Development of Seismic Source Zonation Maps, Seismic Hazard Maps (Using Multiple Approaches) and GMPE for North East India

A thesis
submitted in partial fulfilment of the requirements
for the degree of

Doctor of Philosophy

By

Niranjan Borah

Department of Civil Engineering

Indian Institute of Technology Guwahati

Guwahati-781039, India

April 2024

Dedicated to my parents……..

and

to them ……..

who are working hard to make this world

a better place

for all the species

Statement

I do hereby declare that the matter embodied in this thesis entitled "**Development of Seismic Source Zonation Maps, Seismic Hazard Maps (Using Multiple Approaches) and GMPE for North East India**" is the outcome of investigations carried out by me in the Department of Civil Engineering, Indian Institute of Technology Guwahati, Guwahati, Assam, India.

In keeping with the general practice of reporting scientific observations, due acknowledgements have been made whenever the work described is based on findings of other investigators.

Niranjan Borah

Place: IIT Guwahati

Niranjan Borah

Date: 18-04-2024

(Enroll no. 186104024)

Certificate

This is to certify that the thesis entitled "**Development of seismic source zonation maps, seismic hazard maps (using multiple approaches) and GMPE for North East India**" being submitted by **Mr. Niranjan Borah**, Enrollment number **186104024**, to the **Indian Institute of Technology Guwahati** for the award of the degree of **Doctor of Philosophy** is a record of bona fide research work carried out by him. The thesis work in my opinion has reached the requisite standard for the degree of **Doctor of Philosophy**.

The results embodied in this thesis have not been submitted in any other University/Institute for the award of any Degree or Diploma.

Date: 18-04-2024

Place: Guwahati

Dr. Abhishek Kumar
Associate Professor
Indian Institute of Technology Guwahati,
Assam, India-781039

Acknowledgement

The work presented in this **thesis** has been made possible by the generous and unobstructed support from numerous people. I feel the opportunity to pen down an acknowledgement, is one of the finest privileges and an equally great responsibility that one can have. It is a privilege to remember all those who helped in making this scientific journey possible.

First and foremost, I would like to express my sincere gratitude and heartfelt thanks to my mentor and guide, **Dr. Abhishek Kumar** for giving me an opportunity to be a part of his amazing group. I am grateful to his tremendous support, invaluable guidance in the technical and non-technical field and for the patience showed. I am grateful to the motivational speech received from him and for understanding me in the whole journey. The wonderful support, valuable suggestions and encouragement are the main factors behind shaping my thesis work.

I would also extend my gratitude to the members of my Doctoral committee, **Dr. Anil Kumar Mishra**, **Dr. Ravi K**, and **Dr. Hanumant Singh Shekhawat** for their guidance, constructive comments and inventive suggestions at every stage of my PhD research.

I would also like to thank the course instructors of the subjects I took during my course work: **Prof. Rajib Kumar Bhattacharjya**, and **Dr. Rishikesh Bharti**. The knowledge obtained in their classes are directly or indirectly applied in my research.

I would like to thank former and present Heads of the Department of Civil Engineering: **Prof. Chandan Mahanta, Prof. Sharad Gokhale, and Prof. Rajib Kumar Bhattacharjya**. I would like to acknowledge the Department of Civil Engineering for providing the infrastructural facility for my research work. I thank all faculty members of the Department of Civil Engineering for their support and teachings towards the building of a greater technical society. Besides, I would like to thank all the technical and non-technical staff, and workers of the Department of Civil Engineering for their visible and invisible supports.

My next thanks go to IIT Guwahati which gave me an opportunity of embarking on this fine endeavour of learning. I thank all former and present directors of IIT Guwahati and the entire IIT Guwahati family for providing me with an intellectually apt environment to carry out my PhD research. I acknowledge the aesthetic pleasure I got from the beauty of IIT Guwahati campus, the lakes, the hills, and the flora and fauna.

Next, I would thank the funding agencies that provided me the financial support to conduct my research. I am grateful to Ministry of Education (formerly Ministry of Human

Resource Development) for the fellowship during my PhD tenure. I thank 17th WCEE Organizing Secretariat for the Grant-in-Aid that I received.

I am really grateful to my seniors **Dr. Olympa Baro**, **Dr. Harinarayan NH**, **Dr. Joy Kumar Monda**l for the sharing of knowledge, experiences and for taking caring of me. I also thank all the friends I found in the Department of Civil Engineering: **Himanshu**, **Akhila**, **Manash**, **Prinza**, **Safik**, **Deepak**, **Alok**, **Ankti**, **Chandrashekhar** and **Surender**. I thank our current research group members **Surender**, **Abhishek**, **Hrik**, **Sudeepta**, **Jenny**, **Abhileen**, **Aman,** and **Prajnadeep** for sweet memories. I also thank my Juniors from Bachelor and Masters, **Avinash**, **Sujata**, **Amit**, **Argha**, **Tharunkumar, Rajkamal, Tawfig**, and **Rajat** for the serious and sweet interactions related to technical, non-technical and other life related discussions.

I acknowledge the technical help I got from my friend **Dr. Ashish Bahuguna** and software related help from my friend **Ankur Saikia**. I thank my friends **Alions Basumatary** and **Pawan Kr Singh** for the help I have been receiving since my B.E.

I also thank my labmates: **Dr. Tanmay**, **Dr. Subhra**, **Dr. Rahul**, **Bari**, **Prem**, **Nayan, Banchiva, Rajib**, **Aditya, Samrat, Sureka, Anulekha, Naveen, Jenny and Manohara** for making the workplace an excellent one for doing work as well as to enjoy the environment and throwing parties outside.

A very big thanks to all the people who have made my stay in IIT Guwahati as a memorable and enjoyable journey. I am grateful for the presence of my friend **Arnab Sarmah** during this period in the Campus. I thank my friend **Sumit Singha** for making an enjoyable environment near me. I thank sister **Udangshree** for making my stay in the campus enjoyable. I would like to thank **Animesh Borah** for sharing information related to academic world which lead my path to the current location. I thank all the people who joined me during badminton games. Your **smiles** during the game made the **bright** memories which I can recall in future.

I want to thank all my B.E. and MTech teachers who developed my foundation in the field of Engineering. My thanks go to my B.E. project supervisor **Dr. Atanu Kumar Dutta** and MTech supervisor **Dr. Bijan Kumar Roy** for introducing me to the world of scientific research.

Here, I want to thank all my teachers irrespective of the domains for making me who I am today. I specially thank **Mr. Paramananda Baruah** for helping me to generate curiosity towards science and mathematics, and **Mrs. Madhu Mita Baruah** for generating my curiosity towards the Assamese Language and literature.

I want to thank Grammarly, Chatgpt, QuillBot for the support while writing manuscripts, by not using any form of AI for writing this acknowledgement section.

It is hard to maintain the mental and physical fitness during this period of life. There are some people I will give my credit to. I thank **Rashmi Madam** for teaching me how to do stress reducing Yoga, specially, the Pranayama. I thank the entire medical section of IIT Guwahati for their care during the time of my low health. I thank **Dr. Joe Dispenza** and **Frank Herbert** for their books that gave me hope and courage to face the real world.

Finally, I would like to express my heartfelt gratitude to my parents for their constant support and encouragement in every phase of my life. I am deeply indebted to them for all the love and care I received during my hardships. I thank my sister, grandmother, cousins and all of my relatives for standing with me during necessity.

Lastly, I thank my special friend, **Dawa Zangmu Sherpa**, for her constant support during this phase of my life. I am grateful for her invaluable life advice and motivation. The motivation I received from her is the reason for my goodwill in research and personal life.

Niranjan Borah

Abstract

The Northeast India (NE India) region is one of the most seismically active regions in India. The high seismic activity of the NE India region can be seen in its long history of earthquakes (EQs), which have triggered widespread damage. In the past, two great EQs, 1897 and 1950 Assam EQs, occurred in the area within a very short span of time. Most of the places in the region are situated in Zone V, as per IS-1893 part 1 (2016). Due to its high seismicity, many researchers studied different parts of the area and performed seismic hazard analyses to find the seismic potential of the area. For these researches, they considered the past available EQ events and sources information. It is seen that such information is incomplete for an extended length of time. This data gap results in an underestimate of the area's seismic hazard. Despite the fact that these researchers computed with both linear and areal sources, each study considers only one source model.

In this study, an attempt is made to delineate seismic source zones based on past EQ locations and seismicity parameters. The hierarchical clustering technique is used to cluster past EQs based on their latitudes and longitudes, resulting in 12 seismic source zones. This seismic source zonation enhances the likelihood of a comprehensive representation of seismic activity by segregating zones with a more uniform distribution of seismicity throughout each zone. Further, the spatial distribution of seismicity parameters is determined based on the Least Square Method and Maximum Likelihood Method. By utilising this spatial distribution of seismicity parameters, seismicity-based zonation is carried out. The seismicity-based zonation is combined with the EQ location-based zonation to obtain the final seismic source zonation of NE India. This process divided the entire NE India into 11 source zones having a more uniform distribution of past earthquakes and seismicity parameters.

The present study conducts PSHA for NE India using both traditional PSHA methodology and Kernel density-based methodology. Results from traditional PSHA methodology show that mean peak ground acceleration (PGA) values for Design Basis Earthquake (DBE) range from 0.23g to 0.51g, whereas for the Maximum Credible Earthquake (MCE), the range is from 0.40g to 0.85g. The eastern region of Manipur and the southern region of Nagaland exhibit the most significant seismic hazard level due to the substantial occurrence of prior EQs and the size of seismic events in these areas. Sikkim and certain Arunachal Pradesh and Assam regions have the least seismic hazard level. The deaggregation of seismic hazards for eight important cities

in NE India reveals the contribution of different magnitudes and distances to the seismic hazard level. The deaggregation results help to understand the contribution of seismic sources around cities. Strong EQs mainly control seismic hazards in NE India, both for MCE and DBE. Itanagar receives most contributions from major EQs occurring at a larger distance, while other cities have significant contributions from larger EQs happening at close distances.

The present study utilises both isotropic and anisotropic Kernels for PSHA. The results obtained from Kernel-based PSHA suggest that the overall seismic activity is higher for the Indo-Burma Ranges, Sikkim Himalaya, Arunachal Himalaya, Shillong Plateau, and Assam Valley Region. It is observed from the PSHA result that the mean PGA range for the MCE and DBE scenario of NE India is from 0.45 to 0.85g and 0.24 to 0.57g, respectively. The location-wise difference in PGA value is observed in the case of the isotropic and anisotropic Kernel. For some locations near faults, PGA values obtained in the case of the anisotropic Kernel are more than those obtained in the case of the isotropic Kernel. Similarly, it is also found that in the case of the catalogue, where EQs complete with respect to magnitude and time are present, the PGAs obtained are higher than that obtained by using the catalogue where all EQs with magnitude more than 7Mw but incomplete with respect to time are present.

The present study also develops a GMPE for NE India for different SCs based on regional parameters. The developed GMPE is based on synthetic ground motions generated for the bedrock condition using the EXSIM program. The present work also estimates the site coefficients, which can be used to obtain the ground motion level at the surface of different SCs. The effectiveness of the proposed GMPE is verified by comparing the predicted ground motion response spectra with regional records. The results show that the proposed GMPE is able to predict response spectra for SCs A, B, C, and D across NE India close to the actual record, giving confidence about the outcome of the GMPE.

Table of Content

List of Figures

List of Tables

List of Symbols

Vs	Shear wave velocity
Su	Undrained shear strength
V_{s30}	Average Shear wave velocity at top 30 m
t_s	Time window
d_s	Distance window
M	Magnitude of earthquake
MAXC	Maximum curvature
λ	Unbiased mean rate
σ_λ	Standard deviation
n	Number of time bins
T	Time interval
KL	Krzanowski and Lai index
K	Cluster number
d	Dimensions of data
WK_K	Pooled within cluster sum of squares of K partition
a, b	Seismicity parameters
m_i	Value of the central magnitude in i^{th} interval
t_i	Periods of completeness in years for the i^{th} interval.
$\bar{m}$	Average magnitude of the events in the catalogue.
N_{Mc}	Total numbers of EQs per year greater than Mc
N_T	the total number of EQ events
x	Raw value of a parameter.
$\bar{x}$	Mean of the value of the variables.
x_N	Normalised value of the parameter.
W_k	Weight considered to redistribution of the value of seismicity parameter from a zone to source *k*
Mw	Moment magnitude of EQ
M_{max}^{obs}	maximum observed magnitude
MmaxRWC	maximum possible magnitude estimated by using Robson–Whitlock–Cooke procedure

M_{n-1}	second largest observed magnitude.
Mp	maximum possible magnitude
Mmax-fault	maximum possible magnitude estimated based on fault length
v_z	annual frequency of exceedance of ground motion level "z"
$\lambda_n(m_i)$	annual frequency of EQ occurrence of magnitude m_i on n^{th} seismic source.
$P_n(R = r_j \| m_i)$	probability of occurrence of an EQ having magnitude m_i on n^{th} source at a distance r_j away from the site of interest.
$P(Z > z \| m_i r_j)$	probability of exceedance of ground motion level z for an EQ of magnitude m_i located at a distance r_j from the site of interest.
$P_E(z)$	probability of exceedance of ground motion of level z at a site, in a given exposure period
w_k	weights for *k-th* GMPE
LLH_k	LLH of the k-th GMPE
$g(x_i)$	likelihood of the model for observation x_i,
v_{PGA}	annual frequency of exceedance of PGA
m_{mb}	the mean magnitude of the magnitude bin
r_{mb}	the mean distance of the distance bin
$\lambda_n(m_{mb})$	annual frequency of EQ occurrence of magnitude m_{mb} on n^{th} seismic source.
$P_n(R = r_{mb} \| m_{mb})$	probability of occurrence of an EQ having magnitude m_{mb} on n^{th} source at a distance r_{mb} away from the site of interest.
$P(Z_{pga} > PGA \| m_{mb} r_{mb})$	Probability of exceedance of PGA for an EQ of magnitude m_{mb} and located at a distance r_{mb} away the site of interest.
Z_{pga}	Value of PGA obtainded by utilising GMPE for magnitude m_{mb} and epicentral distance r_{mb}..
$U_{ij}(f)$	Acceleration spectrum of the shear-wave portion of the ground motion of j^{th} sub-fault at i^{th} recording station
Q_s	Quality factor
Gr	Geometric spreading
f_0	Corner frequency
F	Free surface amplification

R^{θ}	Radiation pattern
ρ	Crustal density
β_s	Crustal shear wave velocity
V	Term that deals with the partitioning of energy in orthogonal directions
H_{ij}	Scaling factor responsible for converging the energy at the high-frequency spectral level of the sub-fault
$\Delta\sigma$	Stress drop
κ_H	Kappa
S_a	Spectral acceleration
R	Hypocentral distance
Amp	Amplification factor
$SA_{surface}$	Surface level ground motion level
SA_{rock}	Bedrock-level acceleration
A', B'	Site coefficients
G	Shear modulus
β	Damping ratio
$\lambda(M, st)$	Seismicity rate density
$K(M, r_I)$	Smoothing Kernel function
PL	Kernel fractal scaling index
δ	Degree of anisotropy
$\emptyset$	Angle subtended at the site of interest between the intersection of the fault plane with the Earth's surface and the epicentre location
$H(M)$	Bandwidth function
$\lambda_o(m_i, st)$	The annual rate of occurrence of a particular magnitude m_i
$P(Z > z \mid m_i r_{st})$	Probability of exceedance of ground motion level z for an EQ of magnitude m_i and located at the grid source *st*

List of Abbreviations

EQ	Earthquake
BIS	Bureau of Indian Standard
BP	Before Present
BSSC	Building Seismic Safety Council
CMF	Churachandpur–Mao Fault
CoV	Coefficient of variation
CUVI	Visual Cumulative
DBC	Density-based clustering
DBE	design basis earthquake
DBSCAN	Density-based spatial clustering of applications with noise
DSHA	Deterministic Seismic Hazard Analysis
DSPC	dynamic soil property curves
ELGRA	Equivalent Linear Ground Response Analyses
EMS	European macroseismic scale
EQ	Earthquake
EXSIM	Extended Finite-Fault Simulation
FEMA	Federal Emergency Management Agency
FMD	frequency-magnitude distribution
GCMT	Global Centroid Moment Tensor
GMDA	Guwahati Metropolitan Development Authority
GMM	Gaussian mixture model
GMPE	Ground Motion Prediction Equation
GMT	Generic Mapping Tools
GR	Gutenberg-Richter
GRA	Ground Response Analysis
GSI	Geological Survey of India
HFT	Himalayan Frontal Thrust
IMD	Indian Meteorological Department
IS	Indian Standard
ISC	International Seismological Centre
ISC	International Seismological Summary
IST	Indus Suture Thrust
LHS	Latin hypercube sampling
LLH	log-likelihood
LSM	Least square method
MASW	Multichannel Analysis of Surface Waves
mb	Body wave magnitude
MBT	Main Boundary Thrust
Mc	Magnitude of completeness
MCE	Maximum Credible Earthquake
MCT	Main Central Thrust

ML	Local Magnitude
MLM	maximum likelihood method
MMD	Mean minimum distance
Ms	Surface wave magnitude
MSK	Medvedev–Sponheuer–Karnik
Mw	Moment Magnitude
N	standard penetration resistance value
NDMA	National Disaster Management Authority
NE	North-East
NEHRP	National Earthquake Hazard Reduction Program
NEIC	National Earthquake Information Center
NRSC	National Remote Sensing Centre
PGA	Peak Ground Acceleration
PHA	Peak Horizontal Acceleration
PI	Plasticity Index
PSHA	Probabilistic Seismic Hazard Analysis
SC	Site Class
SHA	Seismic Hazard Analysis
SM	Source Model
SOM	Self-organizing maps
SRTM	Shuttle Radar Topography Mission
UHS	Uniform Hazard Spectra
USGS	United States Geological Survey
USGS	United States Geological Survey
USNRC	United States Nuclear Regulatory Commission

Chapter 1. Introduction

1.1 General

Earthquakes (EQs) are one of the most dangerous natural hazards as they happen unexpectedly and can cause instantaneous devastation to society. Apart from wreaking havoc on the physical world, it also affects people's psychology and instilling terror. Since the dawn of human civilisation, EQs have been occurring in various locations around the globe, some regularly, others infrequently. The seismic activity is much higher near plate boundaries than in the intraplate zone. For measuring the strength of an EQ, there are many magnitude scales, such as local magnitude (ML), body wave magnitude (mb), surface wave magnitude (Ms), moment magnitude (Mw) etc., available. In all these magnitudes, the higher the magnitude, the higher the energy released by the EQ will be, and the higher the destruction potential will be. When one examines human history, it becomes clear that many devastating EQs have occurred till now. As per Guha-Sapir and Vos (2011), some of the most destructive EQs (between 1970 and 2008) in terms of human impact include the 1976 Tangshan EQ of 7.8Mw magnitude (more than 242,000 people died), the 2004 Sumatra EQ of magnitude 9Mw (more than 226,000 people died), the 2008 Sichuan EQ of 7.9Mw magnitude (more than 88,000 people died), and the 2005 Kashmir EQ of 7.6Mw magnitude (more than 75,000 died). Recently (in 2023), in Turkey, a 7.8Mw magnitude EQ has occurred, resulting in more than 50,000 casualties (USGS). Out of all the past EQs, the EQ with the highest magnitude occurred in 1960 at Valdivia, with a magnitude of 9.5Mw (Kanamori and Anderson, 1975).

In India also, some areas have experienced very devastating EQs in the past, while others have experienced minor to moderate EQ-induced damages. This can be understood from the fact that the seismic activity varies significantly across the country. As a result, India is currently divided into four seismic zones based on the potential for seismic hazard posed by EQs, as defined in IS-1893 part 1 (2016). India's Himalayan and Northeast (NE) regions are the most seismically active. Within 50 years, the regions have suffered four great EQs (Jain 1998), namely, the 1897 Assam EQ (Mw 8.1) (Oldham 1899), the 1905 Kangra EQ (Ms-8, Mw-7.35) (Ambraseys and Bilham, 2000), the 1934 Bihar-Nepal EQ (ML-8.4, Mw-8), and the 1950 Assam-Tibet (Mw-8.4). On 16th June 1819, an EQ of 8.3 magnitude struck the western coast of India, known as the Kutch EQ. Calcutta was also affected by this EQ's ground-shaking. The event resulted in the formation of a fault scarp termed "Allah Bund", measuring around 10 feet in height and approximately 16 miles in length (Jain 1998). On 4th April 1905, the Kangra EQ occurred, killing around 20,000 people within the epicentre and destroying 100,000 structures (Ambraseys and Bilham, 2000). On 15th January 1934, the Bihar-Nepal EQ took place. This EQ resulted in widespread liquefaction, resulting in catastrophic damage in an area with a mean radius of around 300 kilometres. At other locations, six-foot-high embankments became level with the surrounding area, and the depth of some borrow areas, ponds, lakes, and other depressions decreased (Jain 1998). The Indian subcontinent has also witnessed several big EQs in locations not expected to experience EQ (following earlier versions of seismic zonation map of India), resulting in severe loss of life and property. The 1993 Latur EQ and the 2001 Bhuj EQ are examples of such EQs. On 29th September 1993, the Latur EQ occurred. It was 6.2 Mw in size and killed about 11,000 people (Gupta et al. 1998). On 26th January 2001, the Bhuj EQ occurred, killing around 13,805 people (Jain and Pathak, 2012).

1.2 Past Seismic Activity in NE India

As mentioned in the previous section, NE India is the most seismically active region in India and the world. Located in the eastern part of India, NE India is comprised of eight states. Except for Sikkim, all seven other states of NE India fall in seismic zone V, the highest seismic potential zone as per IS-1893 part 1 (2016). The NE India region is located approximately from 88°E to 97.4°E longitude and 21.9°N to 29.46°N latitude. The region is experiencing repetitive EQs at frequent intervals, including 1869 Cachar EQ (Mw-7.6, Highest EMS intensity- IX), 1897 Assam EQ (Mw-8.1, highest MSK intensity-IX, EMS intensity-X), 1923 Meghalaya EQ (Ms-7.1, Mw-6.8), 1930 Dhubri EQ (Mw-7.1), 1943 Assam EQ (Ms-7.2, Mw-6.9), 1947 Arunachal Pradesh EQ (Ms-7.7, Mw-7.2), 1950 Assam EQ (Mw-8.4), 1988 Manipur EQ (Mw-7.2), 2009 Assam EQ (Mw-5.1), 2011 Sikkim EQ (Mw-6.9) (Oldham 1899; Fahmi et al. 1988; Sabri 2002; Ambraseys and Bilham 2003; Ambraseys and Douglas 2004; Kayal 2008; Raghukanth et al. 2011; Baro and Kumar 2015; England and Bilham 2015). As a result of alarming seismic activity, the region is also experiencing various forms of induced damage such as ground shaking, landslides, liquefaction and building damage (Oldham, 1899; Kayal, 2008;

Raghukanth et al., 2011; Baro and Kumar, 2015). During the two great EQs, 1897 Assam EQ (Mw-8.1, highest MSK intensity-IX, EMS intensity-X) and 1950 Assam EQ (Mw=8.4), damages extended for more considerable distances. On 28th April 2021, an EQ of 6Mw (as per USGS) occurred in Dhekiyajuli, Assam, causing significant liquefaction. Before this EQ, another EQ (Mw-6.7) occurred near Imphal (in 2016), causing considerable damages (Gahalaut and Kundu 2016).

1.3 Earthquake Features

The earlier sections highlight that numerous features were generated during past EQs (particularly major and great EQs). These EQ-associated features also preserved the characteristics of the EQs. The features generated due to EQs are termed Seismites (Vittori et al. 1991). Sand dikes, sand blows, Fault scarps etc., are some examples of seismites. The seismites are divided into different categories based on genesis (or origin), location and timing (McCalpin and Nelson 2009). Depending on their origin, these seismites may be Primary, which results from tectonic deformation, or Secondary, which results from EQ shaking. Based on their positions, these Seismites can be classified as on-fault or near-field features and off-fault or far-field features. In addition, based on their timing of occurrence, these Seismites can be coseismic (origin during the time of EQ) or postseismic (originated after EQ shaking). Based on their preservations, seismites can also be classified as geomorphic and stratigraphic evidence. Finally, combinations of the preceding types can result in 16 distinct categories (following McCalpin and Nelson 2009). The studies and investigations of these different types of seismites are beneficial in understanding the past EQs in terms of their origin, size and potential effects. The following section briefly describes these types of studies and investigations associated with the seismites.

1.4 Paleoseismological Study

As the seismites preserve the characteristics of the EQs, studying them makes it possible to know more about the past EQs during which such seismites were formed. This study of collecting information on past EQs by using trench surveys of the seismites is known as paleoseismology. Paleoseismological investigations have grown in popularity among seismologists. Seismic activity or hazard of an area is found based on the past EQ information and available tectonic information of the area. In this case, paleoseismology aids in understanding the long-term active seismic features that are most important for hazard assessments, and the seismic history of a fault may then be accessed (McCalpin and Nelson

1996). The paleoseismological study has three main objectives: finding the features, i.e. the Seismite, determining the time of origin of those Seismites, and determining the magnitude of the EQ that produced those Seismites. Besides these objectives, paleo seismological study also attempts to determine the potential focal mechanisms of the EQ. As a result, the outcome of the paleoseismological study is beneficial for the determination of the seismic hazard level of a region.

1.5 The Motivation for the Present Work

Considering the expanding infrastructure and the damage situation witnessed during previous EQs in NE India, accurate seismic hazard analysis (SHA) is critical. Proper assessment of seismic hazards at the regional/ site-specific level not only aids in building design but also plays a key role in reducing the loss of human lives and infrastructure during EQs. For assessing seismic hazards, understanding the seismic activity of an area is very important. To accurately determine the seismic activity of an area, precise knowledge of past EQ in terms of magnitude and its recurrence interval is very important. As per Jacoby (1997), for most places of the world, the actual record of EQs is concise and incomplete. This incomplete knowledge of past EQs will be the reason for underestimating seismic activity and, subsequently, the seismic hazard values. With a limited EQ catalogue, it is almost impossible to determine the return period of major to great EQs as these EQs have a longer return period, but the damage induced by them is enormous. For these reasons, it is essential to include information about past unrecorded EQs, obtained from paleoseismological studies in seismic activity determination. It should be noted here, however, that the information obtained from paleoseismological studies contains huge uncertainties regarding the EQs' location, magnitude and time. Though many past studies considered some paleo EQ events in the SHA (mentioned in Chapter 2), the studies neglected the associated uncertainties. Hence, the present study mainly aims to perform SHA with an updated EQ catalogue consisting of instrumental and paleo EQs both for the NE India region.

