Bai, 2018) were conducted to improve seismic research and these have resulted in better characterization of geometry of the Main Himalayan Thrust (MHT), more complex method of delineating seismic sources and more robust seismic hazard maps by employing recent seismic catalogue. Despite PSHA's popular use, some sources of uncertainty in the framework are present hence the need to fully encompass the uncertainties through further studies. In this paper, uncertainty deriving from 'b' value will be studied in a scenario based format. Similar to that of monte-carlo simulation, 'b' values are sampled and their relation with final rate of exceedance is studied to understand it's uncertainty contribution.

Contents

1	Introduction	4
1.1	Background	5
2	Theory	5
2.1	Magnitude Distribution	5
2.2	Distance Distribution	6
2.3	Ground Motion Prediction Equation	7
2.4	Total Probability Theorem	8
3	Data and Methodology	8
3.1	Introducing uncertainty in G-R relationship	8
3.2	Calculating standard deviation of Rate of Exceedance	9
3.3	Simplifying codes: Matrix Multiplication	10
4	Results and Discussion	11
4.1	Varying 'b' values	12
4.2	Calculating standard deviation	13
5	Conclusion	13
6	MATLAB Code	14
6.1	PSHA Code: Basic	14
6.2	Uncertainty in b	15
6.3	Magnitude	17
6.4	GMPE Function	17

List of Figures

1	Map of Nepal District Boundaries, with an inset map of Nepal's location on a world map (source: EERI, 2016)	4
2	An example of a area source with r as the varying distance (source: Baker, 2015)	6
3	An example of a line source (source: Baker, 2015)	7
4	Rate of Exceedance of PGA values with variance	10
5	Different MATLAB codes to compute $P(X > x)$ for each possibility	11
6	Written note on the process of using matrix for code simplification	11

7	Plot of rate of exceedance vs. PGA values	12
8	Plot of rate of exceedance for multiple PGA values with varying b values	12
9	Plot of rate of exceedance for different PGA values with standard deviation	13

1 Introduction

Nepal is a landlocked country lying in the middle of the Himalayan belt between China to the North and India to the south. The east-west length of the country lies along the Himalayan range and is about 800km whereas the width of the narrow north-south is roughly 190km. Nepal is known globally as the home of the Mount Everest. It is also known as one of the most active seismic zones in the world. The convergence of two continental plates namely Indian plate and Eurasian plate is the prime reason for frequent occurrence of earthquakes in Nepal. These plates converge at a relative rate of 40-50mm per year, resulting in the net uplift of the Himalayan mountain ranges by roughly 18mm per year (EERI, 2016). Such continuous tectonic activity has resulted in a long history of strong earthquakes hitting the region. So far, six large earthquakes (1255, 1408, 1505, 1833, 1934 and 2015) with magnitudes exceeding 7.5 has affected the region. Out of these earthquakes, the Great Nepal-Bihar earthquake in 1934 with M_w 8.7 was recorded as the strongest earthquake, magnitude wise. The epicenter of the earthquake was about 200km east from Kathmandu, but severe damage was experienced in Kathmandu Valley. It was estimated to have caused around 10,600 fatalities, severe damage of more than 126,000 houses and complete collapse of more than 80,000 buildings in the three main cities of the valley: Kathmandu, Bhaktapur and Patan (EERI, 2016).

In 2015, the catastrophic earthquake with moment magnitude M_w 7.8 hit Gorkha, Nepal on April 25 and followed by another strong aftershock of magnitude 7.3 on May 12 to the east of the initial epicentral area. 8,790 lives were lost, 22,300 people were injured and about 100,000 people were displaced as their houses were severely damaged by the earthquake and subsequent earthquake shocks. The total economic and financial loss incurred by the disaster was estimated to be 7 billion USD according to the post disaster assessment by Nepal's National Planning Commission (NPC). This is about half of Nepal's annual GDP! It was also reported that almost all sectors were severely affected by this unfortunate disaster.

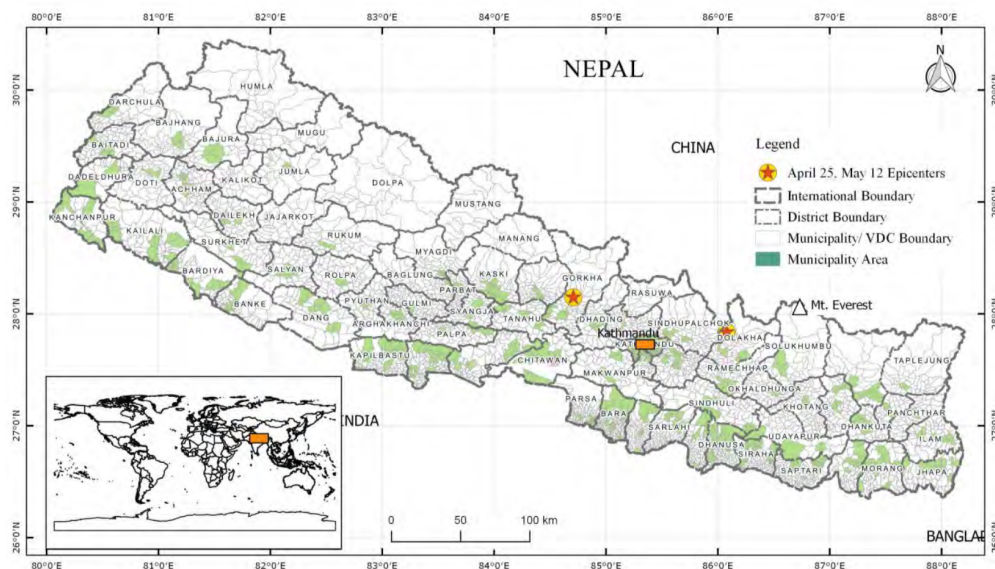


Figure 1: Map of Nepal District Boundaries, with an inset map of Nepal's location on a world map (source: EERI, 2016)