1.6 The Organisation of the Thesis

This thesis is presented in eight chapters. A brief overview of each of these chapters is given below:

- **Chapter 1 Introduction:** This chapter provides basic information about past EQ occurrences, damages and features. The past seismic activity in NE India is also briefly described in this chapter. The chapter highlights the motivation of the present work.

- **Chapter 2 Literature Review and Research Gaps:** This chapter deals with the literature review conducted to understand the research area and its seismotectics better. The chapter provides insight into SHA methodology, previous SHA work and paleoseismological studies performed in the study area. Further, based on the literature review, the research gaps are identified based on which the objectives of the present thesis are defined.
- **Chapter 3 EQ Data Collection and Data Processing for the Study Area**: This chapter explains the type of data collected for the study area. The chapter also discusses the sources of data and the basic interpretation of the data. Further, the methodology used for the required data processing is also discussed in the chapter.
- **Chapter 4 Seismic source zonation for NE India:** Seismic source zonation for NE India based on past EQ location and distribution of seismicity parameters is attempted in this chapter. The chapter also describes the methodology used for the zonation of seismic sources for NE India. Further, the characteristics of identified source seismic zones (based on the current study) are described.
- **Chapter 5 Probabilistic Seismic Hazard Analysis for NE India:** The chapter initially explains the methodologies used for the PSHA and the uncertainty associated with the methodologies and models used for the PSHA of NE India. This chapter then deals with the PSHA for NE India. Further, the chapter contains the deaggregation of EQ hazards carried out for eight important cities in NE India.
- **Chapter 6 Seismic Hazard Analysis (SHA) of NE India by Utilising Kernel Density Method to Deal with the Uncertainty in Location and Magnitude of Past EQs:** The chapter deals with the kernel density-based PSHA. The chapter explains the methodology and describes the entire process for the estimation of seismic hazard levels for NE India.
- **Chapter 7 Development of GMPE for NE India, applicable to different site classes:** This chapter presents the stochastic modelling of the ground motion records for the study region. Further, the chapter describes the methodology adopted for the development of the GMPE for NE India.
- **Chapter 8 Conclusions:** The overall summary and conclusions of the work are presented in this chapter. Furthermore, the future scope of the present study is highlighted.

Chapter 2. Literature Review and Research Gaps

2.1 Background

This chapter provides an overview of the seismotectonics of the study area. In addition, the chapter discusses existing SHA methodologies and previous seismic hazard studies conducted either for the parts or the entire NE India. Identifying historic EQs in the region is very important to understand true seismicity. Such information, however, is not readily available and requires detailed in-situ investigations (such as paleoseismological studies). Accordingly, a discussion on paleoseismological studies around the world and for the present study region is presented in this chapter, clearly explaining how seismologists carry out paleoseismological studies and the outputs obtained from such studies. In addition, this chapter also highlights the research gaps and presents the objective of the thesis.

2.2 Study Area

As the name suggests, the NE India region is located in the northeastern side of India, sharing borders with Nepal, Bhutan, China, Myanmar and Bangladesh. Geographically, the region lies approximately between 88°E to 97.4°E longitudes and 21.9°N to 29.46°N latitude (refer to Figure 2.1). The Himalayas surround the region in the north and east directions and have rugged terrain. The region shows diverse geomorphology. Both hilly areas, such as Eastern Himalayan Region, Mikir Hill, Shillong Plateau, Naga Thrust, Mishmi Hill and Indo-Burma ranges, and plain areas, such as the Assam Valley region, are present in this region. Two major rivers, the Brahmaputra and the Barak, flow through the region.

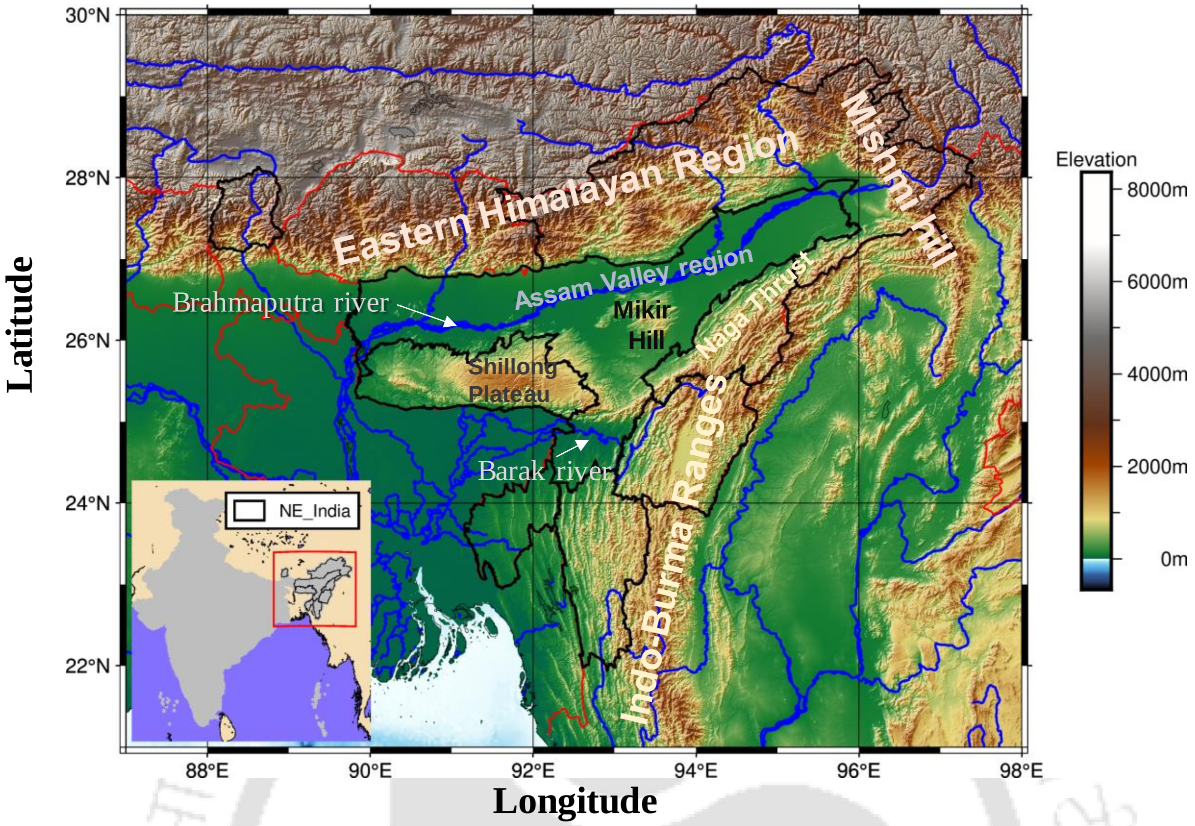

Figure 2.1 NE India map showing hills, valleys and major river systems developed with the help of Generic Mapping Tools (GMT) by using Shuttle Radar Topographic Mission (SRTM) data

2.2.1 Geomorphology of NE India

Figure 2.2 shows the geomorphology of NE India collected from Bhuvan and Bhukosh (GSI and NRSC 2012). Throughout the region, structural origin, denudational origin, fluvial origin and glacial origin landforms are visible. Moreover, structural origin landforms include highly, moderately, and low dissected hills and valleys, highly, moderately, and low dissected upper plateau, and highly, moderately, and low dissected lower plateau. In the case of denudational origin, the region has highly, moderately and low dissected hills and valleys, moderately dissected upper plateau, and highly, moderately, and low dissected lower plateau, piedmont slope, and pediment pediplain complex (GSI and NRSC 2012). Fluvial-origin landforms are mainly present in the states of Assam, Arunachal Pradesh, Tripura and Manipur. In fluvial origin landforms, the region has older alluvial plains, younger alluvial plains, older flood plains, active flood plains, and piedmont alluvial plains. Glacial origin landforms, glacial terrain, and snow cover are mainly in Sikkim and Arunachal.

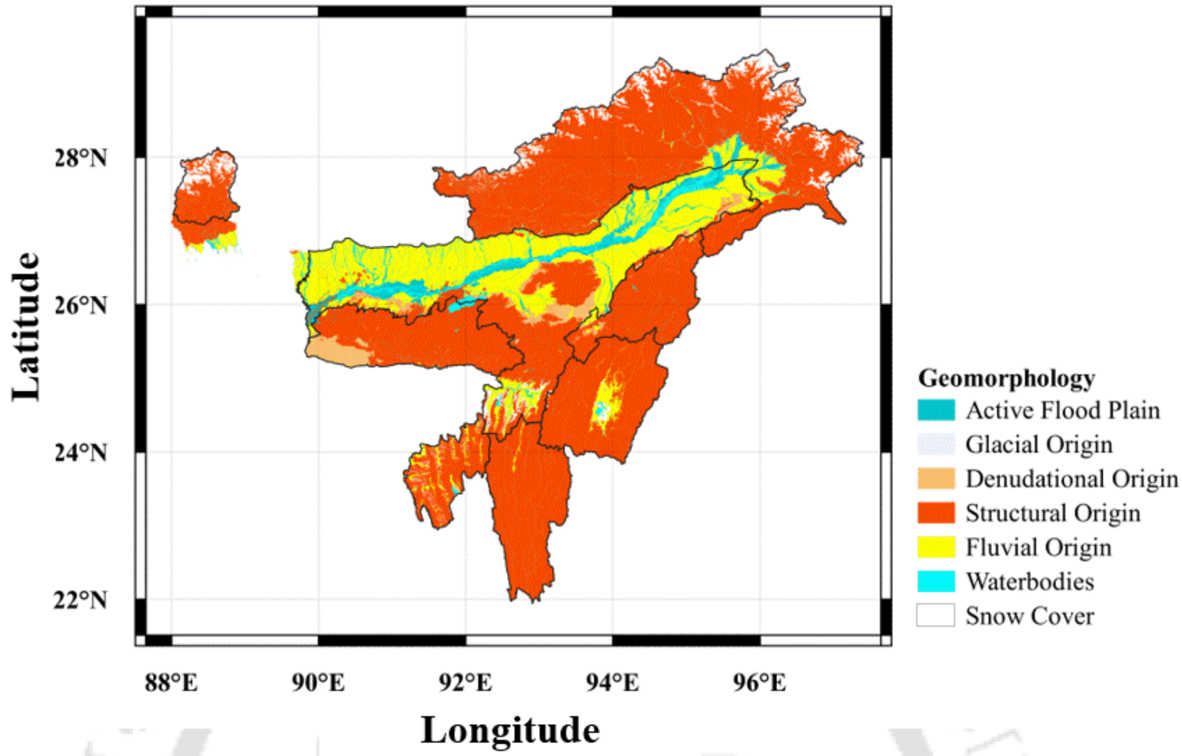

Figure 2.2 Geomorphological map of NE India developed based on the data obtained from Bhuvan and Bhukosh (GSI and NRSC 2012)

2.2.2 Seismotectonics of NE India

The high seismic activity of the NE India region is due to numerous active faults, which are generated due to the collision of the Indian plate with the Eurasian plate in the north and with the Burmese plate in the east. These two collision zones create a complex geotectonic setting for the NE region. As per Bilham (2004), the above subduction process is still continuing in the region at both fronts, thus making the region highly active in terms of seismicity. Considering its diverse and complex nature, past researchers have divided the region into various seismotectonic zones (Kayal 1998; Sharma and Malik 2006; Sil et al. 2013; Das et al. 2016; Baro et al. 2018a). These zones have different characteristics regarding tectonic movements, direction, speed and fault mechanisms (Kayal 1998; Angelier and Baruah 2009). The main seismotectonic zones (Figure 2.3) present in the area include the Eastern Himalayan Region, the Indo–Burma ranges, the Shillong–Mikir Plateaus, the Assam Valley, and the Bengal Basin region (Angelier and Baruah 2009).

In the Eastern Himalayan region, due to the collision of the Indian plate with the Eurasian plate in the north direction, lots of thrusts have been originated (Kayal 2008). The Indus Suture Thrust (IST), the Main Central Thrust (MCT), the Main Boundary Thrust (MBT) and the

Himalayan Frontal Thrust (HFT) are the main thrusts present in the region (Kayal 2008; Thingbaijam et al. 2008). These faults of the Himalayan region are mainly active on the western side of the region. For the northeastern side, the above seismic sources are not that active (Mittal et al. 2012). 1991 Uttarkashi EQ (Mw-6.8), 1999 Chamoli EQ (Mw-6.6), and 1980 Gangtok EQ (Mw-6.3) occurred on the western segment of the MCT (Mukhopadhyay 2011). The slip on HFT generated 1905 Kangra EQ (Ms-7.8, Mw-7.2), 1934 Bihar-Nepal EQ (Mw-8.2) and 1950 Assam EQ (Mw-8.4) (Pandey and Molnar 1988; Ambraseys and Bilham 2000; Bilham 2001). NE Himalayan region shows diversity in seismic activity distribution and major structural units within the region, thus, can further be sub-divided into three subzones: the Bhutan Himalaya, the Arunachal Himalaya and the Mishmi Thrust (Angelier and Baruah 2009). As per Angelier and Baruah (2009), the source mechanisms of the region show normal, reverse and strike-slip faulting. Besides these, Angelier and Baruah (2009) concluded the presence of three types of seismotectonic stresses within this zone namely: N–S compression, E–W extension, and N–S extension.

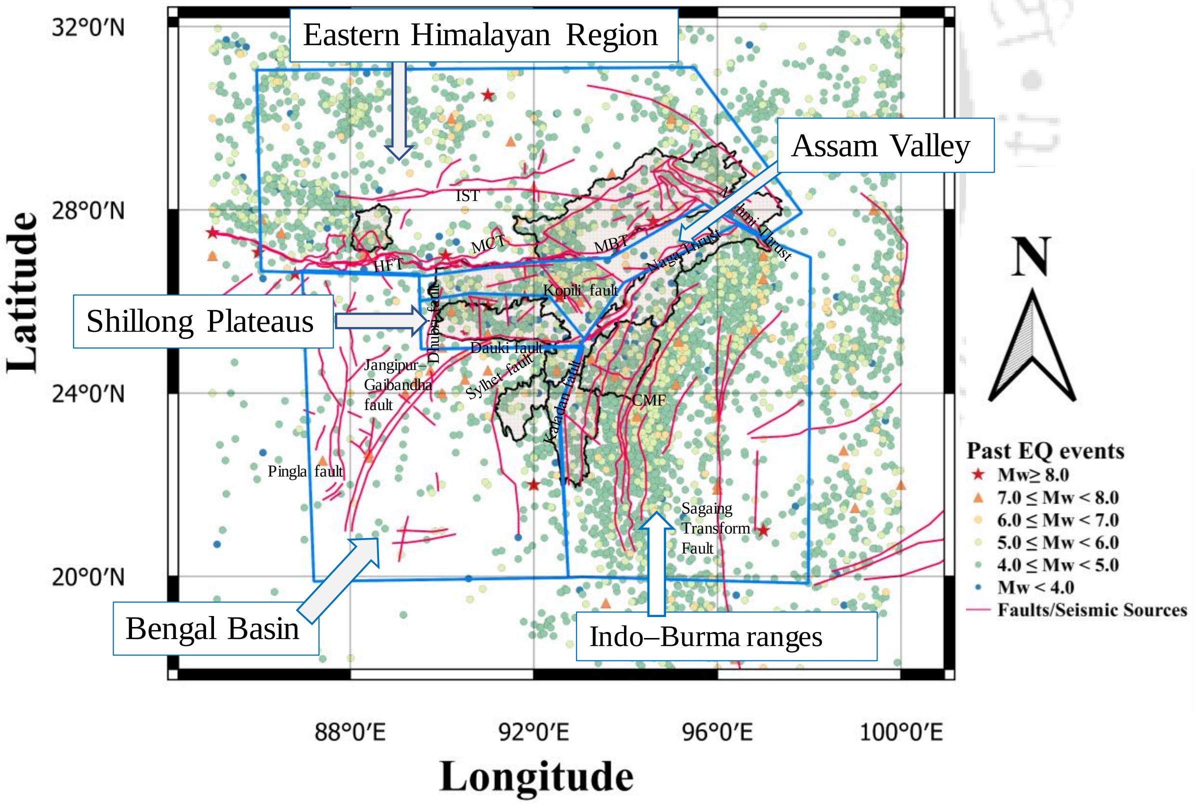

Figure 2.3 Seismotectonic map of NE India showing different zones.

Towards the eastern side of NE India, the collision of the Indian and Burmese plates gave rise to the Indo-Burma Ranges. EQs in this range can be classified as plate-boundary EQs, Benioff-zone EQs and overriding-plate EQs. The Benioff zone EQs with focal depth < 90 km show normal and strike-slip faulting, while EQs with focal depth > 90 km show reversed faulting (Kayal 1998). On the other hand, shallower overriding-plate EQs show both thrust faulting and strike-slip faulting. Based on the fault mechanism of EQs associated with the Sagaing transform fault (present in the Indo-Burma Ranges), it is considered a strike-slip fault. Trending in the northeast-southwest direction in the north and northwest-southeast direction towards the south, the Indo-Burma Ranges has many thrust faults such as the Kaladan fault, the Naga Thrust and the Eastern boundary Thrust (Kayal 2008). The Kaladan fault behaves as a strike-slip fault in the north but towards the south as a thrust fault (Maurin and Rangin 2009). The Naga thrust (according to Wang et al. 2014) is an active 400 km-long thrust fault. The Eastern boundary thrust, known as the Churachandpur–Mao Fault (CMF), is present to the east of the Naga thrust. In their study, Kundu and Gahalaut (2013) found that this CMF was a thrust fault, but now it shows strike-slip fault behaviour. Besides these faults, the Kabaw fault is present on the east side of the CMF. Based on the recent seismicity data, EQ epicentre's location and associated focal mechanism solution, it is found that the Kabaw fault is an oblique reverse fault with a right lateral strike-slip component (Aung 2020). As per Angelier and Baruah (2009), the distribution of stress states is very complex in the Indo-Burma ranges, with four stress states: N–S compression, E–W compression, E–W extension and NE–SW compression.

The Bengal Basin region, which is located far from plate boundaries, is a seismically less active region (Kayal 2008). However, there are some faults present in the region that show some seismic activity. In this region, in the year 1918, Srimangal EQ (Ms-7.6, Mw-6.1) occurred on the Sylhet fault. This Sylhet fault is a strike-slip fault trending in a northeast-southwest direction. Another EQ, 1935 Pabna EQ (M_W-6.2), occurred in the northeast-southwest oriented Eocene Hinge Zone (Kayal 2008; Nath et al. 2014). Besides these two faults, the Jangipur–Gaibandha fault, the Pingla fault, the Garhmayna–Khanda Ghosh Fault and the Sainthia–Bahmani fault are also present in the Bengal Basin region. Angelier and Baruah (2009) observed NE–SW compression and N–S compression stress states in the Tripura belt, located east of the region.

The Shillong Plateau region is located towards the north of the Bengal Basin region, which experienced an uplift during the Late Cenozoic (Rao and Kumar 1997). The northern edge of the Shillong Plateau again uplifted during the 1897 Shillong EQ (Oldham 1899; Bilham and England 2001; Vernant et al. 2014). Analysing the GPS data, Vernant et al. (2014) discovered that the Shillong Block rotates clockwise at 1.15°/Myr rate between 89° and 93°E.

Numerous active faults exist in the region, such as the Dauki fault, the Brahmaputra fault, the Dapsi Thrust, the Barapani Shear Zone, and the Oldham fault. The Dauki Fault, an active north-dipping reverse fault (Morino et al. 2011), separates the Shillong Plateau and the Bengal Basin region. The Brahmaputra fault is in the north of the Shillong Plateau (Dasgupta and Nandy 1982). The Oldham fault triggered the great 1897 Assam EQ (Mw-8.1), while the Dhubri fault was the source of the 1930 Dhubri EQ (Ms-7.1, Mw-6.8) (Bilham and England 2001; Nandy 2001; Kayal 2008). A study by Angelier and Baruah (2009) showed that reversed and strike-slip faults are more common in the region, while EQs with normal faulting are rare. In their work, Angelier and Baruah (2009) also discovered that the area has two forms of seismic stress: N–S compression and E–W compression.

Like the Shillong block, the Assam block also rotates at a 1.13°/Myr rate from 93.5°E to 97°E (Vernant et al. 2014). Seismically, the Assam Valley area, which includes this Assam block, is less active than other regions. The Kopili fault, which was responsible for the 1869 Cachar EQ (Mw-7.5) and the 1943 Assam EQ (Ms-7.2, Mw-6.9), is a critical seismic source in the region (Kayal et al. 2010). As per Angelier and Baruah (2009), the N–S and E–W compression stress states are visible in this region.

In general, the entire NE India is seismically very active and has most of the faults which were responsible for major to great EQs in the past and present (Bilham and England 2001; Nandy 2001; Kayal 2008; Nath et al. 2014). However, some faults segments, such as the NE Himalayan segment, have not been active for some time. As per Mittal et al. (2012), a seismic gap called the ‘Northeast Seismic Gap’ exists between the epicentres of the 1934 Bihar-Nepal EQ and the 1950 Assam EQ. However, Srivastava et al. (2015) considered this seismic gap region consisting of two different gaps: the Sikkim–Bhutan seismic gap and the Arunachal seismic gap. According to Srivastava et al. (2015), the Sikkim–Bhutan seismic gap has the potential to generate 7.5 magnitudes EQ, while the Arunachal seismic gap has the potential to generate 8.5 magnitudes EQ in future. Further, as per Khattri et al. (1983), one more seismic gap exists in the upper Assam Valley region, which is termed the ‘Assam Gap’. The ‘Assam gap’ is not showing any seismic activity for a long time. As a result, this area is considered aseismic and not a seismic gap (Khattri and Wyss 1978; Khattri et al. 1983; Khattri 1987). However, there is some risk in declaring any region of the continent to be an aseismic zone (Kayal, 1998). The sorts and extents of damage from the Dhekiyajuli EQ in 2021 demonstrate the possibility of these seismic gaps, and their presence should not be overlooked.

2.3 Site Classes in NE India

As per the National Earthquake Hazard Reduction Program (NEHRP), a site can be classified into any one of the six site classes (SCs) based on the shear wave velocity (Vs), standard penetration resistance value (N), undrained shear strength (Su), and the type and properties of soil present in the profile (as per BSSC (2003) also known as FEMA 450). These SCs are; SC A, B, C, D, E, and F. Table 2.1 shows the various SCs, as per NEHRP guidelines.

Table 2.1 NEHRP site classification as per FEMA 450

Site class	Detail	Average properties in the top 30m		
		V_{s30} (m/s)	SPT N	Undrain shear strength, S_u (kPa)
A	Hard rock	>1500	-	-
B	Rock	760-1500	-	-
C	Soft rock and Very dense soil	360-760	>50	>100kPa
D	Stiff soil	180-360	15-50	50-100kPa
E	Soil	<180	<15	<50kPa
		Any profile with more than 3m of soft clay defined as soil with PI>20, w≥40%, and S_u<25kPa		
F	This type of soil requires site-specific evaluation	Soils vulnerable to potential failure or collapse under seismic loadings such as liquefiable soils, quick and highly sensitive clays, and collapsible weakly cemented soils		
		Peats and/or highly organic clays (H > 10 ft [3 m] of peat and/or highly organic clay where H = thickness of soil)		
		Very high plasticity clays (H > 25 ft [8 m] with PI > 75)		
		Very thicksoft/medium stiff clays (H > 120 ft [36 m])		

Field investigation-based SCs determination throughout all NE India locations has not been completed until now. However, different researchers such as Nath (2007); Jana et al. (2012); Sil and Sitharam (2013); Sitharam and Sil (2014); Haloi and Sil (2015); Pallav et al. (2015); Biswas and Baruah (2017); Kumar et al. (2018); Kumar and Suman (2020) used geotechnical, geophysical tests for the determination of SCs in the region. These studies are mainly available for some major cities in NE India. Nath and Thingbaijam (2011a) concluded that the locations in Guwahati, Assam fell under SCs CD (intermediate SC between C and D), D, E, and F using the SPT-N value and subsequent data analysis. Apart from this study, Kumar et al. (2018) used SPT N30 and V_{s30} values from the MASW test to determine the SCs in Guwahati. In the study, Kumar et al. (2018) found that both methods gave different SCs for some locations in Guwahati. Additionally, Kumar et al. (2018) found that significant city areas fell under SCs D and E when SPT N30 was considered. Further, when V_{s30} from MASW was considered, significant city areas fell under SCs C and D.

Jana et al. (2012) conducted site response research in the city of Jorhat based on geotechnical investigation and reported V_{s30} values ranging from 222 to 339 m/s for sites with varying depositional conditions of sand, silt, clay, and sandy clay. Sil and Sitharam (2013) conducted MASW survey and geotechnical investigation to examine the distribution of shear-wave velocity in Agartala and Aizwal. In the study, it was found that the region has SC D and E conditions. Pallav et al. (2015), based on borehole data, reported the presence of SCs D and E in Imphal City. In their study, Haloi and Sil (2015) conducted site characterisation in Silchar and reported a V_{s30} in the range of 180–360 m/s.

In NE India, PESMOS has many ground motion recording stations spread across the region. These stations are mainly located on SC B, C and D. Besides, the topography-based approximate V_{s30} map for the entire globe is also available in the United States Geological Survey (USGS) database (https://earthquake.usgs.gov/data/vs30/) as the global V_{s30} map. This global V_{s30} map is developed based on the work done by Heath et al. (2020). This type of map, though it does not provide an exact V_{s30} value for a specific location, is very close to field-based results the majority of the time. Based on the spatial distribution of V_{s30} values for NE India (Figure 2.4) following the USGS global V_{s30} map, it is observed that soft soil profiles such as SCs E and D are mainly associated with the Brahmaputra and the Barak valleys (Figure 2.4). Similarly, relatively hard/ stiff soil profiles and rock surfaces such as SCs B and C are associated with the hilly regions of NE India.

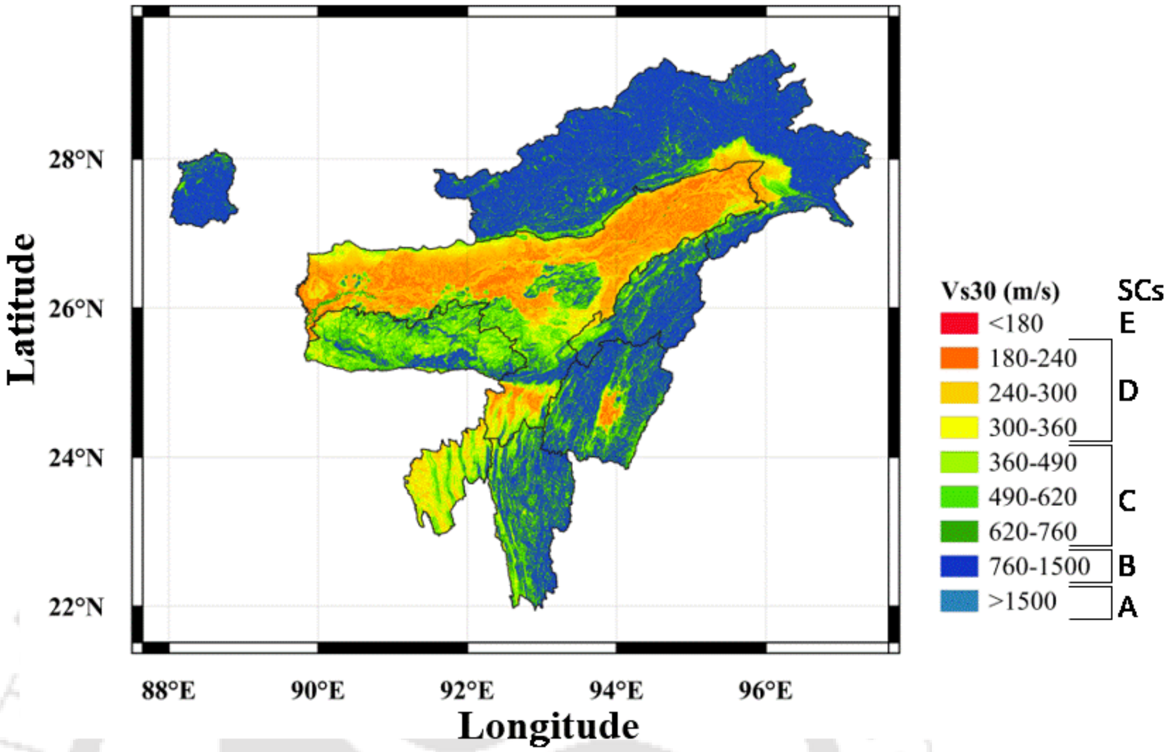

Figure 2.4 Spatial distribution of V_{s30} in NE India region developed based on USGS global V_{s30} (USGS https://earthquake.usgs.gov/data/vs30/)

2.4 Seismic Zonations as Per Indian Standard Seismic Codes

For the seismic safety of structures located in different locations in India, the Bureau of Indian Standards has published many codes. The main code followed for designing seismic-resistant structures is IS 1893 part 1 2016 (BIS 2016). This code is the sixth revision of the IS 1893 code adopted by BIS in 2016. As per IS 1893: part 1 2016, India consists of four seismic zones namely; Zones II, III, IV, and V (refer to Figure 2.5), having zone factors of 0.10, 0.16, 0.24 and 0.36, respectively. Except for Sikkim (Zone IV), all other states of NE India fall under Zone V, having a zone factor of 0.36.

It should be noted here that even though NE India is mostly covered in seismic zone V (single zone), past EQ locations, seismic activity, geology etc., show wide variation throughout the NE India region. Such variation can lead to considerable changes in seismic hazard values across NE India against the same seismic zone as per IS 1893(Part 1): 2016. Thus, several researchers have performed regional-level SHA of either the entire NE or its different parts. The next sections provide steps to be followed while performing SHA of a region/ site.

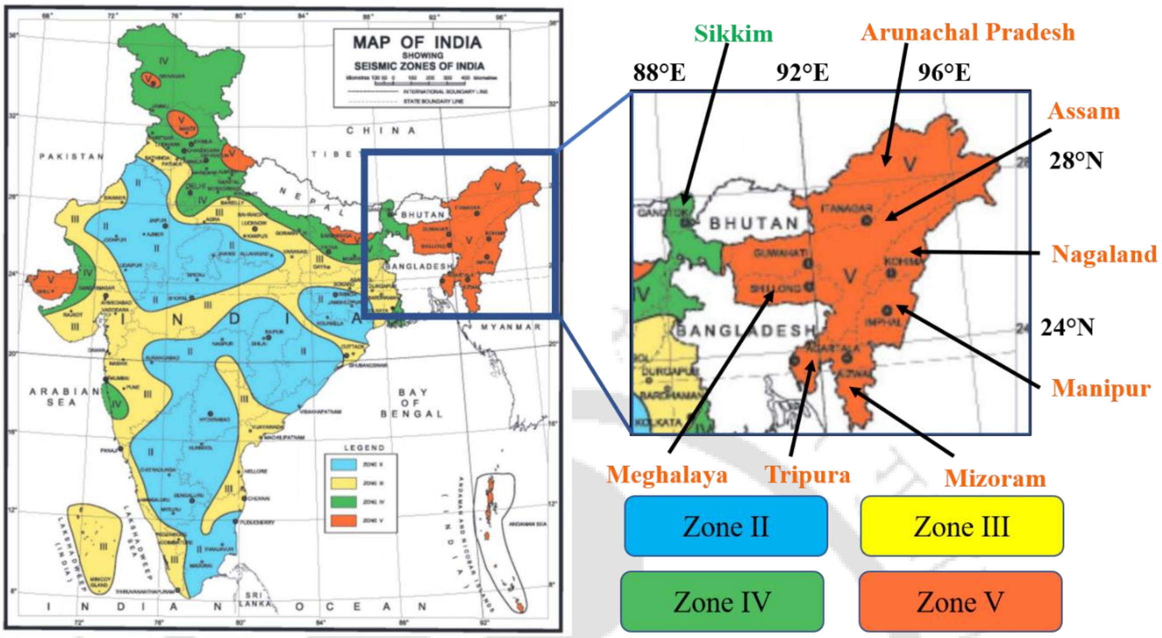

Figure 2.5 Seismic zonation map of India (modified after IS 1893 (part1) 2016). The NE India part is shown along with the names of the states.

2.5 Seismic Hazard Analysis

Based on the earlier discussion, it is clear that NE India is seismically highly active. Thus, consideration of safety of existing as well as proposed infrastructure against future EQ, is utmost important to quantify the seismic hazard in the region. The knowledge of the seismic hazard level for the region is beneficial in minimizing fatalities as well as economic losses for the future. For determining the seismic hazard level for the region, performing SHA is essential.

There are primarily two methods for SHA namely; Deterministic SHA (DSHA) and Probabilistic SHA (PSHA). In DSHA, the ground motion hazard estimation is done based on a particular seismic scenario. This scenario is mainly the occurrence of an EQ of a certain magnitude at a specific location (Kramer 1996). Mainly, in DSHA, the highest magnitude EQ, originated at the minimum distance between the site to the source is considered, which results in the maximum amplitude of ground shaking at the site of interest (worst-case scenario).

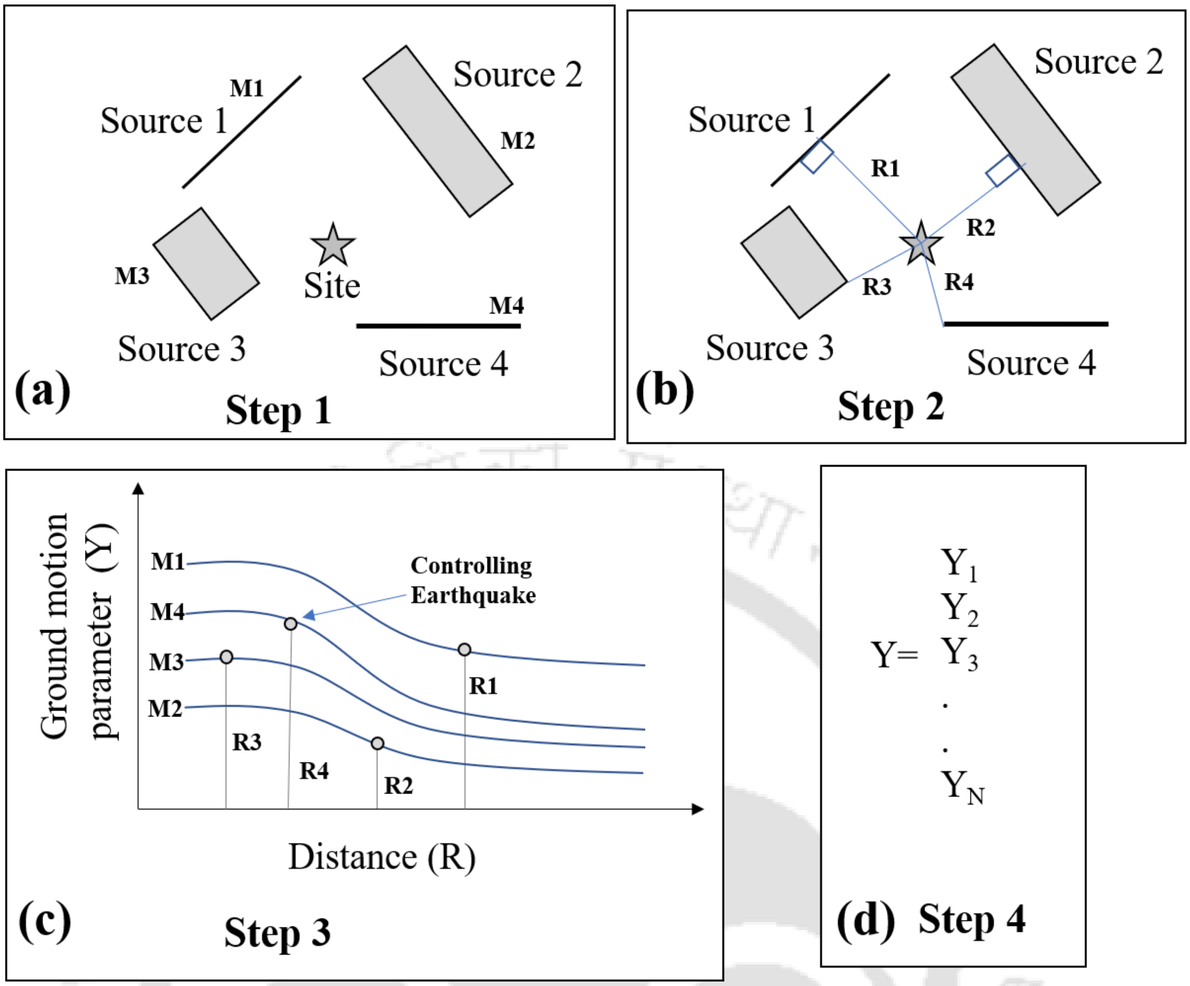

Figure 2.6 Steps in the DSHA (modified after Reiter, 1990)

The methodology of DSHA can be summarized into the following four steps (following Reiter, 1990). These steps are also explained in Figure 2.6:

1) Preparation of a seismotectonic map and identification of all the sources which are capable of producing EQ of significant magnitude around the study area.

2) Estimation of the shortest distance between each of the identified sources and the site of interest.

3) Selection of the maximum possible EQ that can produce the maximum level of shaking for each of the above-identified sources.

4) The overall maximum level of ground motion, which can be produced by comparing the maximum ground motion from each of the identified sources at the site of interest.

DSHA provides a straightforward, comprehensible framework for the determination of seismic hazards. The analyses in DSHA are transparent, discussable, and reviewable. In cases where new information becomes available after an analysis has been performed, determining and evaluating its impact on the results is relatively simple. There are, however, many drawbacks associated with DSHA. DSHA disregards the inherent uncertainty (related to EQ magnitude, frequency of occurrence, location and exposure period) in SHA. Additionally, the analysis ignores the frequency of EQ occurrence of a certain magnitude. Besides, choosing the maximum EQ is a contentious issue in DSHA that often leads to arguments.

As discussed earlier, DSHA does not account for the uncertainty with respect to EQ magnitude, frequency of occurrence, location and exposure period. It only provides the worst scenario of ground shaking irrespective of the design life of the utilities going to be constructed based on seismic hazard values. To account for all the above uncertainties, PSHA is a better choice. PSHA characterises uncertainty in location, size, frequency, and effects of EQs and combines all of them to compute probabilities of different levels of ground shaking in terms of hazard curves. PSHA predicts a relationship between a ground motion value, such as peak ground acceleration (PGA), and the frequency of exceedance of that ground motion to occur at the site of interest. PSHA includes the frequencies of exceedance of different levels of EQ ground motion that are calculated considering the presence of faults and seismotectonic provinces, soil properties around the site, and the EQ history of the fault/ region.

The PSHA study aims to quantify all uncertainties discussed above and combine them to clearly describe the distribution of future ground motion levels that may occur at a site. The uncertainties in the seismic hazard values and in the mathematical model of the hazard are represented by developing a family of hazard curves for each identified source. While developing hazard curves, the Logic tree approach may be used to take care of uncertainty associated with the ground motion prediction models used for calculating the ground motions. In the end, Uniform Hazard Spectra (UHS) can be calculated for a return period of interest, based on above developed hazard curves.

The PSHA methodology is based primarily on the work done by Cornell (1968), Esteva (1969, 1970), Algermissen et al. (1982) etc. Further, based on these works, programs to perform PSHA were developed by McGuire (1976) and Bender and Perkins (1987).

The steps followed in PSHA can be described (also in Figure 2.7) as follows (Reiter 1990):

1) Identification of EQ sources capable of producing significant EQs. Characterise each source in terms of magnitude-frequency and distance-probability.

2) Estimate the recurrence relation (Gutenberg-Richter relation) for each source. This will give the frequency of occurrence of various magnitude values for each source.

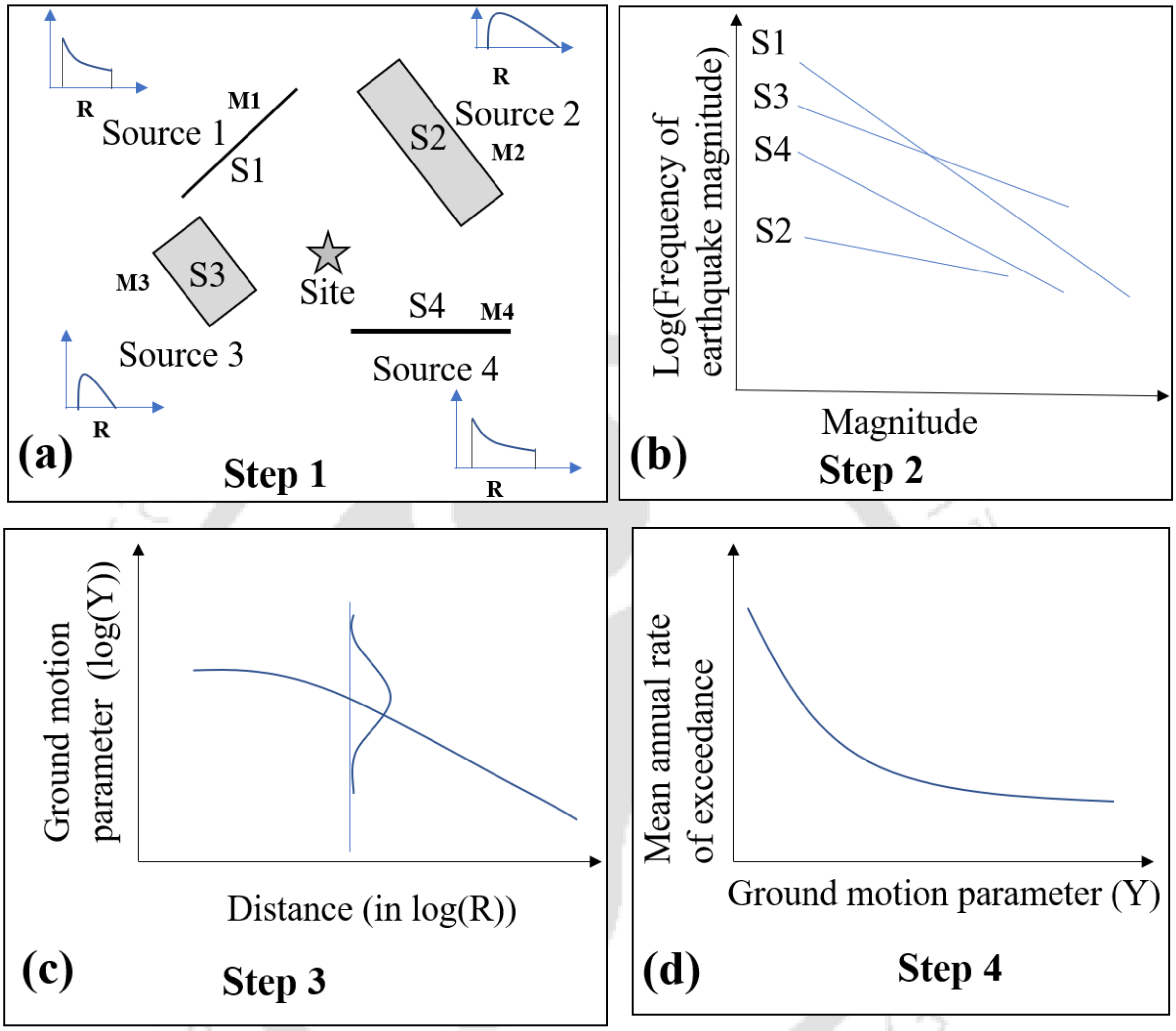

Figure 2.7 Steps in the PSHA (modified after Reiter, 1990)

3) Estimate the ground motion produced considering suitable ground motion prediction equation, using various magnitude and distance values obtained in step 1.

4) Estimate the frequency of exceedance for different levels of PGA from all the sources. This will provide the frequency of exceedance of different levels of ground motions.

Once information about the frequency of exceedance of different levels of ground motions is known in the form of the total hazard curve for the site and for the desired frequency of exceedance (representing Maximum Credible Earthquake/MCE, Design Basis Earthquake/DBE, Operating Basis Earthquake/OBE etc.), the expected level of ground motion can be determined.

2.6 Review on SHA for NE India by Various Researchers

For estimating seismic hazards in the various regions of India, many researchers performed both DSHA and PSHA. Khattri et al. (1984) performed PSHA for India and its adjoining regions to build a seismic hazard map for the Indian subcontinent. This map was developed based on PGA values corresponding to a return period of 475 years. Khattri et al. (1984) divided the whole area into 24 source zones based on past seismicity, tectonics etc. One such zone consisted of the Brahmaputra valley and Shillong massif. For shallow EQs and intermediate EQs, Khattri et al. (1984) considered two separate values of maximum magnitudes. For the first case, the value was taken as 8.75 and for the second case, the value was taken as 7. Master curves prepared by Perkins (1978) were used to estimate the seismic hazard of the areas. Master curves were the curves which show expected peak accelerations due to point sources of particular magnitude, for a particular distance within source zones. Khattri et al. (1984) on the basis of the study found that the Brahmaputra valley and the Quetta region showed highest peak acceleration of 0.8g.

Bhatia et al. (1999) prepared a seismic hazard map for the entire country and its adjoining regions. Bhatia et al. (1999) divided the whole area into 85 seismic zones. Out of which, 10 source zones covered the entire NE India. PGAs were estimated at 0.5°× 0.5° grid locations considering 10% probability of exceedance in 50 years by using the attenuation relation developed by Joyner and Boore (2011). For NE India, the Burmese arc and the Hindukush region showed higher values of PGA in the study (0.35g-0.4g). Though the hazard values obtained by Bhatia et al. (1999) were less than those of Khattri et al. (1984), both studies showed high seismic hazard for NE India.

Sharma and Malik (2006) performed PSHA for NE India, considering four seismic zones for the entire region. Further, these four zones are again divided into 10 seismogenic zones. Sharma and Malik (2006) used the EQ catalogue provided by IMD, having events from the year 1762 to 2001, including the past two great EQs (1897 and 1950). With the help of this catalogue, the hazard parameters, maximum magnitude and return periods of different magnitudes for each zone were estimated. This data was used to estimate ground motion using the CRISIS program. The attenuation relations developed by Youngs et al. (1997), and Sharma and Bungum (2006) were used to develop the seismic hazard maps for 10% and 20% probability of exceedance in 50 years. These two attenuation relations were not developed for the study area but for area with similar tectonic settings. From the study, Sharma and Malik (2006) found that the range of the PGA values is from 0.05g to 0.6g for 10% probability of exceedance in 50 years and from 0.01g to 0.4g for 20% exceedance in 50 years.

Raghu Kanth et al. (2009) performed DSHA for Imphal city in Manipur. The finite-fault seismological model proposed by Motazedian and Atkinson (2005) was used to estimate the time history acceleration in the rock level. The time history was simulated for two EQs; 1869 Cachar EQ (Mw 7.5) and a hypothetical Mw 8.1 event in the Indo-Burma subduction zone. Based on the SPT-N 122 borehole data of the Imphal city; the Vs profile of the soil was estimated. With the help of this profile, nonlinear ground respond analysis was carried out to find at ground surface acceleration time history. For 1869 Cachar EQ, the surface level PGA value obtained for Imphal city was ranging from 0.14-0.24g, and for the Mw 8.1 hypothetical EQ in the Indo-Burmese Arc, the PGA value in Imphal was ranging from 0.25 g to 0.45 g.

The National Disaster Management Authority (NDMA) of India in 2010 published a report on PSHA for entire India. They considered a total of 38,860 EQ events having magnitude ≥ 4 Mw till 31st December 2008. This catalogue contained instrumental data, historical data and paleoearthquake data. Based on the past seismic history, geology and tectonic, the entire country was divided into 32 seismic source zones. Completeness analyses for the catalogue were carried out and based on the complete data, seismicity parameters were estimated for all the seismic source zones. NDMA (2010) used simulate ground motion samples obtained by Motazedian and Atkinson (2005) method to develop attenuation relation for the entire country having different sets of coefficients for different regions. Based on this attenuation relation seismic hazard maps were developed for the entire country considering different return periods.

Pallav et al. (2012) estimated PGA for different locations in the state of Manipur by performing PSHA. Pallav et al. (2012) divided the entire region into seven seismic source zones based on tectonic and geology. These regions were the Indo-Burmese Ranges, the Eastern Himalaya, the Mishmi Thrust, the Shillong Plateau, the Naga-Disang Thrust, the Bengal Basin, and the Assam Valley. The catalogued used by Pallav et al. (2012) contained a total of 2774 main events. The complete events in terms of magnitude and time from the catalogue were used to estimate the seismicity parameter. A total of 27 major faults were identified for the area. The attenuation relations developed by Atkinson and Boore (2003) for subduction zones and Boore and Atkinson (2008) for strike–slip, thrust and reverse faults were used to develop Uniform Hazard Response Spectra for the districts of Manipur for return periods of 100, 500 and 2500 years. Out of the nine districts, the hazard levels for Ukhrul, Senapati, and Imphal east and west districts were very high.

Sitharam and Sil (2014) performed DSHA as well as PSHA for the Tripura and Mizoram, states NE India. The catalogue used for performing SHA contains EQ data from 1731 to 2011. In the study, Sitharam and Sil (2014) also found out the correlation between Mw with other Magnitude scales. After carrying out the completeness test with two methods, Stepp (1972) and CUVI for the catalogue used in the study. It was found that the 4.0 < Mw < 4.9 magnitude class was complete for 49 years, 5 < Mw < 5.9 magnitude class was complete for 54years, 6 < Mw < 6.9 magnitude class was complete for 92 years, 7 < Mw < 7.9 magnitude class was complete for 106 years, and Mw > 8.0 magnitude class was complete for 180 years as per CUVI methods. As per Stepp (1972) method, 4.0 < Mw < 4.9 class was complete for 40 years, 5 < Mw < 5.9 magnitude class was complete for 50years, 6 < Mw < 6.9 magnitude class was complete for 90 years, 7 < Mw < 7.9 magnitude class was complete for 130 years and Mw > 8.0 magnitude class was complete for 200 years. The entire area was divided into six seismic regions/zones, Indo-Burma Range, Eastern Himalaya, Shillong Plateau, Bengal Basin, Naga Thrust and Mishmi Thrust. On the basis of the complete catalogue, seismicity parameters were calculated for each zone. Sitharam and Sil (2014) considered a total of 114 faults/lineaments for the study. The maximum Mmax for the study was found as 9.3 Mw.

Sitharam and Sil (2014) used Ground Motion Prediction Equations (GMPEs) given by Atkinson and Boore (2003) and Gupta (2010) for finding the seismic hazard for the area. GMPE by Gupta (2010) was given higher weight than that by Atkinson and Boore (2003), as GMPE by Gupta (2010) was region-specific. After performing DSHA, the PGA obtained by Sitharam and Sil (2014) for Agartala and Aizwal were 0.29g and 0.24g, respectively. From PSHA it was found for Agartala PGA was varying from 0.11-0.22g and for Aizwal, it was varying from 0.11-0.17g.