Therefore earthquake mitigation, preparedness and planning is a priority in Nepal. One of the fundamental step towards a more resilient community against hazard is understanding the hazard itself. This includes producing hazard maps that identifies the prone areas and design building structure that can withstand future earthquakes. Probabilistic Seismic Hazard Analysis (PSHA) is the standard method used by earthquake scientists and researchers in an effort to understand the hazard. In 2013, Ram and Guoxin (Ram Wang, 2013) for the first time used PSHA to determine hazard map by delineating the entire country into

23 seismic source zones. Similar studies were conducted with fragility function to understand risk as well (Chaulagain et al., 2015). After 2015 earthquake, numerous studies were conducted to continue the work of building resiliency against seismic risk in the future. Two studies in particular (Stevens et al., 2018; Rahman et al., 2018) used PSHA along with other techniques such as multiple seismic source models, logic tree (to reduce epistemic uncertainty), better characterization of the geometry of the Main Himalayan Thrust (MHT) to reach new understanding.

However, the current framework of PSHA has some sources of uncertainties that requires better characterization. One such source of error is 'b' value used in Gutenberg Richter Recurrence law which determines the probability distribution of earthquake magnitude. These 'b' values are obtained from earthquake catalogues. Therefore, an alternative way to study the uncertainty of 'b' and it's relation with the exceedance value is presented in this paper.

1.1 Background

This paper is part of a research initiative of increasing the understanding of seismic risk in Nepal. It is a collaborative research project started by Institute of Engineering, Pulchowk Campus, Nepal and Duke University, US in the summer of 2019. Currently, students from both institutions are researching on Hazard Analysis and Ground Motion Prediction Equations with an aim to build an earthquake early warning system for Kathmandu Valley in the future.

2 Theory

Probabilistic Seismic Hazard Analysis (PSHA) is a method that aims to quantify uncertainties associated with location, size, shake intensities of future earthquakes and combine them to understand the distribution of future shaking that may occur at a site (Baker, 2015). PSHA was first formalized by Cornell (1968) and further studied and popularized by McGuire, 1976, Bender and Perkins, 1987. Now PSHA is widely practiced as a standard part of the process for designing and building critical structure (Mulargia et al., 2017).

At the basic level, PSHA is composed of five steps according to Baker (2015). They are:

- Identify all earthquake sources capable of producing damaging ground motions.
- Characterize the distribution of earthquake magnitudes (the rates at which earthquakes of various magnitudes are expected to occur)
- Characterize the distribution of source-to-site distances associated with potential earthquakes.
- Predict the resulting distribution of ground motion intensity as a function of earthquake magnitude, distance, etc.
- Combine uncertainties in earthquake size, location and ground motion intensity, using a calculation known as the total probability theorem.

Using this framework, the probability of exceedance for different Peak Ground Acceleration (PGA) values are calculated as PGA measures the level of ground shaking intensity. This probability approach is different from deterministic approach where "worst-case" ground motion are needed to be identified for a fault source or an area source even. In this paper, the basic concept of each step are detailed in the following subsections.

2.1 Magnitude Distribution

Sources of earthquake can be categorized into two: line source and area source. Line source typically represent fault lines that passes through the site whereas area source incorporates the background

seismic activity in the region. Depending on the sources, the probability distribution of the magnitudes are calculated.

For tectonic faults or line source, Gutenberg-Richter recurrence law is used to determine the rate of earthquakes greater than some threshold m . The relationship is defined as follows:

$$\log_{10} \lambda_m = a + bm \quad (1)$$

Note that the constant values a and b are empirically derived from the past earthquake catalogues. The values vary depending on the tectonic fault as each of them are capable of producing earthquakes of various sizes.

Based on the G-R law, the cumulative distribution function for the magnitudes of earthquakes that are larger than m and smaller than M is as follows.

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

Finally, these CDF values are discretized to find the corresponding probabilities for each m value.

2.2 Distance Distribution

To predict ground shaking at a site, modeling the distribution of distance from earthquake to site is essential. It is generally assumed that earthquakes will occur with equal probability at any location on the fault (Bakers, 2015). This assumption of uniform distribution of earthquake occurrences simplifies the process of identifying the distance distribution as it is dependent on the geometry of the source. Note that the distribution varies depending on the sources: line source and area source.

For area source, consider a site located in the area source as shown in the figure below. The probability of an epicenter being located at a distance of less than r is equal to the area of a circle of radius r , divided by the area of circle of radius 100km (the value 100km can be changed).

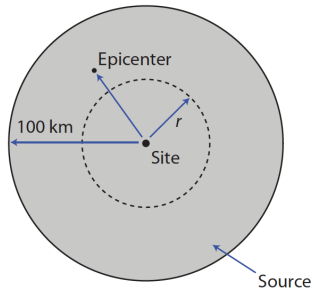


Figure 2: An example of a area source with r as the varying distance (source: Baker, 2015)

$$F_R(r) = P(R \leq r) \quad (3)$$

$$= \frac{\text{area of circle with radius } r}{\text{area of circle with radius } 100} \quad (4)$$

$$= \frac{\pi r^2}{\pi(100)^2} = \frac{r^2}{10,000} \quad (5)$$

It is important to note that equation 5 however is limited to the distance (r) between $0km$ and $100km$. For the line source, several assumptions are made to simplify the formula to calculate it. One such assumption is the position of the site at a certain distance r from the center of the fault. Following formula computes the distance distribution for line source or fault lines and the figure illustrates the location of site with respect to the fault. Note that equation (8) holds true for r values between $10km$ and $51km$ as stated in the example.