Das et al. (2016) prepared a probabilistic seismic hazard map for NE India region based on an improved EQ catalogue. This catalogue was based on the study by Das (2013). The catalogue contained both historical and instrumental data. The historical EQ data (from 1897 to 1962) was taken from Gupta et al. (1986). EQ events data from 1964 to 2010 are taken from global International Seismological Centre (ISC) and the National Earthquake Information Center (NEIC) databases, and data for 1963 was collected from International Seismological Summary (ISS). For the periods 1978–2006 and 1975–2010 moment magnitudes data are taken from Global Centroid Moment Tensor (GCMT) and NEIC respectively. Some EQ events data for the period 1999–2006 are also taken from seismological bulletins of the Indian Meteorological Department (IMD), New Delhi, Bapat et al. (1983) and Chandra (1992). The catalogue used by Das et al. (2016) had event information from 1897 to 2010. For the catalogue, the completeness periods are different for different magnitude classes. For 3.1–3.5 Mw class, the completeness period was 1991–2010, for 3.6–4.0 Mw the completeness period was 1980–2010, for 4.6–5.0 Mw the completeness period was 1960–2010, for 5.1–5.5 Mw the completeness period was 1950–2010, for 5.6–6.0 Mw the completeness period was 1920–2010 and for Mw≥6.0 the completeness period was 1890–2010.

Das et al. (2016) subdivided the whole study area into nine regional source zones, and by using the complete EQ catalogue, seismicity parameters were estimated for each source zone. The GMPE developed the GMPE developed by Gupta (2010) and Boore and Atkinson (2008) were used for the hazard analysis. Das et al. (2016) performed the seismic hazard calculations by using CRISIS2007, a numerical code based on the classical Cornell–McGuire method. In the study, the variation of parameters b, Mc and Mmax were also considered. By using Monte Carlo simulation, 200 sets of values were generated, and by using these values seismic hazard was estimated. By applying this procedure, the mean values of PGA, Sa and corresponding Coefficient of Variation (CoV)s were estimated in the study.

Baro et al. (2018) carried out a Seismic hazard assessment of the Shillong Plateau, NE India. Baro et al. (2018) considered 72 EQ sources (active faults) located within 500 km around the Shillong plateau (SP), considering EQ catalogue data from 1411 to 2015 with Mw> 2.5. The total number of EQs in this catalogue was 2660 (Baro et al. 2018; Baro and Kumar, 2017), and the maximum magnitude of the EQs was 8.1 Mw (1897 Assam EQ). The magnitude of completeness of this catalogue was 4 Mw. In the study, four seismic source zones were considered. For different magnitude classes and different seismic zones, the year of completeness was different. For 4.0 ≤ Mw ≤ 5.0 magnitude class year of completeness was 40; for 5.1 ≤ Mw ≤ 6.0 magnitude class year of completeness was 80-100 years; for 6.1 ≤ Mw ≤ 7.0 magnitude class year of completeness 100-120 and for 7.1 ≤ Mw ≤ 8.0 magnitude class the year of completeness was 120-160 years (Baro and Kumar, 2017). The part of the catalogue, which was complete for both magnitude and time, was taken for calculating the seismic parameters (a and b) of the area.

Baro et al. (2018) performed DSHA for all the seismic zones. GMPEs, which were proposed by Toro (2002), Nath and Thingbaijam (2011), NDMA (2010) and Anbazhagan et al. (2013) were used by Baro et al. (2018). Baro et al. (2018) performed DSHA separately for the three districts, East Khasi Hills, Ri-Bhoi, and West Garo Hills, within the Shillong Plateau. The PHA values for these districts found in the study ranged from 0.27 g to 0.46 g. Baro et al. (2018) found that the Barapani, Oldham, and Dauki faults influence the seismic hazard significantly in the studied region. Baro et al. (2018) proposed that the Oldham fault is a potential source of high seismic hazard in the SP.

Bahuguna and Sil (2018) performed both DSHA and PSHA for the state of Assam. The EQ catalogue used by Bahuguna and Sil (2018) had data from 1761 to 2015 and consisted of 6663 EQ events. This catalogue contained great EQs of 1897 Assam EQ (Mw=8.1) and Assam EQ 1950 (Mw=8.6). Bahuguna and Sil (2018) identified a total of 29 linear (17 active and 12 inactive) seismic source zones but considered 17 potential seismic source zones based on seismic parameters and maximum potential magnitude for the hazard analyses. Bahuguna and Sil (2018) used two GMPEs for the study. One was proposed by Nath et al. (2009), and the other one was modified from the GMPE proposed by Jain et al. (2000). The equation of the modified GMPE was the same as that of Jain et al. (2000), but the coefficients of the equation were calculated by Bahuguna and Sil (2018) based on the strong ground motion data available for the study area.

From DSHA, Bahuguna and Sil (2018) found that Peak Horizontal Acceleration (PHA) for different parts of the state of Assam was between 0.5g and 1.63g. By performing PSHA, Bahuguna and Sil (2018) found that PHA values for the cities of Assam for 2% probability in 50 years were between 0.43g and 0.78g and for 10% probability in 50 years the PHA values were between 0.32g and 0.58g.

2.7 Review on Seismic Source zonation:

As mentioned in the previous section, numerous researchers (Sharma and Malik 2006; Pallav et al. 2012; Sitharam and Sil 2014; Das et al. 2016; Baro et al. 2018; Bahuguna and Sil 2018) have performed SHA of different parts of the NE India region. These were based on dividing the seismotectonic region into various source zones. Sharma and Malik (2006) performed PSHA of NE India considering four seismic zones for the entire region. These four zones were further divided into 10 seismogenic zones. The zones that Sharma and Malik (2006) considered were mainly based on fault locations and alignments. As a result, no uniformity in the distribution of EQs within each source zone was ensured. In addition, demarcation of boundaries of such zones was also based on visual observations of the tectonics of the region, which may vary from person to person. In another work where seismic hazard of Manipur region was attempted, Pallav et al. (2012) divided the entire seismotectonic region into seven seismic source zones based on the tectonics and geology of the region. Source zones considered by Pallav et al. (2012) were based on the works done by Goswami and Sarmah (1982), and Nandy (2001). Sitharam and Sil (2014), while performing SHA for Tripura and Mizoram, divided the entire area into six seismogenic regions/ zones namely; Indo-Burma Range, Eastern Himalaya, Shillong Plateau, Bengal Basin, Naga Thrust and Mishmi Thrust. All these six zones were delineated based on visual observation of event distribution patterns and source orientations. In another work, Das et al. (2016) developed a probabilistic seismic hazard map for the NE India region considering nine regional source zones. Source zones considered by Das et al. (2016) were based on the work done by Angelier and Baruah (2009), where tectonics, regional geology, focal mechanisms and spatial distribution of seismicity were considered for zonation. Baro et al. (2018) and Baro et al. (2020) carried out DSHA and PSHA of the Shillong Plateau, respectively, considering four seismic source zones following Baro and Kumar (2017). As per Baro and Kumar (2017), the 500km radial distance based seismotectonic map for the Shillong Plateau consists of four source zones namely; the Eastern Himalaya zone, the Indo-Burmese zone, the Shillong Plateau-Assam Valley Zone and the Bengal Basin zone. Identification of the above source zones was based on variations in geology, overburden thickness and seismotectonics. In addition, as per Baro and Kumar (2017), these four source zones show differences in terms of rupture characteristics and rate of movement. In another work, while performing DSHA and PSHA of Assam, Bahuguna and Sil (2018) identified a total of 29 linear (17 active and 12 inactive) seismic source zones. However, for SHA, Bahuguna and Sil (2018) considered 17 potential seismic source zones based on seismic parameters and the maximum potential magnitude associated with faults present in each source zone. It should be highlighted here that the source zonation by Bahuguna and Sil (2018) also were based on visual observation of fault alignment and location of the seismic sources present in the region.

In addition to the above-mentioned work, Gupta (2006) considered seismicity along with the tectonic features for the delineation of probable seismic sources in India and its neighbourhood. Out of 81 seismic sources defined by Gupta (2006), almost 20 seismic source zones were near or within NE India. It can be highlighted here that since the demarcation of whole zonation was done based on visual observation, some zones, such as zone 22 (see Figure 7 as per Gupta (2006)) had different clusters of EQs within the same zone. In another work, Kolathayar and Sitharam (2012) divided the entire India into 104 zones based on the pattern of seismic event distribution. Since Kolathayar and Sitharam (2012) took Gupta (2006) work as a reference, similarity in terms of zones in these two studies can be seen. Later, zones proposed by Kolathayar and Sitharam (2012) were used by Sitharam and Kolathayar (2013) as areal sources for SHA of the entire India.

2.8 Review on Paleoseismological Studies

Different methods have been used by numerous researchers for different terrain types and for different regions to relate EQ features with paleoearthquakes. Following subsections provide a brief review on paleoseismological studies around the world and for the NE India region.

2.8.1 Paleoseismological study around the globe

Berryman et al. (1989) conducted a paleoseismological study on the Hikurangi subduction zone in eastern North Island, New Zealand. Berryman et al. (1989) attempted to understand the origin of the Holocene paleoseismicity marine terraces that were abundant in the area and whether their origin was coseismic. Based on morphology, Berryman et al. (1989) concluded that the terraces came into existence due to the uplift caused by past EQs. Based on past seismic activity, possible behaviour in the future was also evaluated. Berryman et al. (1989) applied ^{14}C dating on the beach deposits on top of marine-cut shore platforms to find the age of individual terraces. At least 21 paleoseismic events in the study area had magnitudes between 7.3 and 8.0. In the study, Berryman et al. (1989) also found that the uplift recurrence intervals of the terrace varied from 0.4 ka to 2.0 ka and individual uplift varying from 1.0m to 4.0m.

Adams (1990) studied turbidites (sea features formed during EQ) in Cascadia (North America) to find out the past EQs in that area. Based on the above study, almost 13 great EQs, which had occurred in the Cascadia subduction zone, were detected. Dating the pelagic intervals between the turbidites, Adams (1990) found that the last EQ occurred 300±60 years ago. Based on these findings, the probability for a great EQ to occur was found to be 2-10% in the next 50 years in Cascadia.

The Surface rupture patterns and displacements along the Greendale fault (New Zealand) were studied by Hornblow et al. (2014). It must be mentioned that the Greendale fault was a hidden fault and was not known till 2010 Darfield EQ (Mw-7.1). Based on the study, Hornblow et al. (2014) concluded that the Greendale fault could generate larger EQ in the future. In addition, Hornblow et al. (2014) attempted to characterise the Greendale fault based on past seismic activity from the study for preparing EQ hazard model. Further, based on LiDAR and

GPR, suitable sites for trenching were identified to find displaced cultural features. Based on OSL dating, the age of these features was estimated in the range of 21.6±1.5ka to 33.3±2.0ka for the first site and 24.7±1.5ka and 20.2±1.9ka for the second site. Thus, findings of past EQs helped in a more reliable estimate of the seismic hazard of the study area.

In another work, Vallejo et al. (2003) studied a series of clastic dikes and tubular vents in southern Tenerife (Canary Islands of Spain). Before the study, Canary Island was believed to be a less seismically active zone, but there was almost no information about past EQs and a minimal instrumental record database. Based on geological evidence, Vallejo et al. (2003) found that such evidence was of seismic origin due to liquefaction. Further, based on thermos luminescence dating methods, paleoearthquakes that were not known earlier were detected. Results from dating suggested that the liquefaction features were formed 3490 ± 473 years Before Present (BP), on sand beach, which was, formed 10,081 ± 933 years BP. In addition, four methods namely; the cyclic stress method, the Ishihara (1985) method, the magnitude bound method and energy based solutions were adopted to find the strength of paleoearthquakes. First two methods were used to estimate the PGA which caused the liquefaction while rest two methods were used to estimate the magnitude of paleoearthquake. Based on above methods the PGA was computed to be in a range of 0.22g to 0.30g while the magnitude was estimated in between 6.4 to 7.2. In addition, suitability of indirect evidence in identifying such paleoearthquake (in terms of time as well as strength) was highlighted in a more reliable estimate of the seismic hazard of the study area.

2.8.2 Paleoseismological study in India

In addition to studies performed in different parts across the globe, as discussed above, several studies related to the identification of paleoearthquakes based on both direct and indirect features were also conducted in India. Rajendran and Rajendran (2001) surveyed the Kutch region (north-western India) to study the 1819 Kutch EQ (Mw 7.7 to 8.2), which was not studied earlier. Based on the work, Rajendran and Rajendran (2001) explored the Allah Bund area of Kutch where the surface scarp was observed for the elevation and morphology to identify seismically induced features and the ages of corresponding paleoearthquakes. Based on the work, the height of uplift during 1819 Kutch EQ was found to be 5.3m. Further, based on the correlation between epicentral distances and size distribution of liquefaction features, the strength of the 1819 Kutch EQ was found to be approximately 7.5 in Mw scale. In addition, study by Rajendran and Rajendran (2001) concluded that the Allah Bund is a compound scarp, which has been formed due to repetitive EQs and thus may represent a surface fold. In addition, excavation trenches in the region revealed various liquefaction features which when dated back suggested ages of liquefaction in between 800 to 1000 years BP.

In another work, Kumar et al. (2006) conducted paleoseismological investigation along the HFT system in India to understand the stress accumulation and strain release relationship, to find out the possible seismic hazard associated with HFT. Trenches were excavated on fault scarps near the cities and villages of Chandigarh, Kala Amb, Rampur Ganda, Lal Dhang, and Ramnagar, and collected samples were radiocarbon-dated to find the timing of the events. Based on the sample analyses, vertical slips of 11 to 38m were determined on the HFT. Further, Kumar et al. (2006) showed that for such a slip to occur, 1330 to 3250 or more years would be required. In addition, the future possibility of experiencing larger EQs, having rupture lengths even beyond 250km, was highlighted.

Joshi et al. (2009) highlighted that the majority of Himalayan EQs have occurred on blind faults, and thus, the evidence of surface traces of such faults or even liquefactions is very rare. For the first time, Joshi et al. (2009) found the paleoseismic features in frontal belt of the Western Himalayas. In addition, paleoseismic investigations of liquefaction features near the Yamuna Tear Fault in Delhi were conducted. Dikes and other small-scale deformation features in the area were studied to find the morphology and the mechanism which led to the formation of such features. Carbon dating revealed that the age of formation of the dikes was between 930 and 670 years BP. Further, based on Ishihara (1985) method and considering MCT as source, the magnitude was found to be greater than 8 in Mw scale. However, in case Yamuna Tear would be the EQ source, the magnitude was more than 7 in Mw scale.

Malik et al. (2010) studied the areas which were severely affected during 1905 Kangra EQ. Due to limited information about past EQs in the Himalayas, accurate determination of seismic hazard is challenging. With this problem definition, Malik et al. (2010) tried to find out the EQ associated with the faults in the above area. By using CORONA KH-B data, SRTM data and toposheet from Survey of India, two active fault traces were identified namely; Hajipur faults (HF1 and HF2) on the left bank of the river Beas. Based on the morphology and considering the length of these two faults, Malik et al. (2010) concluded that HF2 is younger and more active than HF1. Further, two trenches across HF1 were excavated, and samples for dating were collected. Firstly, based on stratification in the trenches, younger and older stratum were identified, which helped in finding two EQs along HF2 and the sequences of faulting and depositional phases that lead to forming the present features. Samples collected from trenches, when analysed based on OSL and radiocarbon dating methods, Malik et al. (2010) found information on the two events that had occurred 2.6ka and 550 BP. In addition, the slip rate, uplift rate, shortening rate and recurrence interval of the Hajipur fault were determined as 7.6±1.7mm/year, 3.2±0.6mm/ year, 6.9±1.4mm/ year and 1160±250year respectively. Thus, study by Malik et al. (2010) resulted in identification of Hajipur fault as well as finding two paleoearthquakes which were not known earlier and are important in understanding ongoing seismic activity of the region.

2.8.3 Paleoseismological study in NE India

Sukhija et al. (1999) investigated the meizoseismal area of the 1897 Shillong EQ to explore the paleoearthquakes in NE India as well as Himalaya and adjoint area for the first time. Earlier, the region had no data on the recurrence of EQs as well as information about EQs happened earlier than one hundred years. With this objective, Sukhija et al. (1999) mainly focussed on finding the paleoearthquake information in the area. Primary features of EQ in this region are very less. For this reason, paleoliquefaction in the over-bank silts and sand deposits of river in plains were adopted. During the field investigation, well-preserved liquefaction, as well as deformation features at different sites within the alluvial deposit, were found. Carbon dating of samples collected suggested that such geological features came into existence during three major EQs in addition to the great 1897 Assam EQ. Results showed that before the great 1897 EQ, events of great intensities had occurred during 1450–1650,700–1050 AD and near 600 AD. Considering all such EQs, the return period of major EQs in the Shillong Plateau was found as about 400–600 yr as per Sukhija et al. (1999).

Thomas et al. (2007) studied the liquefied sand structures formed during major EQs in Upper Assam near the River Buri-Dihing with an objective to find the time of occurrence of these EQs. Optically stimulated luminescence (OSL) test was performed on the quartz extracted from samples, which were taken from the sand dyke of the liquefaction structure. The result showed that the liquefaction features were formed between 1430 AD and 1630 AD. The outcome of Thomas et al. (2007) can be compared with one paleoearthquake found by Sukhija et al. (1999) that had occurred, between 1450 AD and 1650 AD.

Similarly, Reddy et al. (2009) carried out field investigation on different locations of the meizoseismal area of the great 1950 Assam EQ to find out the paleoseismological history of the area. Direct paleoseismological features were found to exist in this region. As this area was ideal for liquefaction; samples were collected from the trench along the rivers Buri-Dihing and Dibang. ^{14}C dating of the organic material collected from the study area indicated a paleoseismic record of about 500 years. The study by Reddy et al. (2009) was helpful for seismochronological study of the area.

Kumar et al. (2010) studied the paleoseismic evidence of Nameri in the northeast along with the entire northeastern Himalaya. Since no study was attempted prior to Kumar et al. (2010), the objectives were to find out the timing, size rupture, return period etc., for the fault associated with the seismicity of the area and EQ mechanism. Based on the work, Kumar et al. (2010) mapped active faults in the region by using SRTM imagery. Further, trenches were excavated across active fault scrap to understand the structure and stratigraphy of the trenches. Later, collected samples from the trenches when analysed using ^{14}C dating suggested possible EQ after 1100 AD.

Mishra et al. (2016) presented the paleoseismic evidence of 1255 A.D. EQ in eastern Himalaya near Darjeeling and suggested this megathrust EQ produced a surface rupture of nearly 800 km through, which needs to be further crosschecked by more paleoseismic investigations in this area.

Kumar et al. (2016) studied the liquefaction features at the Kopili Fault Zone, previously identified as the "Assam seismic gap". Paleoseismological investigation in that area was performed by Kumar et al. (2016) to understand the past seismicity of the area and the possible seismic hazard. Samples were collected from trenches in two locations. One of these trenches was excavated for highway construction, and another one was excavated for new road construction near the bank of Kalang River. For the collected samples, both OSL and ^{14}C dating methods were used to determine the ages of the EQs. Collected sample of sand dyke when analysed, suggested the ages of formation were found 250±25 years BP, between 400 to 770 years BP and 900±50 years BP.

The most recent work carried out by Priyanka et al. (2017) solves the ambiguity over the year regarding the surface rupture of the 1950 Great Upper Assam EQ. The study did paleoseismic investigation in the lower Brahmaputra valet near Pasighat, where the active fault deformed the recent terrace of the river. The radiocarbon dating of the deformed using along the fault and the bracketing unit which covers the deformed unit suggest that the last EQ in the fault is 1950 Upper Assam EQ of magnitude Mw=8.6. If we combine the previous study by Jayangondaperumal et al. (2011), which was intended to study the near-surface expression of late Holocene EQ in NE Himalaya near Marbang, the surface rupture of 1950 Upper Assam EQ is about 20-25 km in Arunachal Pradesh Himalaya.

Table 2.2 summarises all the work so far performed to find paleoearthquakes, the methods used, and typical findings in terms of EQ events in the northeastern part of India, as discussed above.

Table 2.2 Summary of the literature about paleoseismic study around NE India

Studied by	Study Area	Paleoseismological Features	Methods of Dating	Findings	Magnitude of the closest EQ occurred during the period
Sukhija et al. (1999)	Meizoseismal area of the 1897 Shillong EQ	Paleoliquefaction features (in over-bank silts and sand deposits of river in plains)	Carbon dating	Found the possible EQs in 1450–1650AD ,700–1050 AD and close to 600 AD with Return period of 400–600 yr	1. 7Mw (1596 AD) 2. 8Mw (825AD) 3. No EQ data available close to 600AD)
Thomas et al. (2007)	Upper Assam near the River Buri-Dihing	Sand dyke	OSL test	Found possible EQ between 1430 AD and 1630 AD	1. 7Mw (1548AD)
Reddy et al. (2009)	Different location of the meizoseismal area of the great 1950 Assam EQ	Sand dyke	Carbon dating	Possible EQ was found of about 500 years from 2009	1. 7Mw (1548AD)
Kumar et al. (2010)	Nameri in the northeast along with the entire	Fault scrap	Carbon dating	Possible EQ after 1100 AD is found	1. 8.5Mw (1100AD)

	northeastern Himalaya				
Mishra et al. (2016)	Eastern Himalaya near Darjeeling	Fault scarp	Carbon dating	Possible EQ near A.D. 1152 which may be the evidence of 1255 A.D. EQ in the area	1. 8.8Mw (1255AD)
Kumar et al. (2016)	Kopili Fault Zone.	Sand dyke	Both OSL and Caron dating	Possible EQs in about 250±25 years BP, between 400 to 770 years BP and 900±50 years BP	1. 8Mw (1714AD) 2. 7Mw (1548AD) 3. 8.5Mw (1100AD)
Priyanka et al. (2017)	Lower Brahmaputra valley near Pasighat	Fault scarp	Carbon dating, radioactive isotopic dating	Found the surface rupture of 1950 great Upper Assam EQ	1. 8.6Mw

2.9 Existing GMPEs for NE India

Numerous GMPEs (Jain et al. 2000; Toro 2002; Atkinson and Boore 2003; Boore and Atkinson 2008; Nath et al. 2009; Gupta 2010; NDMA 2010; Anbazhagan et al. 2013) have been used by various researchers to determine the seismic hazard level of the NE India (NDMA 2010; Pallav et al. 2012; Sitharam and Sil 2014; Das et al. 2016; Bahuguna and Sil 2018; Baro et al. 2018). A careful observation, however indicates that not all the GMPEs used in the earlier works were developed for the NE India region. Toro (2002) developed GMPE for Central and Eastern North America for hard rock site conditions that can be used for EQ of magnitude between 5 to 8 (Mw), having distances less than 1000km. GMPE by Atkinson and Boore (2003) was developed for subduction zones around the world, including both interface and in-slab events. This GMPE is applicable in rock site condition for EQ of magnitude 5 to 8 (Mw) and distances less than 550km. In another work, Boore and Atkinson (2008) developed GMPE from globally recorded EQ accelerograms. Due to the inclusion of many EQ accelerograms, this GMPE is applicable for EQs having magnitude from 5 to 8 (Mw), distance less than 200km, and for V_{S30} values between 180 and 1300m/s.

GMPEs discussed in the last section were developed for other regions but were used for NE India. In a similar manner, GMPEs, which were developed for NE India, also exist. Jain et al. (2000) developed a GMPE to predict the PGA for the entire Himalayan region, considering the accelerograms of 9 EQs having magnitudes from 5.2 to 7.3 (Mw). In their work, besides 100 station-recorded (distance from 2km to 327km) accelerograms, Jain et al. (2000) also used 300 accelerograms from Structural Response Recorders (in the distance ranging from 4km to 772km). In another work, Nath et al. (2009) developed a GMPE for Guwahati region of NE India based on synthetic ground motions generated using the properties extracted from recorded strong-motion waveforms. This GMPE was developed by considering EQs of magnitudes from 4.8 to 8.1 (Mw), distances up to 100km, and can be used for different Site Classes (SCs) based on V_{s30} values. In another work, while performing PSHA for the entire India, NDMA (2010) developed a GMPE applicable to SC A within the NE India region. To develop the GMPE, NDMA (2010) utilized synthetic accelerograms generated for EQ of magnitude 4 to 8.5 (Mw), considering distances up to 500km. Gupta (2010) in an attempt considered the functional form of the GMPE by Atkinson and Boore (2003) to develop a new GMPE for NEHRP SC B within NE India. Work by Gupta (2010) utilised recorded accelerograms of 3 EQ events having magnitudes from 6.3Mw to 7.2Mw corresponding to distance range from 150 to 375km. Further, these accelerograms were mainly from SC B and SC C. Work by Anbazhagan et al. (2013) developed a GMPE for the entire Himalayan region by utilising both recorded as well as synthetic EQ accelerograms. The developed GMPE was applicable for SC A condition with magnitude range from 5.3Mw to 8.7Mw and up to epicentral distance of 300km. Bajaj and Anbazhagan (2019) also developed a GMPE for rock level for the Himalayan region. Bajaj and Anbazhagan (2019) utilised simulated synthetic ground motion generated considering regional parameters and properties faults.

Besides above-discussed GMPEs, some more GMPEs are also available for NE India. In one attempt, Baruah et al. (2009) developed a GMPE which can be used to calculate only the PGA value for Shillong plateau region. Baruah et al. (2009) utilized the recorded ground motions from 82 EQ events recorded mostly on rock sites to develop the GMPE. EQ events having magnitude between 2.5 and 5 (Mw) and up to distance of 145km. Singh et al. (2016) developed a GMPE for NE India crustal EQs using synthetic ground motions generated with a stochastic point source model. This GMPE can be utilized for different site classes also. In another work, Ramkrishnan et al. (2022) developed a GMPE for NE India considering magnitude range of 4.2 to 6.9 (Mw) and maximum distances up to 640km. In their GMPE, Ramkrishnan et al. (2022) considered recorded acceleration data of 24 EQ events from NE India occurred from 1986 to 2014. Besides this, Ramkrishnan et al. (2022) considered all type of SCs together and developed a single GMPE for all types of SCs. In another work, Rahman and Nunia (2020) developed new GMPEs for NE India region considering recorded ground motion from a total of 10 EQs. The EQ magnitude considered in the GMPE development process had magnitude from 4.8 to 6.9 (Mw). As most of the recorded ground motions are located on the SC-C, the GMPE developed by Rahman and Nunia (2020) is only valid for SC-C. Besides the above discussed works, a completely different type of GMPE was developed by Dhanya and Raghukanth (2020). Dhanya and Raghukanth (2020) developed a neural network-based hybrid GMPE model for NE India. This model, based on the input magnitude, distance, shear wave velocity, fault type and region, can predict the possible spectral acceleration.

2.10 Research Gap

From the literature mentioned above for SHA in NE India, it is seen that the numbers of great EQs are much less (only two 1897 and 1950 Assam EQs), which were considered for hazard analysis purposes. Moreover, if we notice the seismotectonic maps of the researchers, we can see that for some regions, EQ events data is absent, though some sources were shown in the maps (see Figure 1 in Das et al. (2016)). Similarly, for some regions, EQ events are there, but sources are absent (See Figure 3 in Sitharam and Sil (2014), Figure 2 in Baro et al. (2018) and Figure 4 Bahuguna and Sil (2018)); for some regions, neither sources are there, and nor adequate events information is available (see Figure 2 Baro et al. (2018)). This incomplete information is affecting the estimation of the seismic hazard. For proper estimation of seismic hazards, knowing all seismic sources of an area and their highest potential is required. There is a need here to find events of great magnitude, if available, in the area and if any event occurred in the area that had more magnitude than the available knowledge of EQ magnitude.

Table 2.3 Past Work on SHA for different parts of NE India

Studied by	EQ catalogue duration	Data completeness	4 ≤Mw< 5	5 ≤Mw< 6	6 ≤Mw< 7	7 ≤Mw< 8	8 ≤Mw
Sharma and Malik (2006) (NE India)	1762-2001 AD (1220 events)	Not performed					
Kanth et al. (2009) (Imphal)	1869 cachar EQ, 8.1Mw in Indo-Burma	No catalogue					
Pallav et al. (2012) (Manipur)	825-2009 AD (2774 events)	yes	1959–2009	1909–2009	1809–2009	1609–2009	Couldn't find
Sitharam and Sil (2014) (Tripura Mizoram)	1731-2011 AD (3251)	yes	1962-2011	1957-2011	1919-2011	1905-2011	1831-2011
Das et al. (2016) (NE India)	1897-2010 AD (2674)	yes	4.6–5.0 Mw (1960–2010)	5.1–5.5 Mw (1950–2010)	5.6–6.0 Mw (1920–2010)	Mw≥6.0 (1890–2010)	
Baro et al. (2018) (Shillong Plateau)	1411-2015 AD (2359)	yes	1975-2015	1915-2015	1895-2015	1855-2015	Couldn't find
Bahuguna and Sil (2018) (Assam)	1762-2015 AD (2882)	Yes	1971-2015	1961-2015	1921-2015	1881-2015	1811-2015

Another problem with the past study is the completeness periods of EQs with different magnitudes. Table 2.3 shows the completeness periods of EQs for the past studies. It is seen in

Table 2.3 that only Pallav et al. (2012) had an EQ catalogue dated back to 825 AD. Besides, Pallav et al. (2012) considered an EQ catalogue complete from 1609. Other than Pallav et al. (2012), other studies only had a complete EQ catalogue from the 19th century. Moreover, the studies did not utilise information on great EQs that occurred in an adequate period. This unavailability of information can be dealt with by including paleoearthquake information. Though most of these studies utilised some of the paleoearthquakes, those are not enough as more information is available nowadays. Besides, these paleoearthquakes have some uncertainties not addressed in past studies.

Based on the past literature, it is observed that numerous studies on seismic source zonation of parts of NE India exist. Further, few studies exist that focus on seismic source characterisation of the entire NE region. However, limitations in existing studies in terms of not ensuring uniform distribution of EQs as well as seismic activity within considered source zones can be observed. Further, demarcations of zones in many existing studies were done based on visual perception, which can vary from person to person. It should be noted that the primary objective of source zonation is to identify regions having the same seismic activity throughout (Hashemi and Karimi, 2016). However, source zonation done by previous researchers has some parts with more EQs while other parts of the same zone have less numbers of EQs. Visual observation-based demarcation of zone boundaries based on the distribution of EQs may be the attribute. Based on the above discussion, it can be clearly stated that SHA considering source zones from existing studies may not be fully appropriate for the NE India region.

Almost all the past SHA studies follow the traditional PSHA where areal or linear sources have been used, considering a uniform distribution of seismicity throughout the sources. A careful observation of the above-mentioned studies for NE India highlights that most of the SHA works are performed based on one source zonation definition. However, as per Roe et al. (2014), significant uncertainties exist regarding the position and geometry of seismic sources such as faults. Thus, considering multiple source models would be more accurate in accounting for these epistemic uncertainties (Grünthal et al. 2018).

Most of the past SHA works considered a uniform distribution of seismicity throughout the faults and areal seismic sources. However, it is seen that no matter the method used to delineate the seismic source zones, the seismicity is not uniform throughout the entire seismic source. Seismic activity can vary significantly throughout a fault or areal source, leading to localised regions of higher seismicity or low seismicity. These pose a significant challenge to SHA, as traditional PSHA methods may not accurately capture them. As per Beauval et al.

(2006), the assumption of homogeneity of seismic activity showed a bias for higher hazard values. The same effect becomes severe if the hazard level is considered for a longer return period. Therefore, it is crucial to consider the spatial variation of seismicity and the presence of seismic gaps when performing SHA, particularly in seismically active regions like NE India.

For almost all the past SHA studies, the hazard levels were obtained for bedrock or surface rock conditions. The reason is due to the absence of adequate GMPE for the region. It is observed that existing GMPEs for NE India have limitations in terms of magnitude range, distance range and consideration of site conditions. While looking at the history of past damaging EQs, it is clearly understood that the NE India region has the potential of great EQs with a magnitude as high as 8.7Mw. However, out of all the earlier mentioned GMPEs, only some GMPEs, such as Nath et al. (2009), NDMA (2010) and Anbazhagan et al. (2013), are applicable to great EQs. Some of the GMPEs are not applicable for lower magnitudes. It is also observed that some of the GMPEs were developed considering significantly fewer numbers of EQ events. This may lead to the inefficiency in capturing the randomness of the ground level shaking that may relate to the different source properties of different EQs to predicted ground motions. Another limitation is the range of epicentre distance or hypocentral distances for many of the GMPEs are applicable. Only GMPEs by NDMA (2010), Gupta (2010), Rahman and Nunia (2020) and Ramkrishnan et al. (2022) were developed considering distances greater than 300km. In NE India, during major to great EQs, the damages have been seen at large distances. In such a case, the GMPE should be capable of estimating ground shaking levels for large distances. Further, keeping in mind various forms of induced damages which have been witnessed during various EQs in the past in NE India, none of the existing GMPEs for NE India except the GMPEs by Nath et al. (2009) and Singh et al. (2016) considered the effect of different SCs while developing GMPE. Other GMPEs are limited to specific SC and not for all SCs. As a result, this also restricts the use of existing GMPEs to determine seismic forces for building design as well as quantification of induced effects.

2.11 Objectives of the Present Work

The main target of the present work is to estimate the seismic hazard levels for NE India by considering an updated EQ catalogue. Besides, the present work aims to understand the influences of different magnitudes and sources in the result. Thus, the objectives of this study are:

1. Delineation of seismic source zones for NE India.

2. Determination of the seismic hazard level for the entire NE India considering various types of uncertainties.
3. To identify the critical seismic sources and magnitudes of EQs that significantly influence the seismic hazard level in NE India.
4. To estimate the seismic hazard levels for NE India by incorporating paleo EQs with their uncertainties.
5. To develop a GMPE for NE India applicable to different SCs.

Chapter 3. EQ Data Collection and Data Processing for the Study Area

3.1 General

For the evaluation of seismic hazard levels and understanding the seismic scenario of a particular region, information about past EQs and seismic sources (such as faults) is essential. This chapter provides details of the past EQ data collected and its processing for NE India.

3.2 Development of EQ Catalogue for the NE India Region

Information about past EQs (including magnitude, epicentre coordinates, time of occurrence, focal depth etc.) that had previously occurred in and around NE India are collected from different sources namely; USGS, ISC, IMD, NDMA (2010), National Center for Seismology (NCS) (https://riseq.seismo.gov.in/riseq/earthquake/archive) etc. NDMA (2010) has EQ information collected from Oldham (1883), Iyengar et al. (1999), Rajendran et al. (2004) etc. Considering the fact that even a distant EQ can cause a significant seismic hazard in a region, past EQs from surrounding regions should also be considered while developing the EQ catalogue for a region (Kumar et al. 2016a, 2017). For this reason, EQ data in this work are collected for a longitude range of 85° E to 100 °E and a latitude range of 18°N to 32°N. A total of 13,204 EQ events (till the year 2018) are obtained from the above-mentioned sources in different magnitude scales (Mw, ML, Ms, mb etc.). For further processing, all these EQ magnitudes (which were in different magnitude scales) are converted to the Mw scale using the following correlations developed by Sitharam & Sil (2014) for NE India to obtain a homogenised EQ catalogue for the region;

$$Mw = 0.862mb + 1.034 \tag{3.1}$$

$$Mw = 0.673ML + 1.730 \tag{3.2}$$

$$Mw = 0.625Ms + 2.350 \tag{3.3}$$

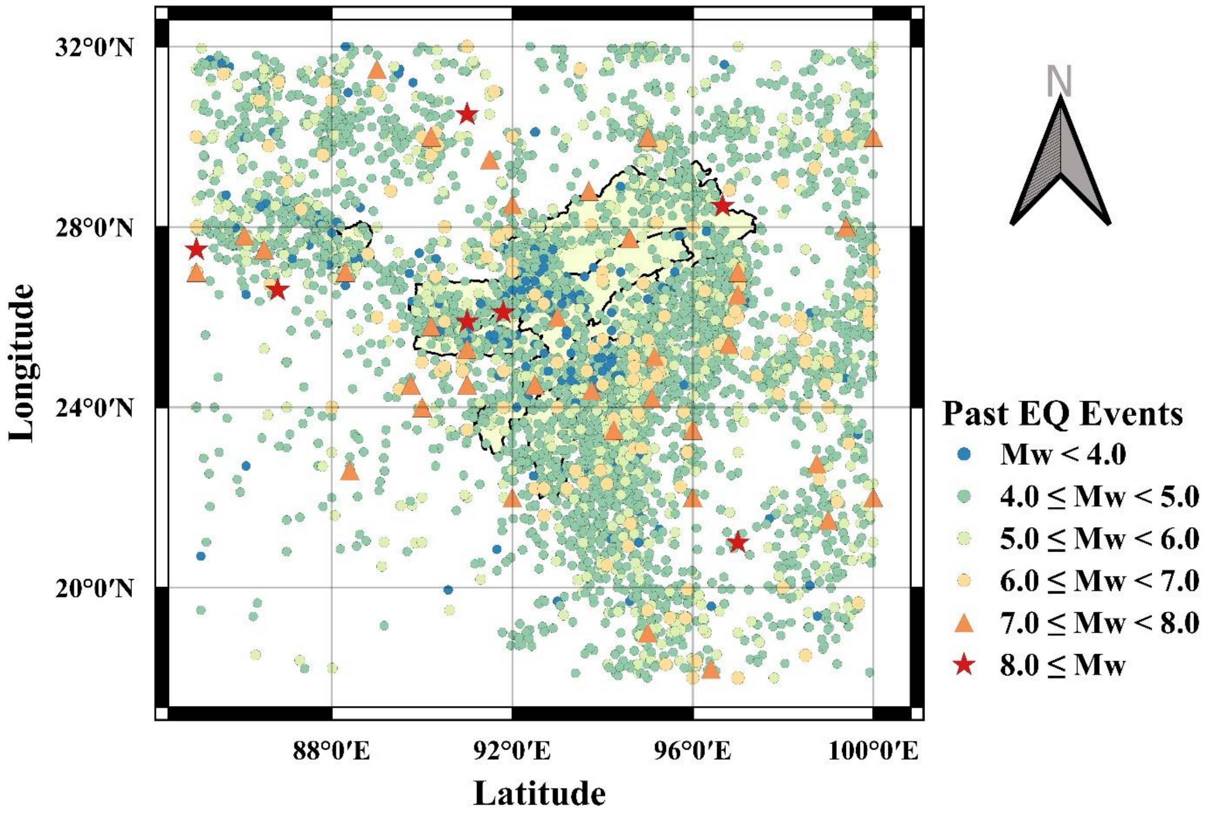

Figure 3.1 Past EQ events in and around NE India obtained from different sources

Once the entire database is converted to the Mw scale, the magnitude range of collected EQ events is found to vary from 2.5-8.6 (Mw). Further, because the past EQ information for the same region is collected from multiple sources, it might be possible that the same EQ event is reported by more than one source. As a result, the above-collected database may have duplicate events. Hence, such duplicate events should be identified and removed before analysing the database further. To do so, duplicate events are identified by searching events with the same occurrence time, then epicentre coordinates and then magnitude. This way, identified duplicate events are removed from the EQ catalogue. After the removal of duplicate events, a total of 7,323 numbers of past EQs are left in the catalogue (as shown in Figure 3.1). Out of these 7,323 events, EQ events having $Mw < 4.0$, $4.0 \le Mw < 5.0$, $5.0 \le Mw < 6.0$, $6.0 \le Mw < 7.0$ and $7.0 \le Mw < 8.0$, and $8.0 \le Mw$ are 241, 5660, 1103, 269, 42 and 8 in numbers respectively. Figure 3.2 illustrates Number of EQs vs Magnitude plot for NE India region.

Similarly, Figure 3.3 shows Magnitude vs epicentral distance of the EQs from Guwahati City (91.74°E, 26.14°N) as centre.

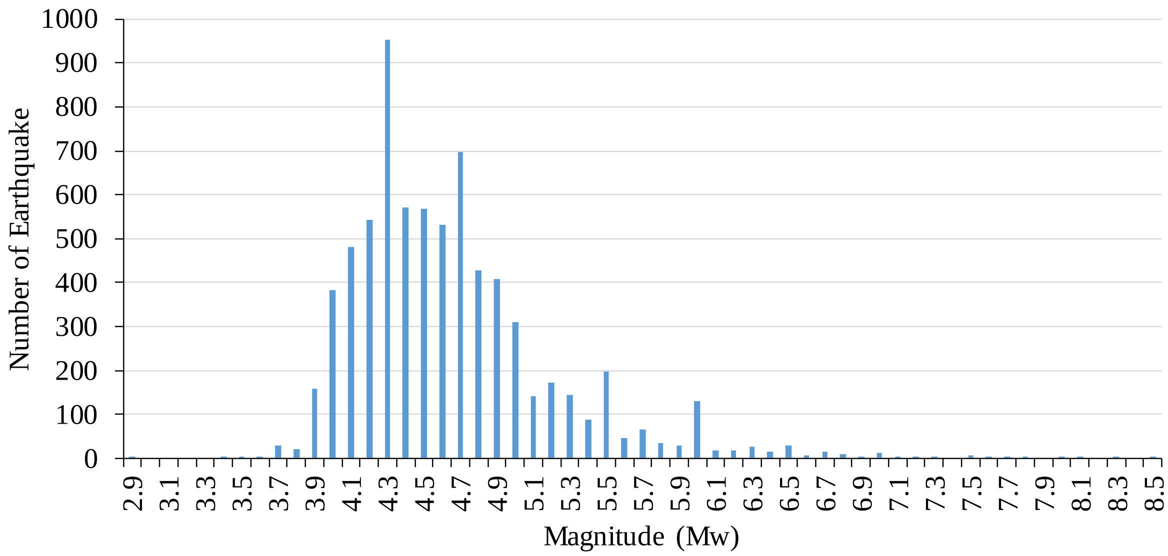

Figure 3.2 Number of EQs vs Magnitude plot for NE India region

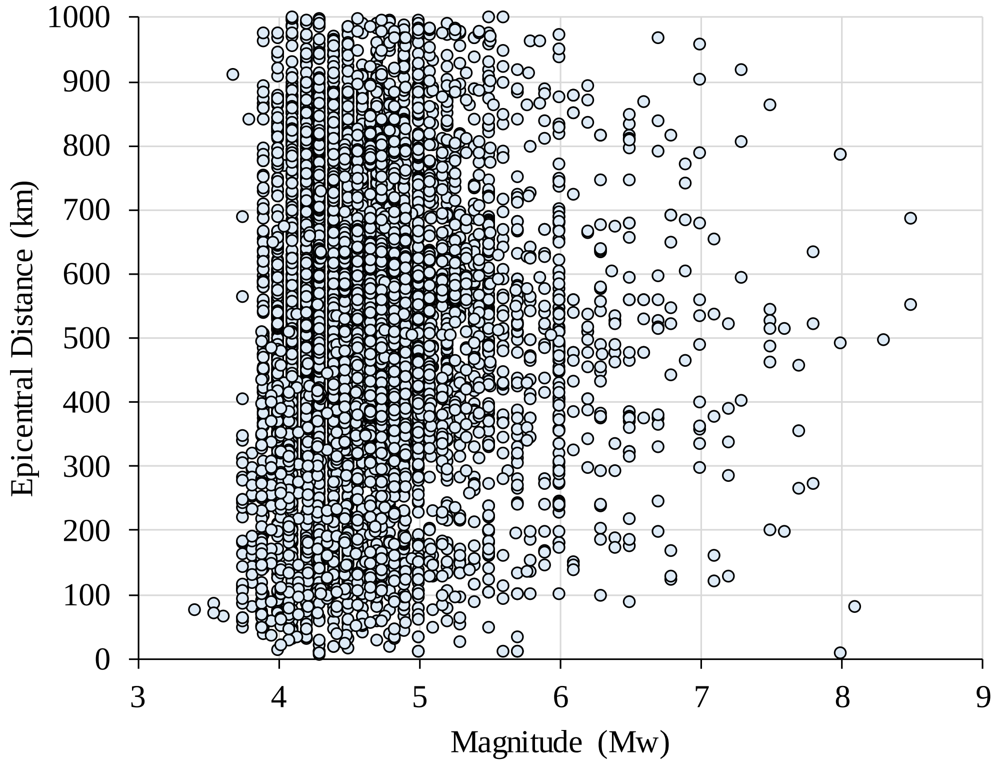

Figure 3.3 Magnitude vs Epicentral Distance plot for NE India considering Guwahati city as centre

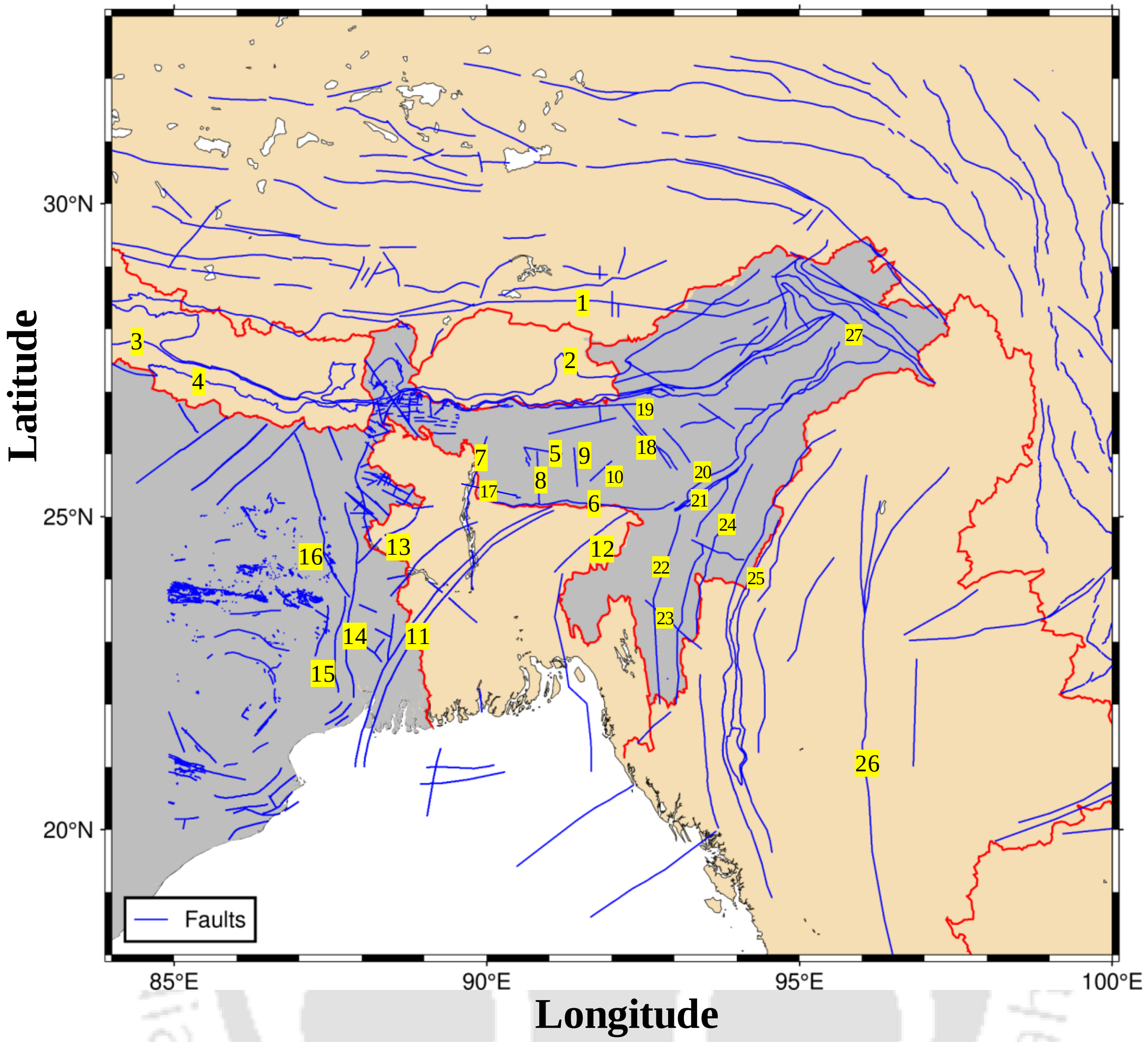

Figure 3.4 The seismic sources/faults present in and around NE India obtained from different sources

3.3 Fault Information

Information about the faults present in NE India is obtained from Bhukosh, Baro and Kumar (2017), Sloan et al. (2017), Chen (2020), and Kundu and Gahalaut (2013). Figure 3.4 shows the faults present in and around NE India. Information about some of the faults and the EQs that occurred on them has already been discussed in Chapter 2. Figure 3.4 shows some of the important faults of the region discussed in Chapter 2 with the numbers mentioned above in the bracket after the fault names. The faults present in the Eastern Himalayan Region are IST (1), MCT (2), MBT (3), HFT (4), and Mishmi Thrust (27). The important faults present in Shillong Plateau are Oldham Fault (5), Dauki Fault (6), Dhubri Fault (7), Dudhnoi Fault (8), Kulsi Fault (9), Barapani Shear Zone (10), Dapsi Thrust (17). Some of the important faults present in the Bengal Basin are Eocene Hinge Zone (11), Sylhet Fault (12), Jangipur–Gaibandha Fault (13), Garhmayna–Khanda Ghosh Fault (14), Pingla Fault (15), Sainthia–Bahmani Fault (16). Some of the faults present in the Assam Valley region are Kopili Fault (18) and Atherkhet Fault (19). In the Indo-Burma ranges, the important faults present are Naga Thrust (20), Disang Thrust (21), Kaladan Fault (22), Mat Fault (23), CMF (24), Kabaw Fault (25), and Sagaing Transform Fault (26).

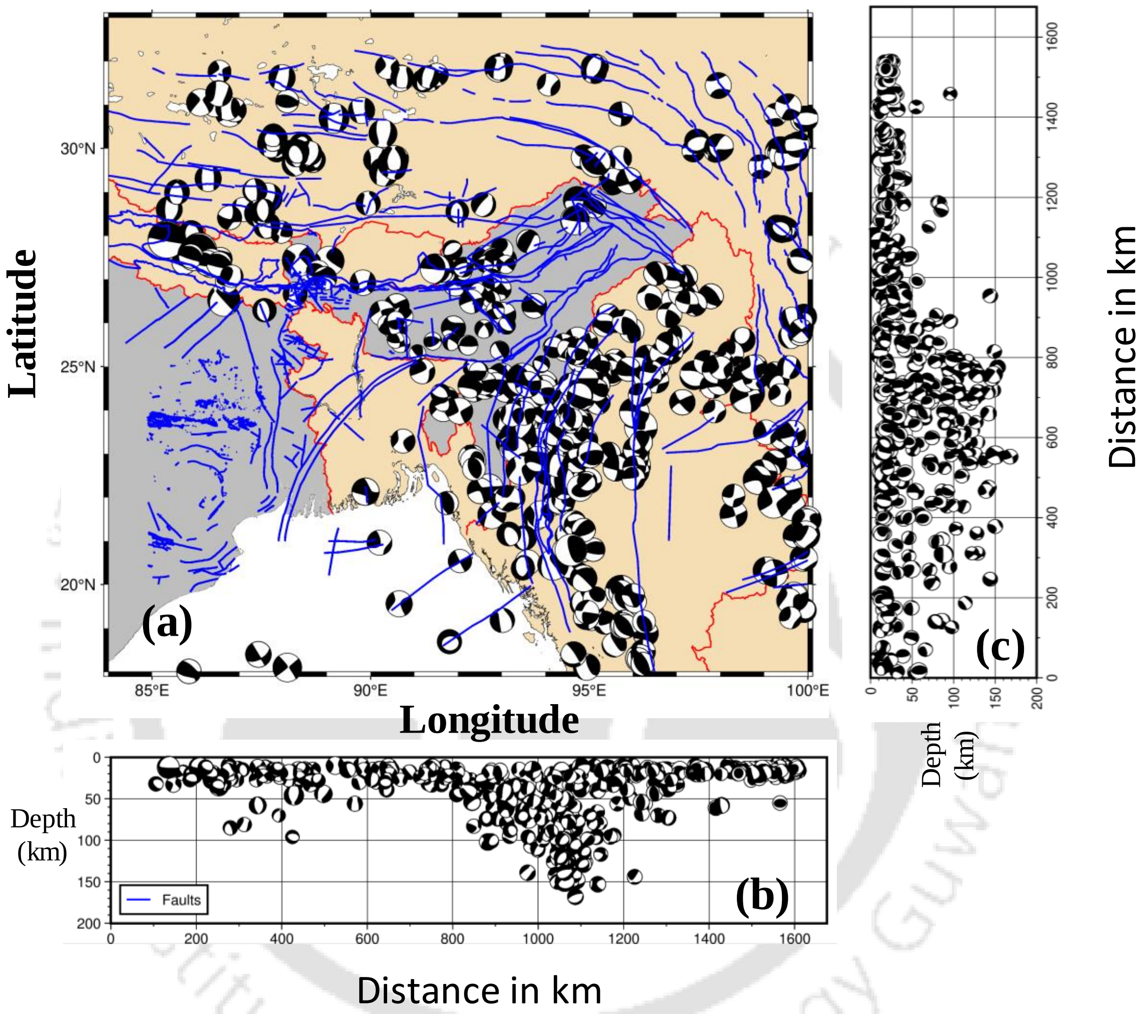

Figure 3.5 Fault mechanism of the Past EQs occurred in the NE India Region (developed by using GMT)

3.4 Fault Mechanism of the Recent EQs in the Region

The focal mechanism of past EQ is important as it can give an idea about the type of faults on which the EQ occurred. The focal mechanisms of the past EQs (from 1976 till 2022) in the region are collected from the GCMT database (Dziewonski et al. 1981; Ekström et al. 2012) and Kayal et al. (2012). Figure 3.5 shows the focal mechanisms as beach ball solutions developed for the NE India region based on the above-collected information. The GMT is used for the development of the beach ball solutions from the focal mechanism information such as strike, dip, rake, moment tensor etc. Figure 3.5 also shows the depth of these EQs. It is observed that only in the Indo-Burma ranges, the focal depths of the EQs are going beyond 50km.