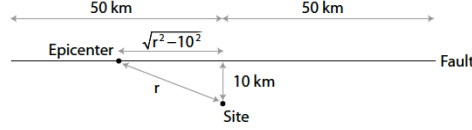


Figure 3: An example of a line source (source: Baker, 2015)

$$F_R(r) = P(R \leq r) \quad (6)$$

$$= \frac{\text{length of fault within distance } r}{\text{total length of fault}} \quad (7)$$

$$= \frac{2\sqrt{r^2 - 10^2}}{100} \quad (8)$$

This cumulative distribution function will be discretized and used to calculate the corresponding probability distribution in MATLAB.

2.3 Ground Motion Prediction Equation

Ground motion prediction equation or GMPE sheds light on how much the ground shakes during an earthquake. It predicts the probability distribution of ground motion intensity, as a function of many predictor variables such as the earthquake's magnitude, distance, faulting mechanism, the near-surface site conditions, the potential presence of directivity effects, etc (Baker, 2015). At present, Nepal doesn't have a region specific GMPE, rather it uses GMPE developed by neighbouring countries such as China (Ram Wang, 2013).

Generally, prediction models take this form for probability distribution of intensity.

$$\ln IM = \overline{\ln IM}(M, R, \theta) + \sigma(M, R, \theta) \cdot \epsilon \quad (9)$$

where $\ln IM$ is the natural log of the ground motion intensity measure of interest such as PGA (Peak Ground Acceleration) used in this paper. The predicted mean and standard deviation in the (9) are outputs of the ground motion prediction model. For the following example, GMPE proposed by Cornell et al. (1979) will be used to illustrate how it calculates the distribution of the ground motion intensity.

$$\overline{\ln PGA} = -0.152 + 0.859M - 1.803\ln(R + 25) \quad (10)$$

Using the output of mean and standard deviation from GMPE, the probability of exceeding any PGA level is calculated as follows since then natural logarithm of PGA was seen to be normally distributed.

$$P(PGA > x|m, r) = 1 - \phi\left(\frac{\ln x - \overline{\ln PGA}}{\sigma_{\ln PGA}}\right) \quad (11)$$

That's how the probability of exceeding any PGA value is calculated using GMPE.

2.4 Total Probability Theorem

Total Probability Theorem (TTP) is a probability concept that combines conditional probability with it's associated probability. In PSHA, TTP combines the conditional probability of exceedance ($P(IM > x|m_j, r_k)$) with distance and magnitude probability distribution ($P(R_i = r_k)$ and $P(M_i = m_j)$) to find the probability of intensity measure (IM) i.e. PGA exceeding a threshold pga $P(IM > x)$. The equation is as follows:

$$P(IM > x) = \int_{m_{min}}^{m_{max}} \int_0^{r_{max}} P(IM > x|m, r) f_M(m) f_R(r) dr dm \quad (12)$$

In the equation (13), integration symbol is used to solve for the total probability. However, it is needed to be discretized to solve easily in MATLAB. Following is the discretized version of the TTP. Additional variable introduced in equation (13) is the annual rate of earthquake occurrence denoted by $\lambda(M > m_{min})$.

$$\lambda(X > x) = \lambda(M > m_{min}) \sum_{j=1}^{n_M} \sum_{k=1}^{n_R} P(IM > x|m_j, r_k) P(M_i = m_j) P(R_i = r_k) \quad (13)$$

This is how PSHA determines the probability of ground motion shake at a particular site during an earthquake using earthquake magnitude value, source site distance and GMPE. This method is the main framework upon which uncertainty of 'b' value is studied.

3 Data and Methodology

3.1 Introducing uncertainty in G-R relationship

In the usual Probabilistic Seismic Hazard Analysis, a and b values of the G-R recurrence law are derived from earthquake catalogues using maximum likelihood method. In the latest research on PSHA in Nepal using multiple seismic model (Rahman et al., 2018), total of 23 different b values were determined for each seismic sources delineated and identified in the paper. The b values ranged from 0.5 to 0.9. The b-values were then used throughout PSHA method to calculate the final exceedance probability. This shows that the value b could be one source of uncertainty in this seismic analysis framework.

The value 'b' is obtained empirically from earthquake catalogues. These earthquake catalogue changes over time as more seismic record would be collected over time. For some catalogues, aftershock seismic data are removed by declustering method and for others, aftershock data are kept intact. These varying methods affects the b value in addition to the types of sources and faults on which the catalogues are based. Therefore, in this paper, a distribution of b values are considered to study the impact of these changing b values to the final exceedance value.

In this study, the b values with an interval of 0.05 were sampled and they are assumed to have log normal distribution. To find the probability of the value, the mean and the standard deviation of the sample should be given. In reality, the mean (m) of the b values fall close to 1 and standard deviation (std) can

be assumed to be 0.2. Based on these values, the probability distribution of the b values are calculated as follows:

$$\mu_{lnB} = \ln(\text{median of } B) \quad (14)$$

$$\sigma_{lnB} = \ln \left(1 + \left(\frac{\sigma_B}{\mu_B} \right)^2 \right) \quad (15)$$

$$F_B(b) = \phi \left(\frac{\ln b - \mu_{lnB}}{\sigma_{lnB}} \right) \quad (16)$$

In equation 5, the cumulative distribution of the b values are found. They are discretized to find the corresponding probability distribution. The probability distribution of 'b' value is then used throughout the PSHA framework to find the rate of exceedance for a given PGA value.