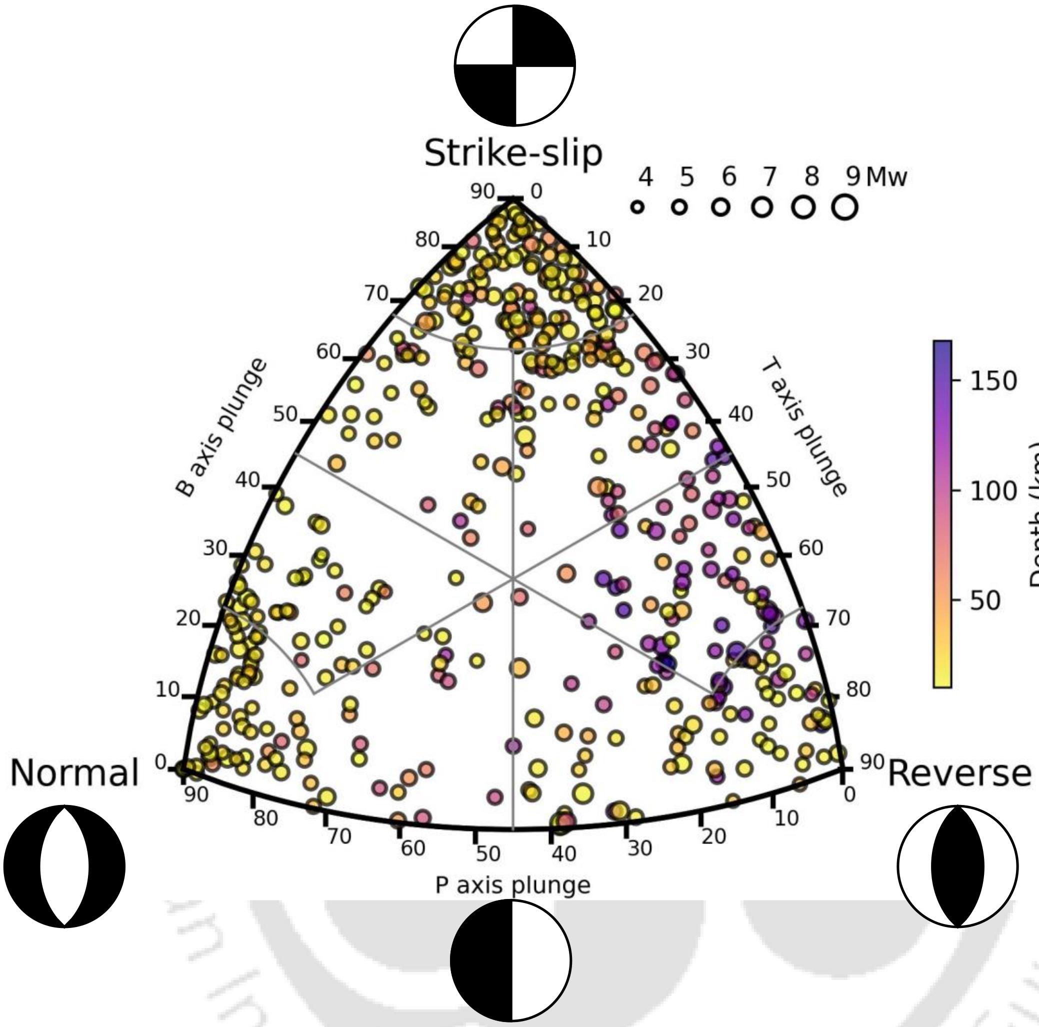

Figure 3.6 Plot showing the classification of fault mechanism of the past EQs in NE India (Developed by using the FMC programme)

Figure 3.6 shows the classification of different types of fault mechanisms for the past EQs collected for the NE India Region. The classification is performed using the FMC program developed by Álvarez-Gómez (2019). This program utilises the focal mechanism information such as strike, dip, rake, moment tensor solution etc., for classifying them into different types. Figure 3.6 shows that all three types of faults, Normal, Strike-slip, and Reverse faults, are present in the NE India region. Most EQs that occurred in the Eastern Himalayan region show Reverse faulting behaviour (Figure 3.5). Some of the EQs, however, exhibit even strike-slip and normal fault behaviour. Normal faulting is visible in the EQs that occurred north of the Eastern Himalayan region. EQs that occurred in the Shillong plateau show strike-slip and reverse faulting. EQs that had occurred in the Indo-Burma ranges exhibit all three types of faulting (Figure 3.5). In addition, majority of the EQs that occurred in the Assam Valley region are strike-slip in nature. Some EQs are also observed to have normal and reverse faulting behaviour (refer to Figure 3.5).

3.5 Study of the Dependent and Independent EQ Events Occurred in NE India

For SHA, it is considered that the EQs occur randomly, and the phenomenon of occurrence follows Poisson's distribution. Major to great EQs are followed by a large number of dependent events (foreshocks and/or aftershocks). Thus, these dependent events must be removed from the catalogue before proceeding with SHA. This process of removing dependent events from an EQ catalogue is called EQ declustering (van Stiphout et al. 2016). There are different methods followed all around the globe for separating dependent events from main events. Knopoff (1964) was the first one to find that after removing aftershocks from the EQ catalogue, the resultant EQ catalogue followed Poissonian distribution (van Stiphout et al. 2016). After this, Gardner and Knopoff (1974) introduced a space-time-distance criterion to remove aftershocks from the EQ catalogue collected for the Southern California region. This method of removing aftershocks from an EQ catalogue was later known as the window method of declustering. The length of space and time windows for declustering is depended on the magnitude of the main events. Further, Uhrhammer (1976) modified these window lengths. Later, many researchers used Uhrhammer (1976) version of windows for different parts of the world, including NE India (Wiemer and Schorlemmer 2007; Sitharam and Sil 2014; Das et al. 2016; Pandey et al. 2017). In this section, the effect of declustering on the EQ catalogue is examined. Further, the effect of declustering in understanding the true potential of seismic sources is also assessed.

3.5.1 EQ Catalogue considered

The EQs that occurred within a longitude range of 85° E to 100 °E and a latitude range of 18°N to 32°N from the catalogue mentioned in Section 3.2 is considered for the study. The total number of events left in the catalogue is 5,472, of which 209 numbers of total EQ events have Mw < 4.0. The total numbers of events having Mw ranging from 4.0 to 4.9, 5.0 to 5.9, 6.0 to 6.9 and 7.0 to 7.9 are 4,177, 839, 210 and 31, respectively. In addition, there are 6 EQ events in the catalogue with Mw ≥ 8.0, with the highest being 1950 Assam EQ (Mw = 8.5).

3.5.2 Declustering of the EQ Catalogue

As mentioned earlier, the window method modified by Uhrhammer (1976) has been used by various researchers for NE India, and thus has been selected for the present work. In Urhammer (1976)'s method, dependent events are identified in space and time windows. The approximate size of the time (t_s) and distance (d_s) windows as per Uhrhammer (1976) are given in Eq. 3.4 (van Stiphout et al. 2016).

$$t_s(in\ days) = e^{-2.87+1.235M} \text{ and } d_s(in\ km) = e^{-1.024+0.804M} \tag{3.4}$$

For each EQ, d_s and t_s are calculated based on their magnitude value M, in Mw scale following Eq. 3.4. If a subsequent EQ has lower magnitude than the previous EQ and falls within the above space and time windows, then such EQ is considered as dependent event and is removed from the catalogue. Similarly, if the subsequent EQ which falls within the space and time windows has a larger magnitude, then the previous EQ is considered as foreshock and time-space windows are recalculated based on the Mw of a larger event (van Stiphout et al. 2016). This way, the largest event is considered as the main event and all other events are removed from the catalogue. After declustering, 4,444 numbers of EQ events remain in the catalogue. In the declustered catalogue, the number of EQ events with Mw < 4 Mw is 190. Similarly, the numbers of events remain in the Mw range of 4.0 to 4.9, 5.0 to 5.9, 6.0 to 6.9, 7.0 to 7.9, and Mw ≥ 8.0 are 3,487, 624, 113, 25 and 5, respectively. Thus, approximately 19% of EQ events are removed from the catalogue after declustering. Table 3.1 shows the percentage of EQ events removed after declustering for different Mw ranges.

Table 3.1 Percentage of EQ events removed for each Mw range after the declustering

Range of Mw	**Percentage of events removed (%)**
Mw < 4	9.09
4 ≤ Mw ≤ 4.9	16.52
5 ≤ Mw ≤ 5.9	25.63
6 ≤ Mw ≤ 6.9	46.19
7 ≤ Mw ≤ 7.9	19.35
8 ≤ Mw	16.67

3.5.3 Completeness analyses of two EQ catalogues

Once a declustered catalogue is available, it must be checked for completeness before determining the seismic activity based on the EQ catalogue. This is done because the catalogue may not contain complete events for all the years and for various ranges of magnitudes. In this study, the completeness with respect to magnitude and time are checked for both the catalogues (catalogues before and after declustering). The minimum magnitude above which a catalogue is complete with respect to magnitude is called the magnitude of completeness (Mc) (Rydelek and Sacks 1989). In this study, Mc is determined based on the maximum curvature (MAXC) method (Wiemer and Wyss 2000). In this method, firstly, the cumulative number of EQ events, which are greater than or equal to a magnitude, is plotted against that magnitude, generating a curve between the two variables. Then, the magnitude corresponding to the maximum curvature point is considered as the Mc. Figure 3.7 shows the cumulative number of EQ versus Mw plots for the catalogue before declustering (Figure 3.7 a) and for the catalogue after declustering (Figure 3.7 b). In both cases, Mc = 4.0 is obtained.

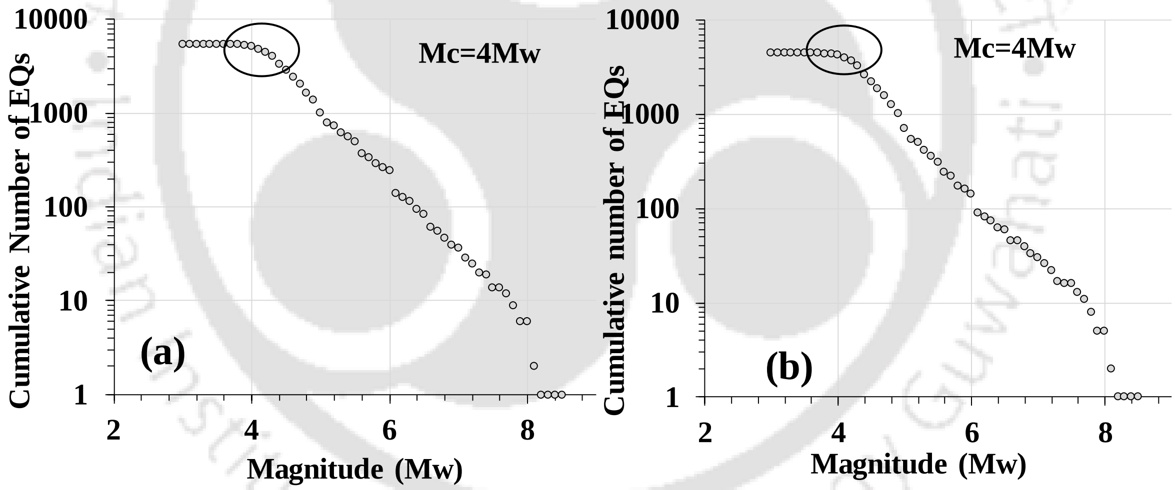

Figure 3.7 Cumulative frequency distribution plots and Mc values for (a) the catalogue before declustering and (b) for the catalogue after declustering

For finding completeness with respect to time, both Stepp (1972) and Visual cumulative (CUVI) (Mulargia and Tinti 1985) methods are used. Firstly, Stepp (1972) method is used, where the EQs are needed to group into certain magnitude classes, and the test is performed on each class separately. The EQ catalogue is divided into time bins from last year up to which the EQ information is available in the catalogue and going back in history with the number of EQs in each time bin. The main assumption of Stepp (1972) method is that the EQ events follow the

Poisson distribution model. If for a definite unit time interval, x_1, x_2, x_3, ….. x_n are the numbers of EQ events, an unbiased mean rate (λ) can be calculated as;

$$\lambda = \frac{1}{n}\sum_{i=1}^{n} x_i \qquad 3.5$$

and for such a dataset, the variance will be;

$$\sigma_\lambda^2 = \frac{\lambda}{n} \qquad 3.6$$

Where n is the number of time bins, if the time interval (T) is considered as one year, then n will be equal to T, where T is the sample length. Then, the standard deviation (σ_λ) will become:

$$\sigma_\lambda = \sqrt{\frac{\lambda}{T}} \qquad 3.7$$

From the study, Stepp (1972) found that for constant λ in each magnitude class, σ_λ behaves as $\frac{1}{\sqrt{T}}$. For the completeness test, the σ_λ is plotted against T, for all the classes of magnitudes and $\frac{1}{\sqrt{T}}$ line is also plotted against T in the same graphs. Stepp (1972) concluded that the time interval up to which the plotted points follow a straight line and parallel to the $\frac{1}{\sqrt{T}}$ line, the data is considered complete for that T.

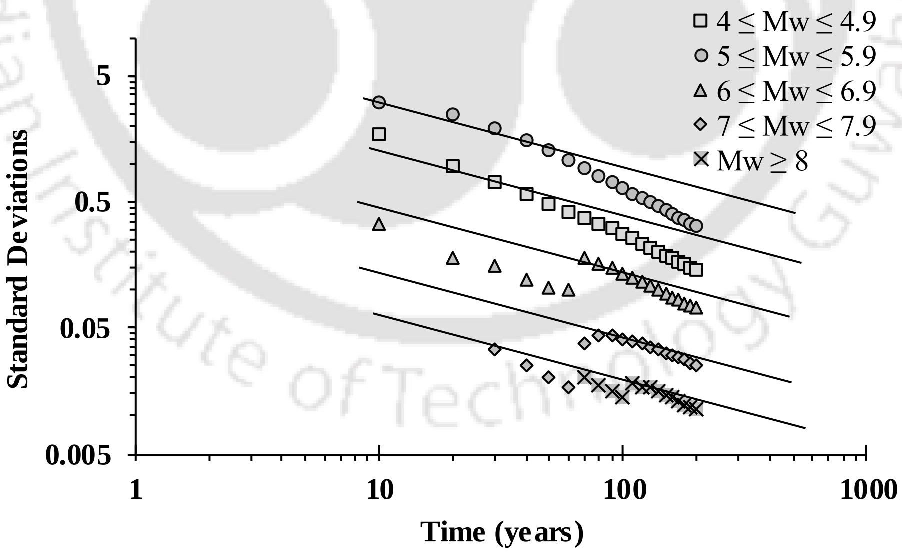

Figure 3.8 Standard deviation versus time interval plots for different magnitude classes for the EQ catalogue before declustering

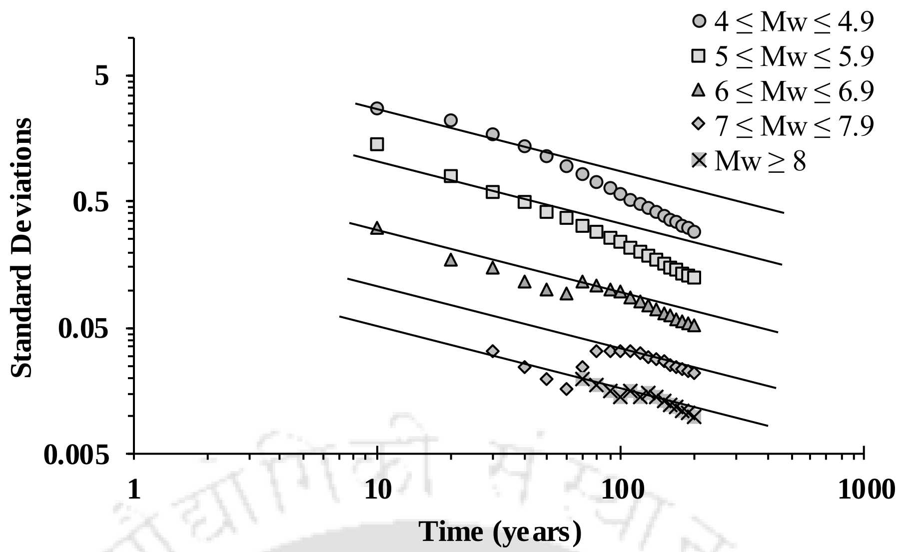

Figure 3.9 Standard deviation versus time interval plots for different magnitude classes catalogue after declustering

In the present study, the EQ events are grouped into five magnitude classes; (1) 4 ≤ Mw ≤ 4.9, (2) 5 ≤ Mw ≤ 5.9, (3) 6 ≤ Mw ≤ 6.9 and (4) 7 ≤ Mw ≤ 7.9 and (5) 8 ≤ Mw. Then, Stepp (1972) method is used for all the five magnitude classes of EQ events. In Figure 3.8, σ_λ versus T for different magnitude classes of the catalogue before declustering are plotted. In Figure 3.9, σ_λ versus T for different magnitude classes of the catalogue (after declustering) are plotted. For each catalogue, year of completeness for different magnitude classes is shown in Table 3.2. It is seen that the completeness periods for the two catalogues are the same for each of the magnitude classes. For magnitude classes, 4.0 ≤ Mw ≤ 4.9, 5.0 ≤ Mw ≤ 5.9, 6.0 ≤ Mw ≤ 6.9, 7.0 ≤ Mw ≤7.9 and 8.0 ≤ Mw, the years of completeness are 60, 70, 130, 190 and 190 years respectively for both the catalogues.

Table 3.2 Years of completeness for different magnitude classes by Stepp (1972) method

Magnitude Classes	**Years of completeness**	
	Catalogue before declustering	**Catalogue after declustering**
4 ≤ Mw ≤ 4.9	60	60
5 ≤ Mw ≤ 5.9	70	70
6 ≤ Mw ≤ 6.9	130	130
7 ≤ Mw ≤7.9	190	190
8 ≤ Mw	190	190

In another attempt for completeness analysis with respect to time, the Visual Cumulative method (CUVI) (Mulargia and Tinti 1985) is used on the catalogue before and after declustering. In the CUVI method, the cumulative number of events per year is plotted against years of occurrence for different magnitude classes. The catalogue is considered complete from the year after which magnitudes are assumed to follow the stationary occurrence process. In addition to completeness analysis, the CUVI method is also used to check the cumulative number of EQs occur per year for both catalogues. Figure 3.10 compares CUVI plots of both catalogues for all the magnitude classes separately. From Figure 3.10, it can be seen that the years of completeness for all the magnitude classes are the same for both catalogues. For $4.0 \leq Mw \leq 4.9$ class, the completeness period found is between 1958 – 2018 (60 years), which is like the duration of completeness obtained based on Stepp (1972) method. For other magnitude classes, however, the durations of completeness obtained from the CUVI method are significantly different from those obtained from Stepp (1972) method. For $5.0 \leq Mw \leq 5.9$, $6.0 \leq Mw \leq 6.9$, $7.0 \leq Mw \leq 7.9$ and $8.0 \leq Mw$ magnitude classes, the completeness duration obtained from CUVI methods are 1922 – 2018 (96 years), 1915 – 2018 (103 years), 1905 – 2018 (113 years) and 1897 – 2018 (121 years) respectively.

Besides finding the periods of completeness, the differences in terms of the cumulative number of EQs per year obtained from both catalogues are also observed in Figure 3.10. It can be seen from Figure 3.10 that for magnitude classes $4 \leq Mw \leq 4.9$ and $8.0 \leq Mw$, the plots of the cumulative number of events per year vs year are the same for both the catalogues. It must be mentioned here that both these magnitude classes contain significantly fewer percentages of dependent events (see Table 3.1). Thus, events for the above two classes remain more or less the same, as can be observed in Figure 3.10. However, for the other three magnitude classes ($5.0 \leq Mw \leq 5.9$, $6.0 \leq Mw \leq 6.9$ and $7.0 \leq Mw \leq 7.9$), differences in terms of the cumulative number of EQs per year for the two catalogues can be observed from Figure 3.10. Further, it can be observed from Figure 3.10 that for these magnitude classes, 1950 Assam EQ generated a larger number of aftershocks. The difference in terms of the cumulative number of EQs per year between the two catalogues can be observed significantly in the year 1950. In other words, a reduction in the cumulative number of EQs after 1950 Assam EQ can be found.

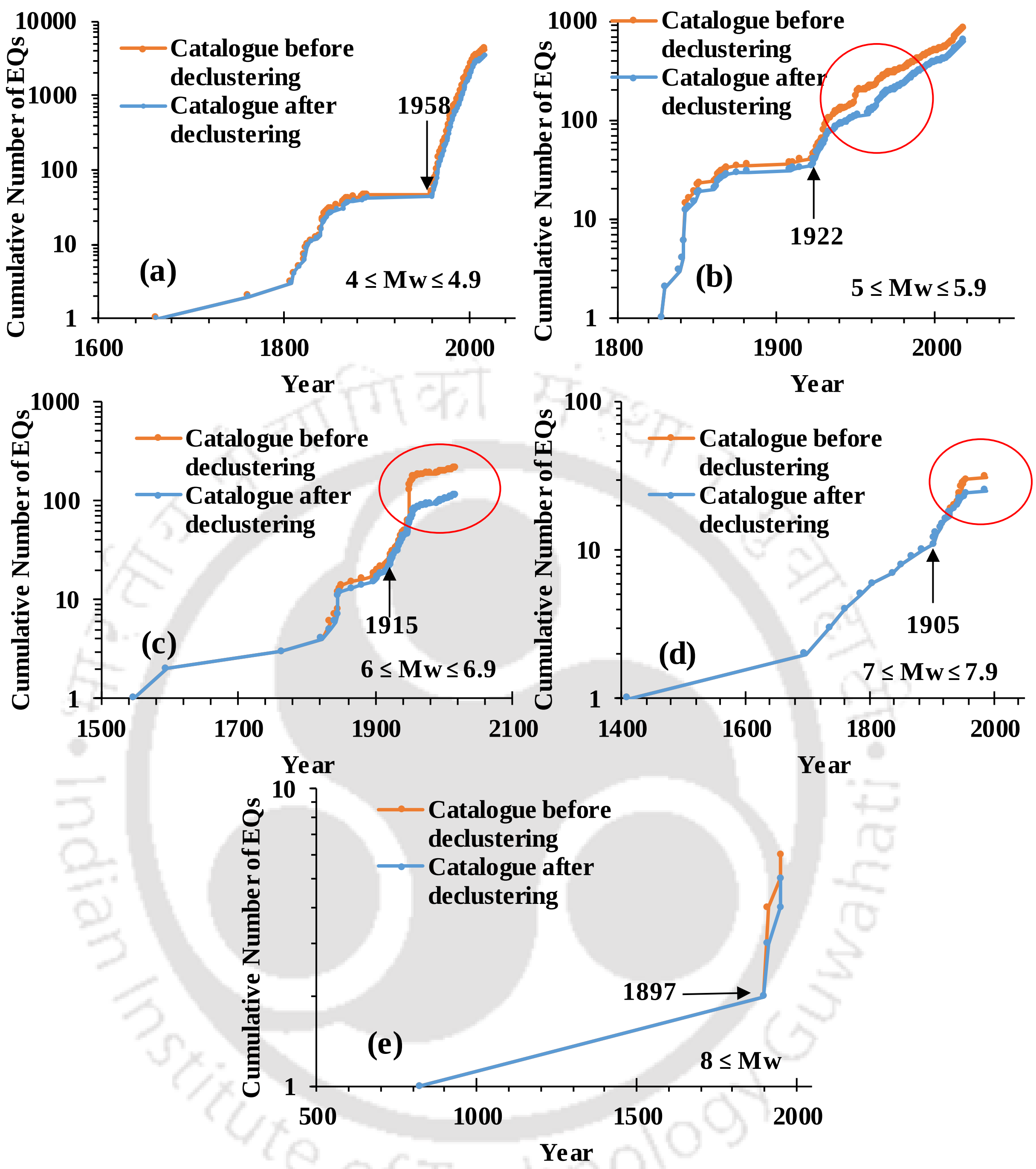

Figure 3.10 CUVI method for determining catalogue completeness (arrow showing the point after which the data are complete with respect to time)

Additionally, the slope of the plot in Figure 3.10 gives the occurrence rate for particular magnitude classes. It can be seen from Figure 3.10 that for the above three magnitude classes, after the 1950 Great Assam EQ, which triggered a large number of aftershocks, the slope is significantly changing for the EQ catalogue obtained after declustering (as shown by the red circles in Figure 3.10). After the 1950 Great Assam EQ, the slope is suddenly reduced, indicating a sudden reduction in the occurrence rate of EQs in the above three magnitude classes. Such reduction in the slope can be attributed to a possibility that many EQs, which in actuality were supposed to occur as independent EQs later, had occurred before the time due to the loading caused by the 1950 Great EQ. Another possibility for the reduction in the slope after 1950 is that during the declustering process, many of the actual independent events were considered dependent events of the 1950 Great EQ since they fell within specific space and time windows. In both cases mentioned above, collectively, a change in the cumulative number of EQs per year is observed because independent events became dependent events. In the later section, more discussion on whether it is just an attribute or actual ground truth to consider such events as independent events will be discussed.

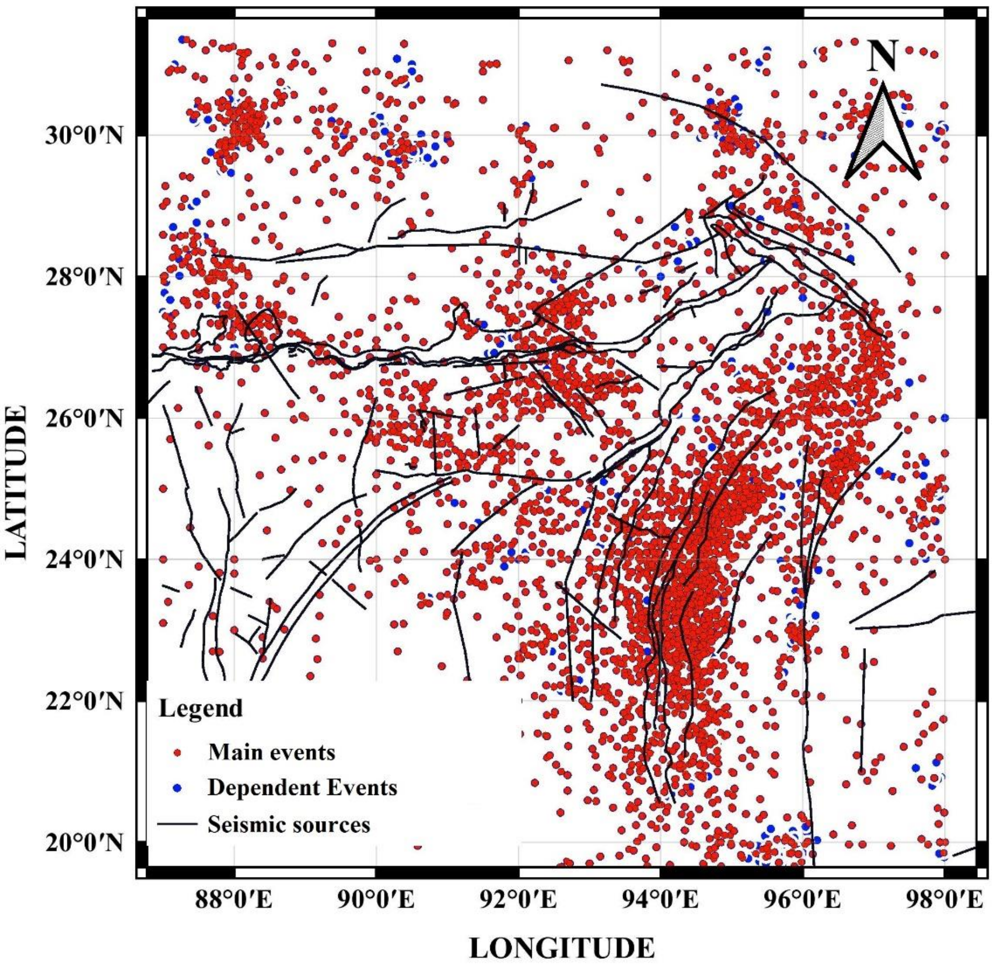

Figure 3.11 Seismic sources in NE India region along with EQ data from both catalogue before declustering and catalogue after declustering

3.5.4 Effect of declustering on the estimation of Mmax

Knowledge of the Mmax on a particular seismic source, which has occurred so far, is very important for SHA. Figure 3.11 shows the seismic sources (total of 121 in numbers) in and around NE India, as well as the past seismic events (Figure 3.1), both main events (red dots) and dependent events (blue dots). Here, main events represent the EQ events from the catalogue after declustering, and a combination of both main events and dependent events represent EQ events from the catalogue before declustering.

Considering both catalogues separately, all the EQ events are assigned to the nearest seismic source. After assigning the EQ events, the Mmax values for all the seismic sources using both catalogues are found separately. For 112 seismic sources (one seismic source with no EQ reported), the Mmax values observed are the same from both catalogues. Further, for 8 other seismic sources, Mmax values obtained from both catalogues are different. In these 8 seismic sources, the total numbers of sources having the differences in Mmax to be 0.1, 0.4, 0.5, 0.8, 0.9 and 2.5 Mw are 2, 1, 2, 1, 1 and 1, respectively. A comparison of both EQ catalogues for these 8 seismic sources shows that Mmax after declustering has become lesser than what was there before declustering. This indicates that the declustering has led to underestimating the Mmax value for a particular seismic source. Observing the locations of the above 8 seismic sources shows closer proximity of many of them with respect to each other. Thus, declustering for a particular seismic source, based on the spatial and temporal window, has considered EQs happening on the adjacent seismic source (having Mmax of lower magnitude than Mmax of the seismic source under consideration) as dependent EQ and removed it. As a result, during declustering Mmax of the adjacent source available will be a different value since the actual Mmax has been removed, as discussed earlier. This scenario of underestimating Mmax of seismic sources based on declustering is possible when seismic sources are close to each other. Thus, in such a scenario, the present work recommends considering Mmax value of each seismic source before declustering. Further, the above analysis also suggests that a larger EQ can also trigger EQs on adjacent seismic sources. As a result, such secondary EQs are found to be dependent events.

3.5.5 Dependent events of major to great EQs

Based on the currently adopted method for declustering (by Uhrhammer 1976), more than 115 numbers of aftershocks for 1950 Assam EQ are obtained. It must also be mentioned here that based on the declustering method, no foreshocks are observed from 1950 Assam EQ. Figure 3.12 shows the location of the main event of 1950 Assam EQ, along with the aftershock's locations. It can be observed from Figure 3.12 that the aftershocks of 1950 Assam EQ are not located on the seismic source responsible for the main event but on other nearby sources. Thus, based on the locations of epicentres and nearby seismic sources, the events which are identified as aftershocks of 1950 Assam EQ might be independent events. Three of such identified aftershocks had Mw > 7, while many had Mw > 6. These so-identified dependent events indicate the seismic activity of nearby sources and can cause significant seismic hazards in the region. As mentioned in section 3.5.3 earlier (Figure 3.10 b-d), 1950 Assam EQ might trigger EQs on other seismic sources before their actual occurrence. Hence, care should be taken while removing high-magnitude EQs from the EQ catalogue; otherwise, the EQ catalogue will underestimate the true potential of many seismic sources while quantifying the seismic hazard of the region.

Similar observations can be made for other EQs that occurred after the 1950 Assam EQ as well. In 1954, 1957 and 1988, three EQs having magnitudes of 7.3 Mw, 7.2 Mw, and 7 Mw occurred in the region. Of these, only 1988 EQ had a significant number (22) of aftershocks. It can also be mentioned here that not all higher-magnitude events cause a significant number of aftershocks. Figure 3.13 shows the main event and aftershocks of the 1988 EQ events. Here, it can be seen that most of the aftershocks occurred within the seismic source. Further, fewer aftershocks occurred on different seismic sources as well. The magnitudes of these aftershock events were not that high (maximum Mw = 5.3). However, this may also contribute considerably to the seismicity of nearby sources.

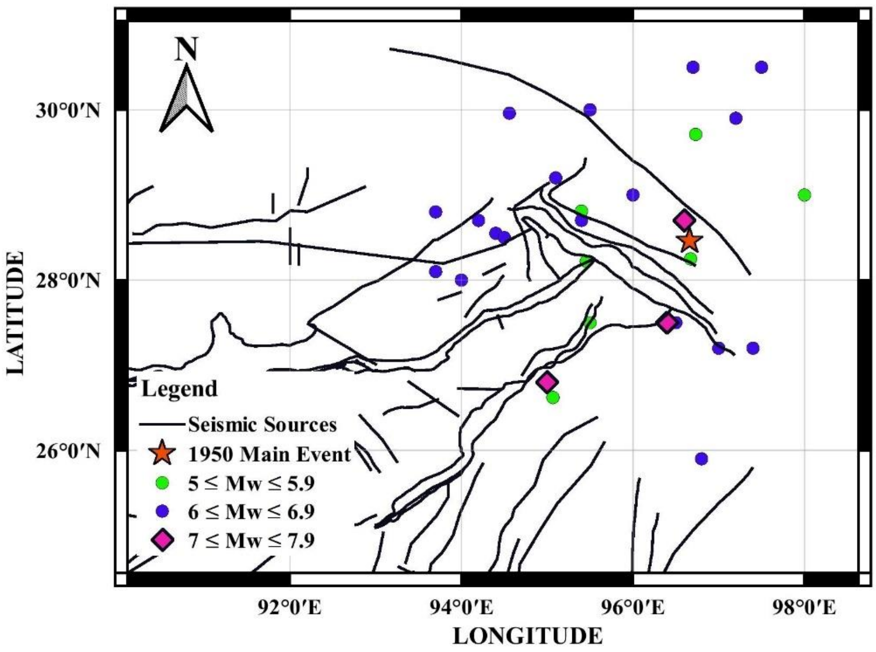

Figure 3.12 The great 1950 Assam EQ and its aftershocks location, along with the seismic sources

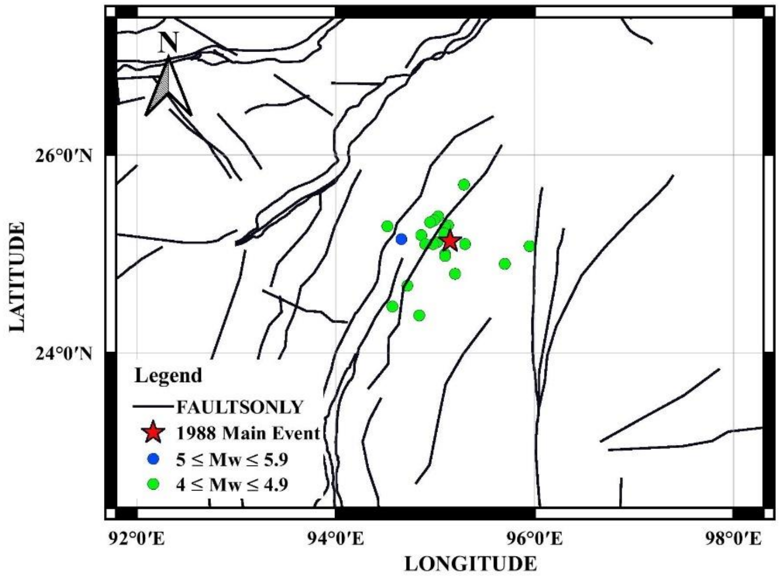

Figure 3.13 1988 EQ and its aftershocks location along with the seismic sources

3.5.6 The overall effect of declustering of EQ catalogue

While preparing the EQ catalogue for SHA, the current practice is to develop an EQ catalogue based on past EQ information, then remove dependent events and determine seismic activity parameters from the complete part of the EQ catalogue. This way, while declustering, a suitable method is applied to the EQ catalogue, and no due consideration is given to the location of nearby seismic sources. In the present work, the EQ catalogue for NE India is initially developed based on collected past EQ information from various sources. When the catalogues before and after declustering are analysed for completeness, both catalogues showed the duration of completeness to be the same for all the magnitude classes. This indicates that the duration of the complete portion of the catalogue is independent of declustering. Further, while examining the cumulative number of EQs per year from both catalogues, a sudden decrease in the number of EQs after 1950 in the declustered catalogue is observed. It must be mentioned here that EQ catalogues for all magnitude classes are found complete after 1950, and thus such a reduction in the cumulative number of EQs per year since 1950 indicates either of the two possibilities. The first possibility is that the 1950 Assam EQ has triggered EQs on other seismic sources, which should have happened later, resulting in fewer EQs in the declustered catalogue. The second possibility is that the events that occurred on different seismic sources but due to larger spatial and distance windows during declustering, were considered dependent events of EQ that happened on nearby sources. Later, it was found that as a result of the second observation, the Mmax of nearby seismic sources has been underestimated after declustering, validated for 8 seismic sources. The first reason mentioned above is also validated for two EQs, namely in the years 1950 and 1988. Whether going with the first or second reason mentioned above, declustering leads to underestimation of actual seismic activity and, subsequently, the seismic hazard level. Thus, in addition to past EQ information, the relative locations of seismic sources should also be given due importance in declustering. Further, the Mmax of each seismic source should be determined before declustering is another important observation from this work.

3.6 Summary

This chapter discusses the EQ catalogue, fault data, and fault mechanism of the past EQ that occurred in the region. Besides, a study is carried out in the region to examine the effect of declustering on the EQ catalogue and determine seismic sources' true potential.

Chapter 4. Seismic Source Zonation for NE India

4.1 General

Researchers perform both DSHA and PSHA for the site/ region to know the worst and most probable ground motion scenario for the structure. While performing such analyses, the seismotectonic province, which can contribute to the seismic hazard of the site/ region, needs to be divided into different seismic zones based on various factors such as geology, tectonic features, rate of movement and direction of tectonic plate movement and many more. According to Vipin and Sitharam (2013), the identification and characterisation of EQ sources play an important role in SHA. As mentioned in Chapter 2, different researchers have considered different sets of parameters to identify seismic source zones which are different from each other, for quantifying the seismic hazard of the NE India region. These source zones for NE India considered for the hazard analyses were mainly based on the tectonics and geological features. While considering tectonics and geological features for source zones, it can be seen that the distribution of EQs is not uniform. But for different kinds of seismological analyses, distribution of EQs is considered as uniform and same seismicity parameters are applied to the entire zones.

In this work, seismic source zones for NE region are tried to identify on the basis of location of Past EQs, so that the distribution in each zone should be more uniform and the zones can be used as areal sources also. Besides this, seismic source zones are also identified based on the spatial distribution of seismicity parameters. Two methods namely; Least square method (LSM) and maximum likelihood method (MLM) are used for finding the spatial distribution of seismicity parameters.

4.2 Clustering Methods

Clustering methods are used to group similar items for different purposes in all fields of science, engineering, medicine, business and many other activities. Besides the visual classification of data by humans, modern computational power can also be utilised to classify items. These classifications can be done in two ways/ methods. In the first method, sample classes or the behaviour of the classes are provided at the beginning of the classification; in the second method, no sample classes are provided at the beginning. The first method is called a supervised classification, and the second one is called an unsupervised classification (Saxena et al. 2017). In the supervised classification method, the classification is done based on initial input information about the classes. Items satisfying a particular class's behaviour are assigned to that particular class. Unsupervised classification is used when the initial class information is unknown. In this method, the items are classified based on similarity and dissimilarity between the items in the data without considering any external input regarding the class behaviour (Saxena et al. 2017).

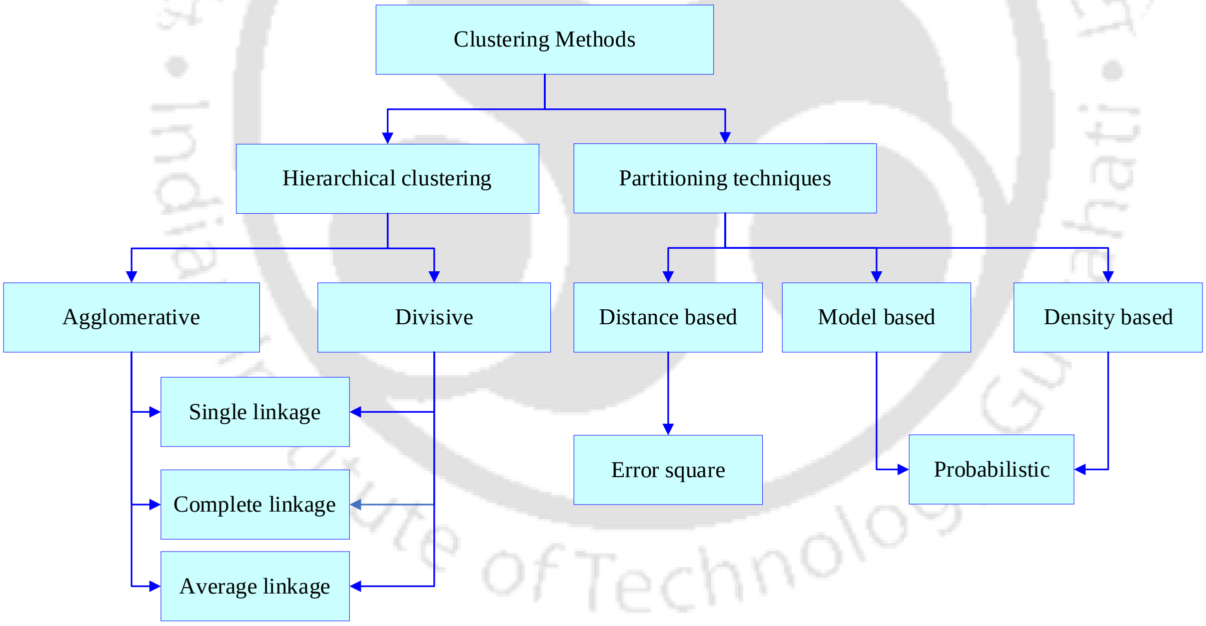

Figure 4.1 Classification of Clustering methods by Fraley and Raftery (1998). (Modified after Saxena et al., 2017)

In present days, there are lots of unsupervised classification or clustering methods available. In each method, ways of obtaining the clusters are different. As a result, clustering methods can be divided into different categories. According to Fraley and Raftery (1998), clustering methods can be divided into two groups, named hierarchical and partitioning techniques. The other sub-categories of the clustering are shown in Figure 4.1.

Many researchers have been using different clustering methods for seismological applications since 1990. Out of all these methodologies, K-means clustering and hierarchical clustering are widely used clustering techniques for seismicity partitioning (Shang et al. 2018). K-means clustering, which falls under partitioning techniques, is a simple iterative algorithm where at first, user-defined k numbers of classes are chosen, and for each class, the initial centroid is defined. Then, the other items are assigned to their nearest centroid class. After assigning all the items to particular classes, the centroids for the classes are recalculated, and the entire procedure is repeated. Such iteration is done till the centroid is not shifted significantly (Saxena et al. 2017). Some of the research on the application of K-means clustering in the field of seismicity are done by Weatherill and Burton (2009), Rehman et al. (2014), Morales-Esteban et al. (2014), Besheli et al. (2015) and Novianti et al. (2017). Besides these, Shang et al. (2017) and Shang et al. (2018) used K-means clustering method for spatial and temporal seismicity analysis of the data obtained for Yongshaba Mine, China. Mainly in these above-mentioned studies, K-means clustering was used for finding seismic clusters and seismic source zones, foreshock classification, interpretation of geological structure, and also for spatial and temporal seismicity analyses.

One of the earliest uses of Hierarchical clustering in seismology can be seen in the work of Frohlich and Davis (1990). In the work, Frohlich and Davis (1990) evaluated the temporal and spatial properties of the EQ catalogues by Single-Link hierarchical clustering analysis. The main idea in Hierarchical clustering is that an item or element is more related to its nearest one than the farthest one. The same idea can be applied to different clusters also. The distance can be calculated mainly in three ways for the linkage between two clusters. These three linkage methods are single-linkage clustering, complete-linkage clustering and average-linkage clustering (Saxena et al. 2017). In the case of single-linkage clustering, the link is made between two clusters by considering the distance between the elements/members of two clusters closest to each other. In the case of complete-linkage clustering, the link is made between two clusters by considering the distance between two elements/ members of two clusters which are farthest from each other. In the case of average-linkage clustering, the link is made between two clusters by considering the average distance between the elements/ members of the two clusters. Some more works on the application of hierarchical clustering in seismology were done by Wardlaw et al. (1990), Davis and Frohlich (1991), Kijko and Funk (1996), Vasak et al. (2004), Hudyma

(2008), Hudyma and Potvin (2010), Hashemi and Mehdizadeh (2015), and Hashemi and Karimi (2016). In the above works, single-linkage, complete-linkage and average-linkage hierarchical clustering were used on both non-spatial variables (magnitudes, occurrence date and depths) and spatial variables (latitude and longitude) for aftershock determination, seismic source identification, rock mass failure mechanism etc.

Besides K-means clustering and Hierarchical clustering, researchers have also used some other clustering techniques such as fuzzy clusters, Self-organizing maps (SOM), Gaussian mixture model (GMM), Density-based clustering algorithm (Density-based spatial clustering of applications with noise or DBSCAN) etc., in seismological studies (Shang et al., 2017, Shang et al., 2018). Benítez et al. (2013) and Ansari et al. (2015) used fuzzy clusters to find the seismogenic zones. Zamani and Hashemi (2004), and Zamani et al. (2011) used the SOM technique for tectonic zoning for Iran. Nanda and Panda (2015) proposed a new fast density-based clustering (DBC) algorithm for clustering of EQ events. Georgoulas et al. (2013) used a density-based clustering scheme that took EQ magnitude into account during the density estimation considering both spatial and time information during cluster. Mukhopadhyay et al. (2010) used a statistical ‘point density’ function for clustering EQ events for the Burmese–Andaman Arc System and West Sunda Arc. Besides the methods mentioned before, Martínez-Álvarez et al. (2015) used a novel triclustering methodology to the Iberian Peninsula for seismogenic zoning.

Researchers have also compared different types of seismic clustering methods discussed above. Leśniak and Isakow (2009) in their work compared mean minimum distance (MMD), K-means clustering and Hierarchical clustering. Konstantaras et al. (2012) used and compared fuzzy C-means clustering, density-based spatial clustering, quantum clustering and dynamic clustering for the vicinity of the Hellenic seismic arc. Shang et al. (2017) compared K-means clustering, GMM and SOM by applying them to the seismicity partitioning of the Yongshaba mine (China).

4.3 Clustering of EQ Catalogue and Seismic Source Zones Identification

In the present work, the Hierarchical clustering technique is used to cluster past EQs based on their latitudes and longitudes. The EQ catalogue developed in Section 3.2 of Chapter 3 is used in the present work. It should be highlighted here that the outcomes of the hierarchical clustering technique are not affected by the initial seeds, but this effect is observed in the case of K-means clustering (Ledolter 2013). For this reason, the Hierarchical clustering technique is used in the present work. This method considers the nearest EQ events in the same cluster. The average link method is chosen for the determination of the distance between two clusters as it is less sensitive to outliers (Hashemi and Karimi, 2016).

One of the most important things in unsupervised clustering is to find the proper number of clusters, as by using the clustering methods, the data can be divided into as many clusters as a user wants. The partition of the clusters should be good partitions while clustering data. A good partition can be obtained by finding the optimum number of clusters of a data set. There are numerous methods available nowadays which can be used to find the optimum number of clusters. Weatherill and Burton (2009) used Krzanowski and Lai (1988) index (KL index) to find the optimum number of clusters for k-means cluster analysis of the delineation of seismic source zones. By testing several methods on artificial data, Weatherill and Burton (2009) found that the KL index was able to identify the optimum number of clusters correctly. The cluster number (K) for which KL is maximum is considered as the optimum cluster number. The KL index can be calculated by using the following formula;

$$KL(K) = \left|\frac{DIFF(K)}{DIFF(K+1)}\right| \quad 4.1$$

Where

$\mathrm{DIFF(K)} = (\mathrm{K}-1)^{\frac{2}{\mathrm{d}}} \times \mathrm{WK}_{\mathrm{K}-1} - \mathrm{K}^{\frac{2}{\mathrm{d}}} \times \mathrm{WK}_{\mathrm{K}}$

d= The dimensions of the data

K=cluster number

$\mathrm{WK}_{\mathrm{K}-1}$ = Pooled within-cluster sum of squares of K-1 partition

WK_{K} = Pooled within-cluster sum of squares of K partition

For finding KL for different K, it is assumed that the minimum optimum number of clusters is greater than 1 (i.e. K>1). Figure 4.2 shows variation in KL for different numbers of clusters (K). From Figure 4.2, it is found that K=12 yields the highest value of KL. Accordingly, the entire study area for which past EQ data was collected in preparing the EQ catalogue is divided into 12 clusters (or zones) (as shown in Figure 4.3). Considering Figure 4.3, the total number of EQs and maximum magnitudes in each cluster are summarised in Table 4.1.