3.2 Calculating standard deviation of Rate of Exceedance

Using Theorem of Total Probability (TTP), varying probabilities of the b values can be incorporated in the PSHA methodology. Referring back to the equation of theorem of total probability in the theory section,

$$\lambda(X > x) = \lambda(M > m_{min}) \sum_{j=1}^{n_M} \sum_{k=1}^{n_R} P(IM > x | m_j, r_k) P(M_i = m_j) P(R_i = r_k) \quad (17)$$

The $P(B = b_i)$ can be introduced easily to this fundamental framework as follows:

$$\lambda(X > x) = \lambda(M > m_{min}) \sum_{l=1}^{n_B} \sum_{j=1}^{n_M} \sum_{k=1}^{n_R} P(IM > x | m_j, r_k, b_l) P(M_i = m_j) P(R_i = r_k) P(B_i = b_l) \quad (18)$$

The simplified version of the above equation is

$$\lambda(X > x) = \sum_l \lambda(X > x | B = b_l) P(B = b_l) \quad (19)$$

Assuming the simplified equation 8 is similar to the following equation of finding expected value of a distribution where the conditional probability $\lambda(X > x | B = b_1)$ is considered to be a value a_i found in the equation below.

$$E[A] = \sum_i a_i P[A = a_i] \quad (20)$$

If this is assumed, then the corresponding variance can be calculate as follows.

$$E[A] = \sum_i a_i P[A = a_i] \quad (21)$$

$$E[A^2] = \sum_i a_i^2 P[A = a_i] \quad (22)$$

$$\sigma_A^2 = E[A^2] - E[A]^2 \quad (23)$$

Therefore, the variance of rate of exceedance for varying b value can be calculated as follows.

$$E[\lambda(X > x)] = \sum_i \lambda(X > x|B = b_i)P(B = b_i) \quad (24)$$

$$E[\lambda^2(X > x)] = \sum_i (\lambda(X > x|B = b_i))^2 P(B = b_i) \quad (25)$$

$$\sigma_\lambda^2 = E[\lambda^2] - (E[\lambda])^2 \quad (26)$$

Following is a figure that depicts how variance can help us explain the impact of changing b value on the final value - rate of exceedance for each PGA value.

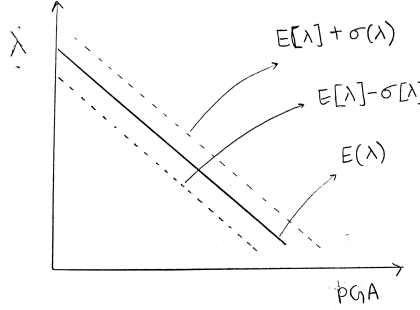


Figure 4: Rate of Exceedance of PGA values with variance

3.3 Simplifying codes: Matrix Multiplication

MATLAB is primarily used to run PSHA. The MATLAB code is written by the author based on Baker's Introduction to Probabilistic Seismic Hazard Analysis paper. For MATLAB calculation, multiple loops are needed to run some parts of the analysis. Here's an example:

When Theorem of Total Probability is applied, the probability of exceedance $P(X > x|m, r)$ is multiplied by corresponding probabilities of $P(M = m)$ and $P(R = r)$. For this multiplication, for loops are used initially to compute the values.

```

%% Multiple for loop version

% PExc_mr = a matrix of P(X>x|m,r) and they need to be multiplied by
% corresponding P(M=m) and P(R=r)

for i = length(PM)
    for j = length(PR)
        PExc_mr1(i,j) = PExc_mr(i,j)*PM(i)*PR(j);
    end
end

%% Simplified version

% multiply each element of PExc_mk by it's corresponding PM and PR values
% P(IM>x) = sum(sum(P(IM>x|m,r)* FM(m)*FR(r))) % [Baker, Eqn. 2.22]

PM_mat = PM'*ones(1,length(PR));
PR_mat = ones(length(PM),1)*PR;
PExc_mr1 = PM_mat.*PExc_mr.*PR_mat; % multiplying by PM and PR at each element

```

Figure 5: Different MATLAB codes to compute $P(X > x)$ for each possibility

Multiple use of for-loop is an inefficient method of computing the needed value as it is time consuming especially when the sample size is large. This process can be simplified and made more efficient by using matrix multiplication. As observed in the simplified version code of the figure 2, the $P(M = m)$ and $P(R = r)$ are transformed into matrix instead of a single array.

$$\begin{aligned}
 D &= (PM' * \text{ones}(1, N_r)) .* P_{Exc_mr} .* (\text{ones}(N_m, 1) * PR) \\
 &= \begin{bmatrix} PM_1 \\ PM_2 \\ \vdots \\ PM_{N_m} \end{bmatrix} [1 \dots 1] .* P_{Exc_mr} .* \begin{bmatrix} 1 \\ PR_1 \\ PR_2 \\ \vdots \\ PR_{N_r} \end{bmatrix} \\
 &= \begin{bmatrix} PM_1 \\ PM_2 \\ \vdots \\ PM_{N_m} \end{bmatrix} .* P_{Exc_mr} .* \begin{bmatrix} PR_1 & PR_2 & \dots & PR_{N_r} \end{bmatrix} \\
 &= \begin{bmatrix} PM_1 * P[X > x | M=m_1, R=r_1] * PR_1 \\ \vdots \\ PM_i * P[X > x | M=m_i, R=r_i] * PR_i \\ \vdots \end{bmatrix}
 \end{aligned}$$

Figure 6: Written note on the process of using matrix for code simplification

4 Results and Discussion

Seismic hazard is generally quantified to understand the severity and intensity based on different factors such as site location, distance from faults and etc. Probabilistic Seismic Hazard Analysis (PSHA) is one such framework that quantifies seismic hazard by determining the annual rate of exceedance for different peak ground acceleration (PGA) values i.e. when the rate of exceedance for 0.2g is 0.02, this means that the annual rate of ground shake above 0.2g is 0.02. Following is a plot of the rate of exceedance for different PGA values.

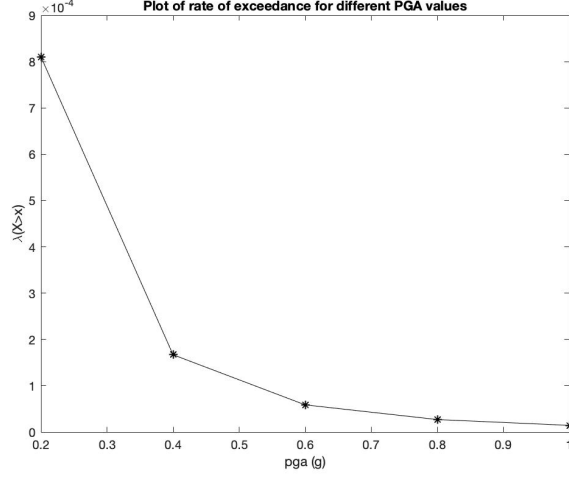


Figure 7: Plot of rate of exceedance vs. PGA values

In figure 7, it is observed that the rate of exceedance $\lambda(PGA > pga)$ decreases with the increase in PGA values. This observation makes sense as higher PGA values corresponds to higher magnitude which has smaller probability of it's occurrence, hence the lower rate of exceedance.

4.1 Varying 'b' values

The value 'b' is one of the factors that can affect the resulting rate of exceedance values. This value is based on the seismic fault as it's catalogues are used to derive 'b'. Determining the correct 'b' value is hard as it is always changing as faults move and has multiple factors that might influence it's final value. Therefore, a sample of b values are considered to study it's sensitivity to the final value of PSHA: the rate of exceedance for different PGA values. The figure on the left is a plot of rate of exceedance for different PGA values with multiple b values whereas the right figure overlays the plot the general rate of exceedance (figure 7) with the plot of rate of exceedance for multiple b values on the left.

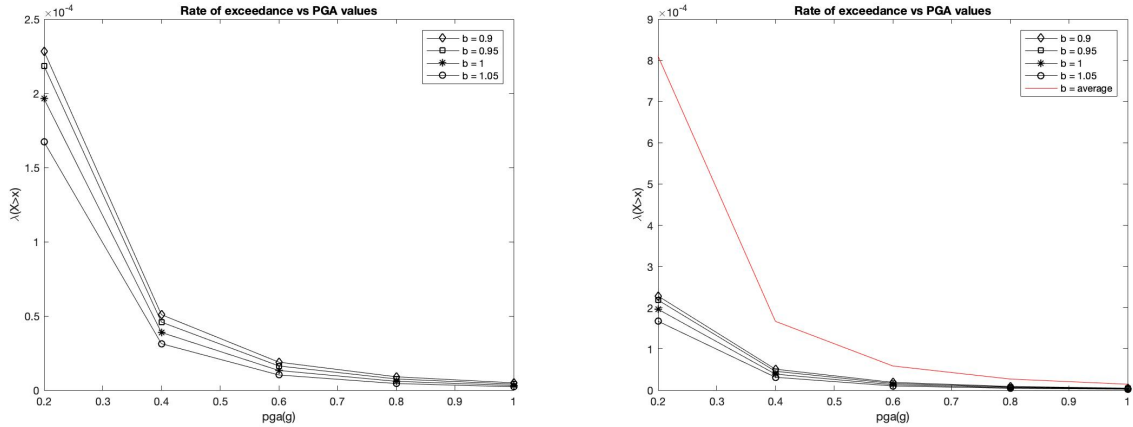


Figure 8: Plot of rate of exceedance for multiple PGA values with varying b values

From figure 8, a clear difference between two plots can be observed. The left photo shows the rate of exceedance plot for the sample b values. On the right side, the hazard curve from Figure 4 is plotted on the same plot as in the left. Before plotting, it was assumed that the additional hazard curve would lie among the cluster of lines formed by different b values such that the variance can be studied. However,

the original hazard curve turned out to have larger values than the rest of the lines from b values. The possible explanation for this observation is that the b values used in the sample are distributed with varying probabilities hence, each of the hazard lines are weighted accordingly. Therefore, the impact of varying b -value on the rate of exceedance can't be interpreted in the above plot.

4.2 Calculating standard deviation

Another approach is assuming the rate of exceedance $\lambda(X > x|b)$ as a value and using probability theorem to calculate the variance and standard deviation. The method of finding the associated standard deviation is explained in the method section of the paper. The plot of the rate of exceedance curve with the standard deviation is presented below.

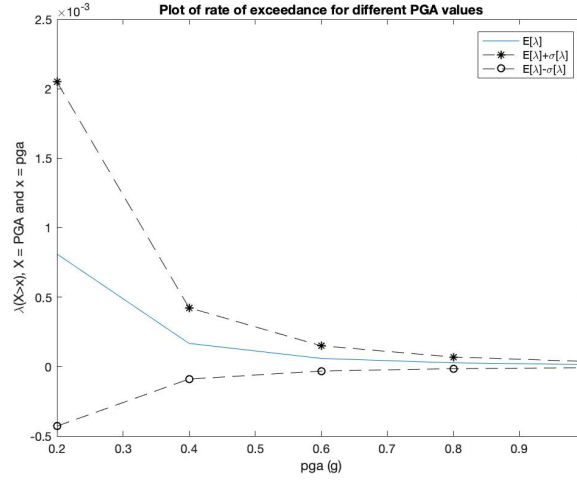


Figure 9: Plot of rate of exceedance for different PGA values with standard deviation

It can be deduced from figure 9 that the variation of the rate of exceedance $\lambda(X > x)$ decreases with the increase in PGA values. This means that varying b value is more sensitive to smaller PGA than bigger PGA value. However, this finding has relatively lower significance than it would have if large PGA values are more sensitive. This is because seismologist and engineers care less about the smaller PGA values than the large PGA values as larger PGA values or ground shakes would result in catastrophic effects.

5 Conclusion

In an attempt to quantify the uncertainty of ' b ' value on the PSHA final value: the rate of exceedance, the conditional probability $\lambda(X > x|B = b_1)$ is considered to be a value so that the general probability theorems of mean and variance can be applied to understand it's spread. The main result of the approach is that the spread is inversely related to the PGA value or the degree of ground shake. With higher PGA, the spread of rate of exceedance due to varying ' b ' value is smaller, hence the smaller sensitivity. For smaller PGA values, the spread is bigger which implies more uncertainty however, this finding smaller significance in reducing seismic hazard. Moreover, in this analysis, Cornell (1976) GMPE is used instead of a GMPE that represents Nepal better.

Moving forward, the analysis will be more robust with the region specific placement i.e. using region specific site, GMPE and fault lines. Since this is an active area of research, present and upcoming resources regarding the topic will be studied further.

6 MATLAB Code

6.1 PSHA Code: Basic

```

1 % Uncertainty Propagation method of PSHA
2 % Last edited: Oct 19th, 2019
3 % Written by Tenzin Yangkey
4
5 % The purpose of this script is to study how uncertainty propagates
6 % through the Probabilistic Seismic Hazard Analysis (PSHA) method.
7
8 % In this code, the magnitude distribution is defined by the G-R relationship
9 % of the fault line. The distance is characterized by a line source
10 % and it is dependent on that distribution. For site characteristics, we
11 % will use GMPE [Eqn. 2.14] in Baker's paper for the analysis
12
13 % Note that the GMPE used in this example is below
14 % lnPGA = -0.152 + 0.859 .* M - 1.803 * ln(d + 25) [Baker, Eqn. 2.14]
15
16 %% PART 1 - PROBABILITIES OF EARTHQUAKE MAGNITUDES AT A GIVEN FAULT
17 % are modeled by the Gutenberg Richter (G-R) relationship
18
19 % In this model, we are assuming that the fault lines are known in case of
20 % Kathmandu valley. We will account for it's uncertainty first.
21
22 m = [5 : 0.25 : 8]; % A set of earthquake magnitude exceedance thresholds
23 b = 1; % The G-R paramter from log(lambda_M>m) = a - b*m
24 rate = 0.02; % "base rate" of magnitudes exceeding the smallest magnitude being considered
25 nM = length(m); % FM(m) CDF of earthquake magnitudes from bounded G-R [Baker, Eqn. 2.4]
26 FM = (1 - 10.^(-b * (m - m(1)))) ./ (1 - 10.^(-b * (m(end) - m(1))));
27
28 % Discretization of CDF to find P(M = m(i)) means P[m(i) < M < m(i+1)]
29 PM = FM(2:length(FM)) - FM(1:length(FM)-1); % [Baker, Eqn. 2.6]
30
31 %% PART 2: DISTRIBUTION OF DISTANCES (Line source considered)
32
33 % LINE SOURCE - let's assume that the nearest site-source distance is 10km and longest distance is 51km.
34 % This means that the overall fault length is 2*sqrt(51^2 - 10^2) = 100km
35
36 rmin = 10;
37 rmax = 51;
38 fault_length = 100;
39
40 r = [rmin: 2 : rmax]; % Viable radii between a site and a line fault
41 nR = length(r);
42
43 % FRL(r) = CDF of line source-site distances
44 FRL = 2 * sqrt(r.^2 - rmin.^2) / fault_length; % [Baker, Eqn. 2.11]
45
46 % Discretization of the CDF of distance: P(R = r(i)) means P[r(i) < R < r(i+1)]
47 PR = FRL(2:length(FRL)) - FRL(1:length(FRL)-1);
48
49 %% PART 3: GMPE EQUATION TO FIND THE EXCEEDANCE RATE
50
51 % using GMPE_Cornell_1979 as a function
52 [meanlnPGA, sdlnPGA] = GMPE_Cornell_1979(m,r);
53
54 % A range of peak ground acceleration (PGA) exceedance thresholds
55 pga = [0.2:0.2:1];
56 PExc_mk = zeros(length(pga),length(PM),length(PR)); % initializing Probability of Exceedance matrix
57
58 % initialize
59 Deg = zeros(length(pga),length(PM),length(PR)); % Deaggregation matrix
60 lamda_pga = zeros(1,length(pga)); % lamda(x > pga)
61
62 for k = 1:length(pga)
63
64     % Evaluate Exceedance Rates for the set of pga thresholds specified above
65
66     zk = (log(pga(k)) - meanlnPGA(:,,:)) / sdlnPGA; % Standardized version of the ln(PGA)
67     PExc_mr = 1 - normcdf(zk); % Probability of exceedance, given magnitude and distance
68     PExc_mr = PExc_mr((1:length(m)-1),(1:length(r)-1)); % Re-adjust vector length
69
70     % PART 4: THEOREM OF TOTAL PROBABILITY - combines all parts of PSHA
71
72     % multiply each element of PExc_mk by it's corresponding PM and PR values
73     % P(IM>x) = sum(sum(P(IM>x|m,r)* FM(m)*FR(r))) % [Baker, Eqn. 2.22]
74
75     PM_mat = PM*ones(1,length(PR));
76     PR_mat = ones(length(PM),1)*PR;

```

```

77 % PExc_mr1 = PM.mat.*PExc_mr.*PR.mat; % multiplying by PM and PR at each element
78 % PExc_mr_annual = rate * PExc_mr1;
79 % lamda_pga(k) = sum(sum(PExc_mr_annual));
80 %
81 % % dividing the PExc_mk_annual by lamda helps find the weighted values
82 % Deg(k, :, :) = PExc_mr_annual/lamda_pga(k); % Deaggregation values from 0 - 1
83 %
84 % easiest path
85 PExc_mr1 = PM*PExc_mr*PR'; % multiplying by PM and PR at each element
86 lamda_pga(k) = rate * PExc_mr1;
87
88 end
89
90 %% SOURCES
91 %
92 % Baker, Jack W. (2015) Introduction to Probabilistic Seismic Hazard Analysis.
93 % White Paper Version 2.1, 77 pp.
94 %
95 % Rachael Lau's Intro codes 1a, 1b, 2, 3 and 4 on Probabilistic Seismic Hazard Analysis (PSHA)

```

6.2 Uncertainty in b

```

1 % Uncertainty in b-value
2 % Last edited: Dec 9th, 2019
3 % Written by Tenzin Yangkey
4
5 % In this case, the montecarlo simulation of the b value is solved via
6 % Theorem of Total Probability method. The distribution of b value is
7 % incorporated in the TTP itself.
8
9 %% Line Source
10
11 % PART 1 – PROBABILITIES OF EARTHQUAKE MAGNITUDES AT A GIVEN FAULT
12 % are modeled by the Gutenberg Richter (G-R) relationship
13
14 % In this model, we are assuming that the fault lines are known in case of
15 % Kathmandu valley. We will account for it's uncertainty first.
16
17 mmin = 5;
18 mmax = 8.5;
19 dm = 0.25;
20 m = [mmin : dm : mmax]; % A set of earthquake magnitude exceedance thresholds
21 n_m = length(m);
22
23 % Distribution of b values
24 b = [0.9:0.05:1.1]; % The G-R paramter from log(lambda_M>m) = a - b*m
25 n_b = length(b);
26 mu = 1; % mu = mean of b value
27 std = 0.2; % std = standard deviation of b value
28 b_mu = log(mu); % log of the mu (log normal distribution)
29 b_std = sqrt(log(1+(std/mu)^2)); % log normal of the std
30 CDF_b = logncdf(b, b_mu, b_std); % Cumulative probability distribution of b values
31 PB = CDF_b(2:length(CDF_b)) - CDF_b(1:length(CDF_b)-1); % discretization of the probabilities
32
33 % Probability of m given b (in a matrix format)
34 PM_b = zeros(length(PB), length(m)-1); % PM_b = probability of m given b
35 for i = 1:length(PB)
36 [PM_b(i, :)] = magnitude(m, b(i)); % probability of m given b
37 end
38 % Probability of b in matrix form (n_b x n_m)
39 PB1 = zeros(length(PB), length(m)-1); % Initialization
40 for j = 1:length(m)-1
41 PB1(:, j) = PB; % probability of b values
42 end
43
44 % Probability of magnitudes
45 PM = PB*PM_b; % probability of magnitude without varied b values – already weighted
46 PMM = PB1.*PM_b; % probability of magnitude with varied b values vertically
47
48 % Base rate
49 rate = 0.02; % "base rate" of magnitudes exceeding the smallest magnitude being considered
50
51 %% PART 2: DISTRIBUTION OF DISTANCES (line source considered)
52
53 % LINE SOURCE – let's assume that the nearest site-source distance is 10km and longest distance is 51km.
54 % This means that the overall fault length is 2*sqrt(51^2 - 10^2) = 100km
55
56 rmin = 10;

```

```

57 rmax = 51;
58 dr = 2;
59 r = [rmin: dr : rmax]; % Viable radii between a site and a line fault
60 fault_length = 100;
61
62 [PR] = line_dist(r,rmin,fault_length);
63
64 %% PART 3: GMPE EQUATION TO FIND THE EXCEEDANCE RATE
65
66 % using GMPE-Cornell-1979 as a function
67 [meanlnPGA,sdlnPGA] = GMPE_Cornell_1979(m,r);
68
69 % A range of peak ground acceleration (PGA) exceedance thresholds
70 pga = [0.2:0.2:1]; % pga values from 0.2g to 1g is considered
71
72 % Initialize
73 PExc_mk = zeros(length(PM),length(PR)); % initializing Probability of Exceedance matrix
74 lamda_pga = zeros(1,length(pga)); % lamda(x > pga): the original values
75 lamda_pgaB = zeros(length(pga),length(PB));
76 lamda_pgaSQ = zeros(length(pga),length(PB));
77
78 for k = 1:length(pga)
79
80     % Evaluate Exceedance Rates for the set of pga thresholds specified above
81     zk = (log(pga(k)) - meanlnPGA(:,:)) / sdlnPGA; % Standardized version of the ln(PGA)
82     PExc_mr = 1 - normcdf(zk); % Probability of exceedance, given magnitude and distance
83     PExc_mr = PExc_mr((1:length(m)-1),(1:length(r)-1)); % Re-adjust vector length
84
85     %% PART 4: THEOREM OF TOTAL PROBABILITY – combines all parts of PSHA
86
87     % multiply each element of PExc_mk by it's corresponding PM and PR values
88     %  $P(IM > x) = \sum(\sum(P(IM > x|m, r) * FM(m) * FR(r)))$  % [Baker, Eqn. 2.22]
89
90     % Part I
91     PExc_mr1 = PM * PExc_mr * PR'; % TTP: multiplying by PM and PR at each element
92     lamda_pga(k) = rate * PExc_mr1; % Rate of exceedance values with varying PGA values
93
94     % Part II – maintain variation in b
95     PExc_mrv = PM_b * PExc_mr * PR'; %  $P(X > x|b)$ 
96     lamda_pga_b = rate * PExc_mrv; %  $\lambda(X > x|b)$ 
97     lamda_pgaB(k,:) = lamda_pga_b * PB'; %  $\lambda(X > x)$  or  $E[\lambda(X > x)]$ 
98     lamda_pgaSQ(k,:) = (lamda_pga_b.^2) * PB'; %  $\lambda(X > x|b)^2 * P(B=b)$  or  $E[\lambda(X > x)^2]$ 
99 end
100
101 %% Variance and Mean Calculation
102 var = lamda_pgaSQ - (lamda_pgaB.^2); % variance of rate with varying b and PGA values
103 var = var'; % transpose
104 variance = sum(var); % PGA variance across PGA values — CAN ONE ADD VARIANCES?
105 std = variance.^0.5;
106 lamda_pgaB = lamda_pgaB'; % transpose
107 mean = sum(lamda_pgaB); % mean of the varying b with PGA values
108
109 %% plotting
110
111
112 % Figure 1: Plot of rate of exceedance for different PGA values
113 figure(1)
114 plot(pga,lamda_pga,'k*-')
115 title('Plot of rate of exceedance for different PGA values')
116 xlabel('pga (g)')
117 ylabel('\lambda(X > x)')
118
119 % Figure 2
120 figure(2)
121 plot(pga,lamda_pga)
122 hold on
123 plot(pga,lamda_pga+std,'k*-')
124 hold on
125 plot(pga,lamda_pga-std,'ko--')
126 legend('E[\lambda b]', 'E[\lambda b]+\sigma[\lambda b]', 'E[\lambda b]-\sigma[\lambda b]')
127 title('Plot of rate of exceedance for different PGA values')
128 xlabel('pga (g)')
129 ylabel('\lambda(X > x), X = PGA and x = pga')
130
131 % Figure 3: Plot of rate of exceedance for different PGA values with
132 % different b values
133 figure(3)
134 plot(pga,lamda_pgaB(1,:), 'kd-')
135 hold on
136 plot(pga,lamda_pgaB(2,:), 'ks-')
137 hold on

```



```

138 plot(pga,lamda_pgaB(3,:), 'k*-')
139 hold on
140 plot(pga,lamda_pgaB(4,:), 'ko-')
141 legend('b = 0.9', 'b = 0.95', 'b = 1', 'b = 1.05')
142 title('Rate of exceedance vs PGA values')
143 xlabel('pga(g)')
144 ylabel('\lambda(X>x)')
145
146 % Figure 4: Plot of figure 1 and figure 3 on the same plot
147 figure(4)
148 plot(pga,lamda_pgaB(1,:), 'kd-')
149 hold on
150 plot(pga,lamda_pgaB(2,:), 'ks-')
151 hold on
152 plot(pga,lamda_pgaB(3,:), 'k*-')
153 hold on
154 plot(pga,lamda_pgaB(4,:), 'ko-')
155 hold on
156 plot(pga,lamda_pga, 'r-')
157 legend('b = 0.9', 'b = 0.95', 'b = 1', 'b = 1.05', 'b = average')
158 title('Rate of exceedance vs PGA values')
159 xlabel('pga(g)')
160 ylabel('\lambda(X>x)')

```

6.3 Magnitude

```

1 % magnitude function
2
3 % This function is used in the main code that studies uncertainty propagation method of PSHA
4 % Last edited: Oct 19th, 2019
5 % Written by Tenzin Yangkey
6
7 function [PM] = magnitude(m,b)
8 % FM(m) CDF of earthquake magnitudes from bounded G-R [Baker, Eqn. 2.4]
9 FM = (1 - 10.^(-b .* (m - m(1)))) ./ (1 - 10.^(-b .* (m(end) - m(1))));
10 % Discretization of CDF to find P(M = m(i)) means P[m(i) < M < m(i+1)]
11 PM = FM(:,(2:length(FM))) - FM(:,(1:length(FM)-1)); % [Baker, Eqn. 2.6]
12 end

```

6.4 GMPE Function

```

1 % GMPE FUNCTION
2 % GMPE_Cornell_1979
3
4 % This function is used in the main code that studies uncertainty propagation method of PSHA
5 % Last edited: Oct 19th, 2019
6 % Written by Tenzin Yangkey
7
8 function [meanlnPGA,sdlnPGA] = GMPE_Cornell_1979(m,r)
9 % finding the mean log PGA to compare it with log PGA to find the z value
10 meanlnPGA = zeros(length(m),length(r));
11 sdlnPGA = 0.57; % Standard Deviation of ln PGA
12 for i = 1:length(m) % looping over the different values of magnitude
13     for j = 1:length(r) % looping over different values of distance
14         meanlnPGA(i,j) = -0.152 + 0.859 .* m(i) - 1.803 * log(r(j) + 25); % [Baker, Eqn. 2.14]
15     end
16 end
17 end

```

References

- [1] Baker, J. W. (n.d.). Introduction to Probabilistic Seismic Hazard Analysis. 2013, 79.
- [2] Chaulagain, H., Rodrigues, H., Silva, V., Spacone, E., Varum, H. (2015). Seismic risk assessment and hazard mapping in Nepal. *Natural Hazards*, 78(1), 583–602. <https://doi.org/10.1007/s11069-015-1734-6>
- [3] Musson, R. M. W. (2000). The use of Monte Carlo simulations for seismic hazard assessment in the U.K. *Annals of Geophysics*, 43(1). <https://doi.org/10.4401/ag-3617>
- [4] Rahman, M. M., Bai, L. (2018). Probabilistic seismic hazard assessment of Nepal using multiple seismic source models. *Earth and Planetary Physics*, 2(4), 327–341. <https://doi.org/10.26464/epp2018030>
- [5] Rajaure, S., Paudel, L. P. (2018). A comprehensive earthquake catalogue for Nepal and its adjoining region. *Journal of Nepal Geological Society*, 56(1), 65–72. <https://doi.org/10.3126/jngs.v56i1.22747>
- [6] Stevens, V. L., Shrestha, S. N., Maharjan, D. K. (2018). Probabilistic Seismic Hazard Assessment of Nepal. *Bulletin of the Seismological Society of America*, 108(6), 3488–3510. <https://doi.org/10.1785/0120180022>
- [7] Thapa, D. R. (2018). Seismicity of Nepal and the Surrounding Region. *Bulletin of the Department of Geology*, 83–86. <https://doi.org/10.3126/bdg.v20i0.20727>
- [8] Ram, T. D., Wang, G. (2013). Probabilistic seismic hazard analysis in Nepal. *Earthquake Engineering and Engineering Vibration*, 12(4), 577–586. <https://doi.org/10.1007/s11803-013-0191-z>