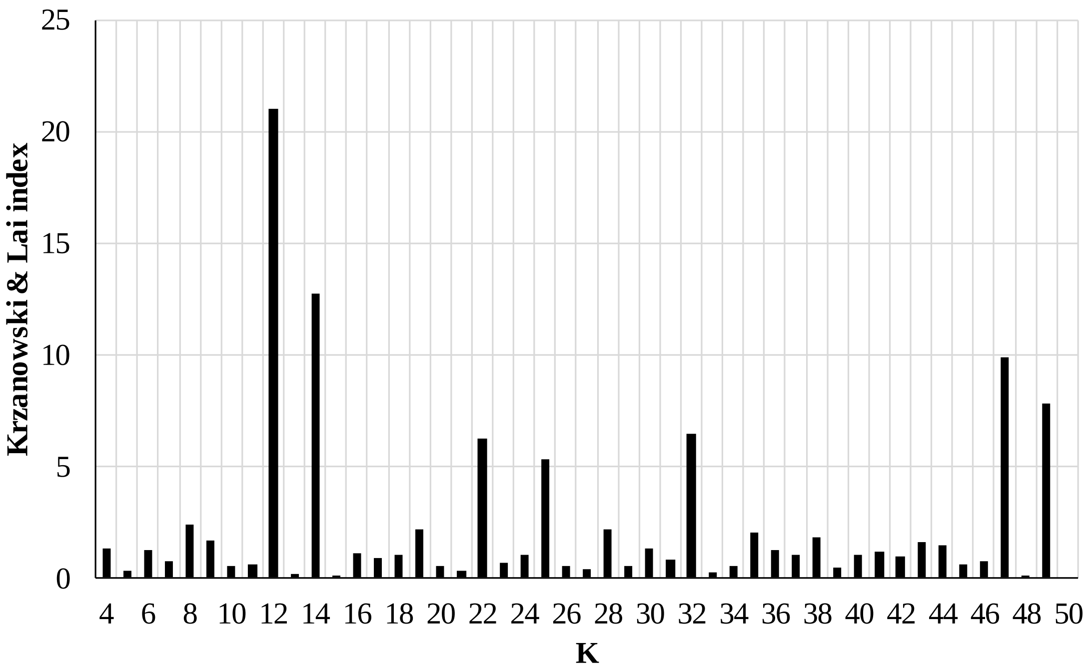

Figure 4.2 Variation in KL index with K for clustering based on past EQ locations

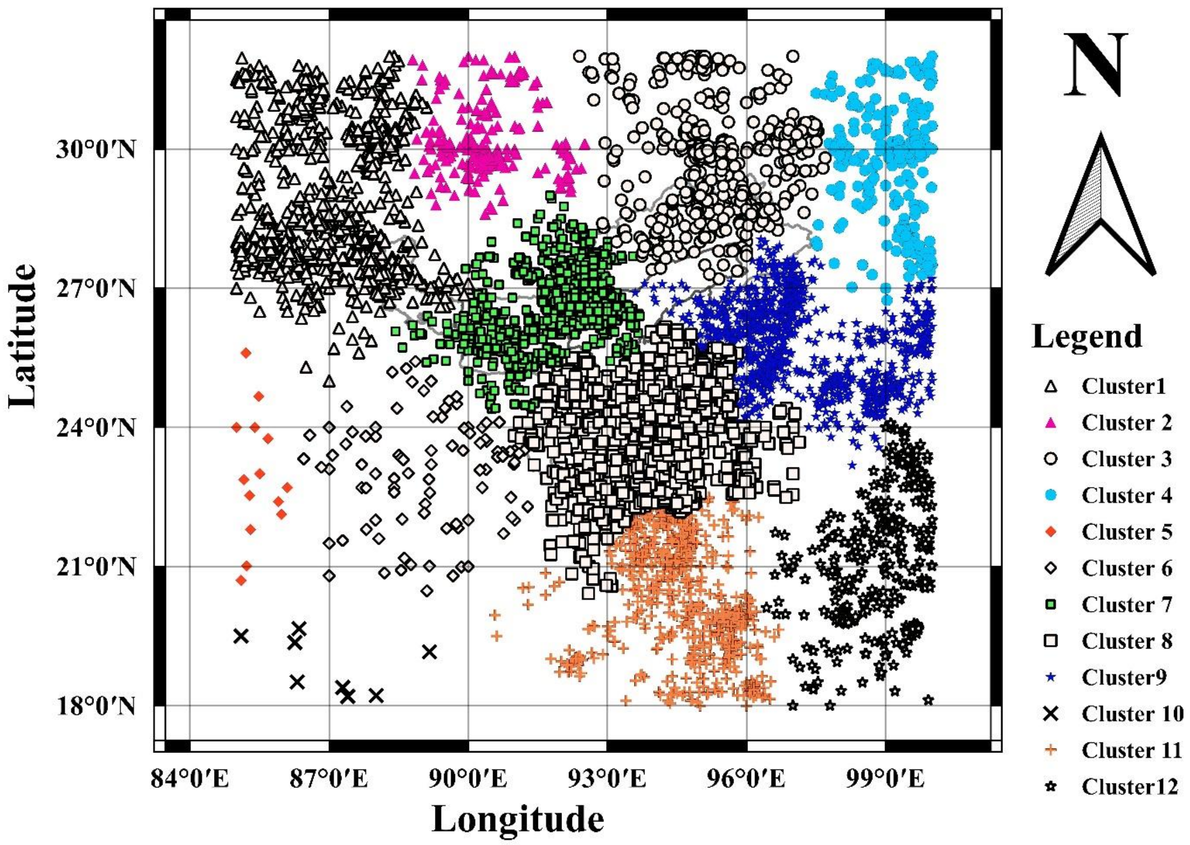

Figure 4.3 Clusters obtained by considering the optimum number of clusters i.e. 12 for clustering based on past EQ locations

It has to be highlighted here that the areas for which zones are identified in Figure 4.3 are much bigger than the NE India region. Figure 4.4 superimposes NE India on the zone map. The zones shown in Figure 4.4 are obtained by dividing the entire region into 0.1°× 0.1° grids and then assigning each grid to the nearest EQ cluster. It can be observed from Figure 4.4 that though the region considered in Figure 4.3 has 12 zones, NE is composed of 9 zones (Zones 1, 2, 3, 4, 6, 7, 8, 9 and 11). Zones 5, 10 and 12 of Figure 4.3 are not a part of the NE India region (NE region shown in Figure 4.4). However, for SHA of the NE India region, the entire 12 zones can be utilised if needed. In the further analyses in this chapter, the above 9 zones of NE, identified in Figure 4.4 will be considered for the discussion.

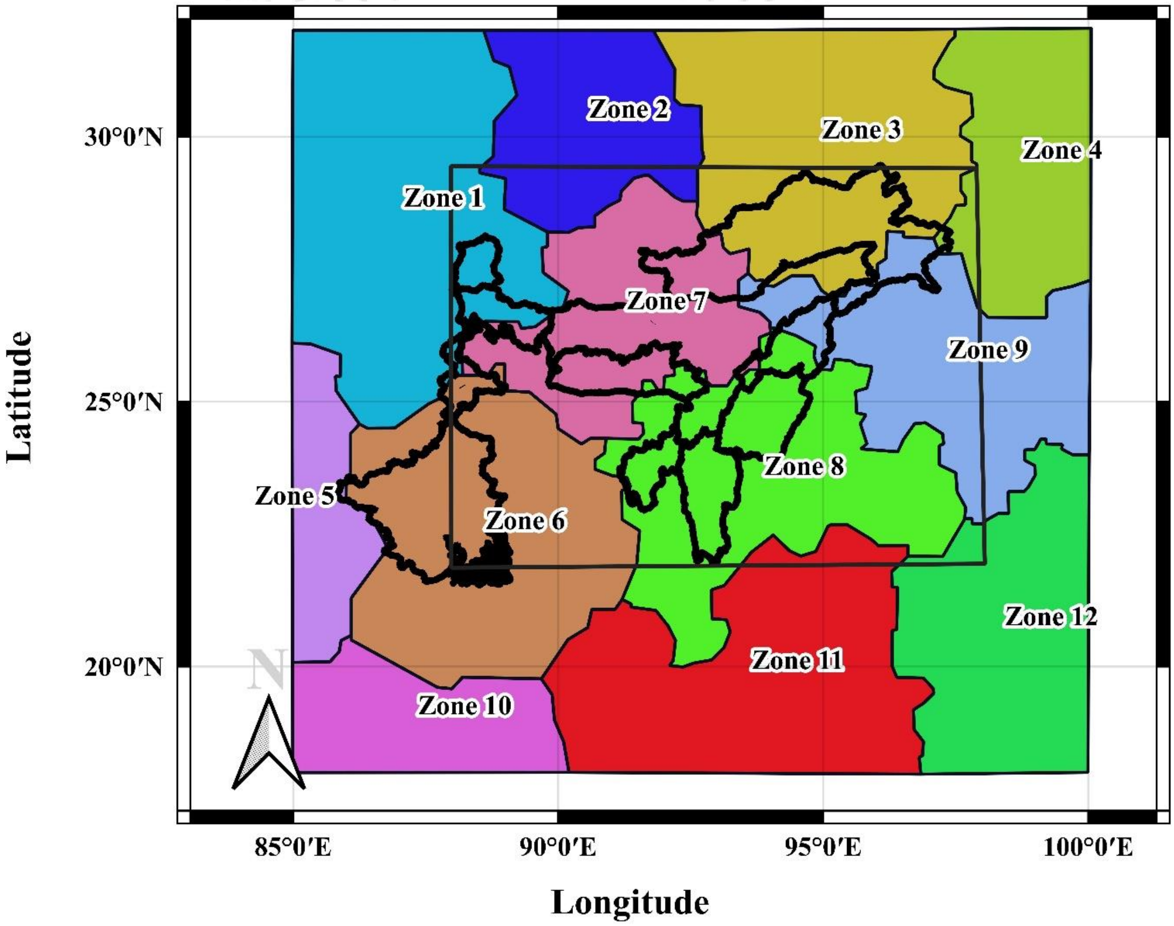

Figure 4.4 12 seismic source zones based on past EQ locations

It can be observed here that Zones 3 and 1 have the highest magnitude of 8.5 (Mw) each. For zone 3, 8.5 (Mw) is the 1950 Assam EQ (Some literature suggests this magnitude is 8.7Mw). Similarly, for zone 1, 8.5 (Mw) EQ occurred in 1100AD somewhere in the Nepal region. Zone 7 comprises Shillong Plateau, Assam Valley and the eastern Himalayan region. Zones 8, 11 and 9 encompass portions of the Indo-Burma Range. Similarly, Zone 6 is located in the Bengal Basin region, Zone 5 is located west of the Bengal Basin region, Zone 1 is in the eastern Himalayas, Zone 3 is in the eastern Himalayan syntaxis and Zone 4 is located in the east of NE India.

On comparing the clusters and zones obtained in this work (Figure 4.3, Figure 4.4) with the source zones considered in earlier studies (Sharma and Malik 2006; Pallav et al. 2012b; Sitharam and Sil 2014; Das et al. 2016; Bahuguna and Sil 2018), a clear difference in terms of location, size and shape from seismotectonic zones can be observed. In addition, it can be observed from Figure 4.3 and Figure 4.4 that the spatial distribution of EQs in each of the 9 zones obtained in the present work is more uniform in comparison to the distribution of EQs within source zones considered in earlier studies. Thus, it can be clearly stated here that the source zones identified in Figure 4.4 are different from the existing literature, which ensures uniform distribution of past EQs.

Table 4.1 Details of the EQ clusters obtained from clustering

Cluster	Total number of EQs	Maximum magnitude of the cluster (Mw)
Cluster 1	721	8.1
Cluster 2	1914	7.8
Cluster 3	15	5
Cluster 4	115	7.8
Cluster 5	276	8
Cluster 6	1242	8.5
Cluster 7	307	7.3
Cluster 8	680	8.5
Cluster 9	1006	7.6
Cluster 10	368	8
Cluster 11	671	7.8
Cluster 12	8	5.86

4.4 Zonation of the NE India Region Based on Seismicity Parameters

The number (N) of EQs having a magnitude equal to or greater than a particular magnitude (M) can be related to M by the Gutenberg-Richter (G-R) recurrence relation as (Gutenberg and Richter 1944);

$$log\, N(M) = a - bM \tag{4.2}$$

Where 'a' and 'b' are the seismicity parameters. These 'a' and 'b' parameters can tell about the seismicity of a particular area. 'a' represents the logarithm of the total number of EQs with $M \geq 0$. The value 'b' shows the relative likelihood of higher and lower magnitude EQs. A lower 'b' value indicates that the number of higher magnitude EQs is more when compared to a higher 'b' value. In other words, it can be said that if the 'b' value for an area is low, stresses are not releasing often and are storing to be released a large magnitude EQs. Thus lower 'b' values are generally associated with areas having high-stress rates (Kalafat and Görgün 2019). Thus, a temporal reduction in the 'b' value can be used to anticipate an impending main EQ event (Görgün 2013). Further, low 'b' values in some areas prior to a major seismic event are also found to be followed by high 'b' values after the rupture (Görgün et al. 2009). High stress is not building up for areas with a high 'b' value as stresses are released continuously as many smaller magnitude EQs (Thingbaijam et al. 2008).

4.4.1 Declustering

Earlier EQ catalogue used for the clustering consisted of the main event, foreshocks and aftershocks. If such a catalogue is used, the obtained seismicity parameters will not be correct. Hence, before proceeding further and determining the seismicity parameters from the EQ catalogue, the dependent events (foreshocks and aftershocks) should be removed by declustering the catalogue. In the present work, Uhrhammer (1976) method is used for declustering the EQ catalogue (As mentioned in Chapter 3).

After declustering the earlier developed homogeneous EQ catalogue, a total of 5,622 numbers of EQ events are found as independent events. The numbers of EQs in the declustered catalogue corresponding to different magnitude ranges are summarised in Table 4.2. The percentage of EQ events removed due to declustering is also summarised in Table 4.2.

4.4.2 Completeness analyses of the declustered EQ catalogue

The declustered EQ catalogue obtained in the previous section is used in the present work. Once a declustered EQ catalogue is obtained, it has to be examined in terms of completeness with respect to magnitude and time before using the same for determining seismicity parameters for the region. In the present work, completeness with respect to magnitude is determined using the MAXC method and Mc (Magnitude of completeness) value for the whole area is determined (followed by the works of Wiemer and Katsumata, 1999; Wiemer and McNutt, 1997; Wyss et al., 2000). Figure 4.5 shows the plot of the cumulative number of EQs greater than or equal to M versus that magnitude value. The maximum curvature point of this frequency-magnitude distribution curve is observed, and the magnitude corresponding to this point is considered as the Mc. Based on Figure 4.5, the Mc value for the present study region is 4.0 (Mw).

Further, in order to find the completeness of the declustered EQ catalogue with respect to time, two methods, namely Stepp (1972) and CUVI methods, are used in this work. The entire catalogue is divided into five ranges of magnitude, namely; 1) 4≤ Mw<5; 2) 5≤Mw<6; 3) 6≤Mw<7; 4) 7≤Mw<8 and 5) Mw>8. First, following Stepp (1972) method (as mentioned in Chapter 3), the standard deviation (σ_λ) is plotted against time interval (T), for all classes of magnitudes (see Figure 4.6). The time interval up to which the plotted points are parallel to the 1/√T line, the data is considered complete for that T (Stepp 1972). In the case of CUVI (described in Chapter 3), which was proposed by Mulargia and Tinti (1985), for different magnitude classes, the cumulative number of events that occurred in a particular year are plotted against the year (as shown in Figure 4.7). It can be seen from Figure 4.7 that for a magnitude class, the occurrence of EQs is stable and significant after a particulate year. Hence, as per Mulargia and Tinti (1985), from that particular year onward, that specific magnitude class is considered complete.

Table 4.2 EQ events left after declustering for different ranges

Range of Mw	Number of main events left in the catalogue	Percentage of events removed (%)
Mw < 4	214	11.2
4 ≤ Mw < 5	4452	21.3
5 ≤ Mw<6	754	31.6
6 ≤ Mw<7	159	40.9
7 ≤ Mw <8	36	14.3
8 ≤ Mw	7	12.5

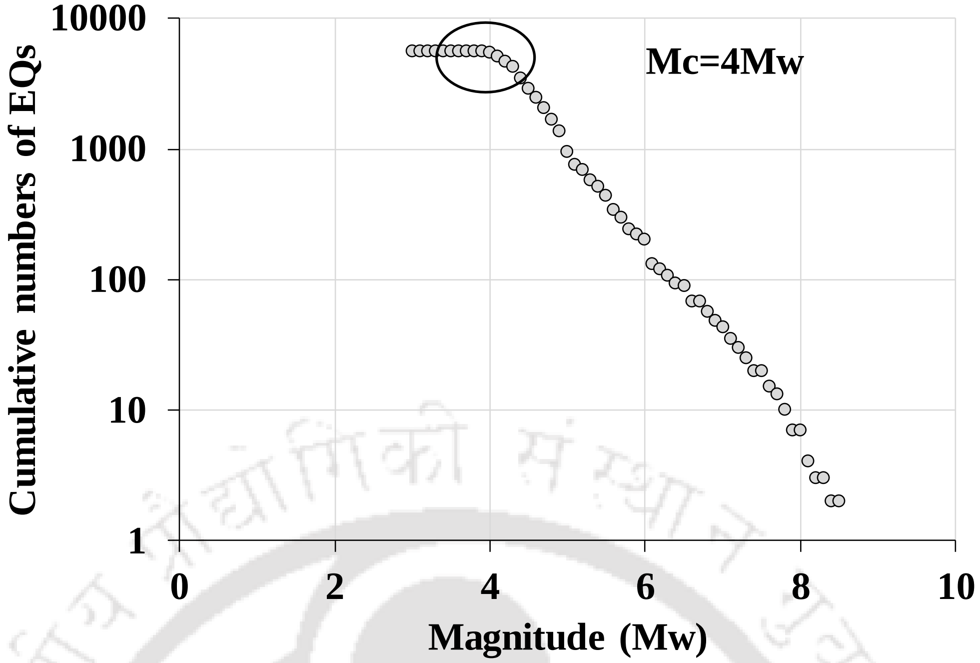

Figure 4.5 Cumulative frequency distribution plots showing the Mc value

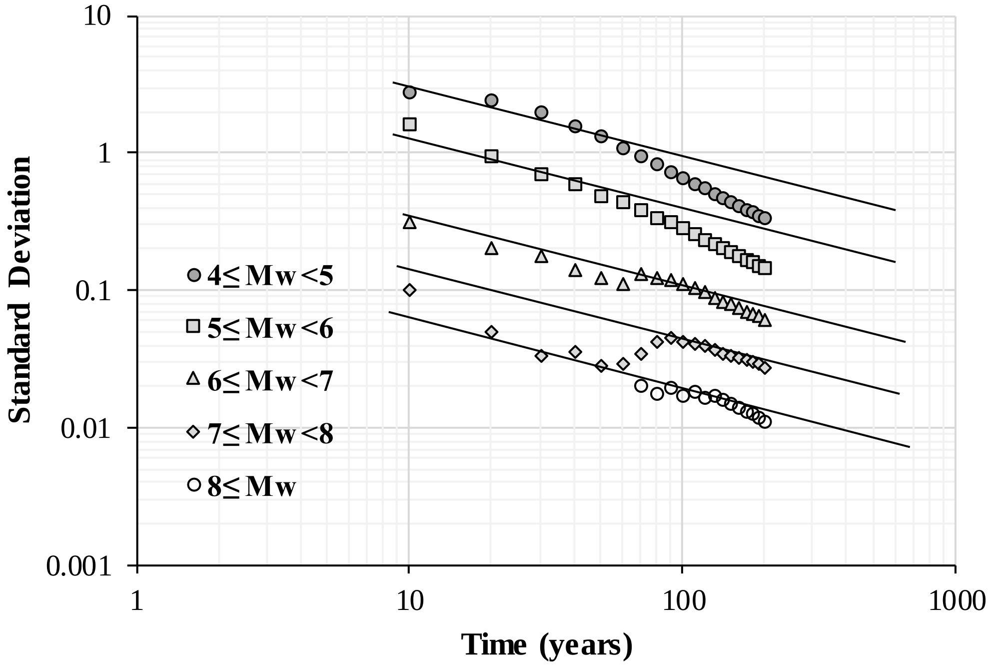

Figure 4.6 Standard deviation versus time interval plots for different magnitude classes

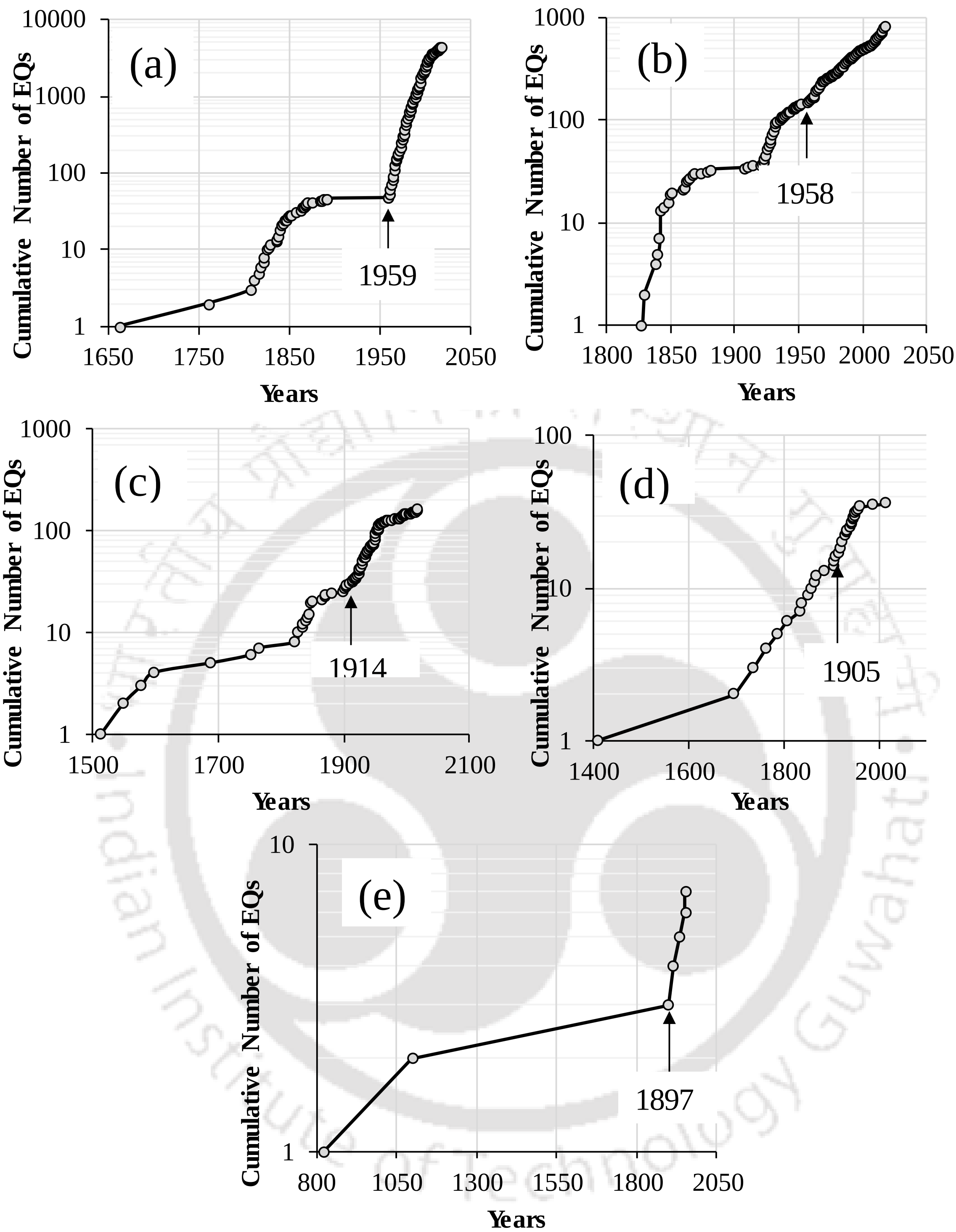

Figure 4.7 Cumulative number of EQs vs time plot magnitude classes (a) 4≤ Mw <5 (b) 5≤ Mw < 6 (c) 6 ≤ Mw < 7 (d) 7≤ Mw <8 (e) 8≤ Mw

Table 4.3 summarises the duration of completeness for each of the five magnitudes classes obtained based on the two methods discussed above. It should be highlighted here that the above two methods of completeness analyses with respect to time are entirely different from each other. The Stepp (1972) method is a statistical method, while CUVI is a graphical method.

The difference can also be seen in terms of handling the EQ catalogue in each of the methods to determine the completeness with respect to time. Stepp (1972) method considers EQ occurrence follows the Poisson distribution model, and those EQs which follow Poissonian distribution are considered as complete. On the other hand, CUVI considers the fact that in present-day technology, the detection and recording of EQs are more efficient when compared to the past. As a result, EQs are considered as complete up to the time these recordings are stable. Two conditions may arise while interpreting the EQ catalogue based on the above-described methods. In one condition, it may happen that as per the CUVI method, the recording is stable from a particular year, but as per Stepp (1972) method, the EQ occurrence is not following Poissonian distribution up to that year. In such a case, the duration of completeness of EQ catalogue obtained by Stepp (1972) method will be lower than that obtained based on the CUVI method. Similarly, another condition may arise in which the occurrence of EQ from a particular year may follow Poissonian distribution while EQ recording is not stable up to that particular year. In such case, the duration of completeness of the EQ catalogue obtained by the CUVI method will be lower than that obtained based on Stepp (1972) method.

For the present study, the duration of completeness of the EQ catalogue is considered up to the year where both stable record and Poissonian distribution are obtained. In other words, the duration of completeness for each of the magnitude classes, obtained based on each of the two methods, are compared, and the minimum duration out of the two methods is taken as the duration of completeness for that magnitude class.

Table 4.3 Completeness duration for different EQ classes

Mw	CUVI method		Stepp method	
	Completeness Period	Completeness duration (years)	Completeness Period	Completeness duration (years)
4≤Mw<5	1959-2018	60	1959-2018	60
5≤Mw<6	1958-2018	61	1959-2018	60
6≤Mw<7	1914-2018	105	1899-2018	120
7≤Mw<8	1905-2018	114	1879-2018	140

4.4.3 Seismicity parameters

Once the completeness of the EQ catalogue with respect to magnitude and time is determined, seismicity parameters are estimated for the region based on two methods. These methods are; the LSM and MLM. Many researchers use LSM methods to fit a straight line to the frequency-magnitude distribution (FMD) plot in order to get the seismicity parameters (Guttorp 1987). The advantages of LSM are that it is simple and easy to apply. As the relation between *M* and the annual rate of exceedance is observed to be linear (Kramer 1996), it is a standard method to fit a linear line to the points by using LSM (Guttorp 1987). According to some researchers (Naylor et al. 2010), as the cumulative EQ data is correlated to all higher magnitude events, regression should never be performed on cumulative data as it assumes that data is not correlated. However, it has also been observed that for estimating the probability of the largest magnitude of EQs, LSM is more suitable (Shi and Bolt 1982). Researchers like Weichart (1980) suggested using MLM for ‘b’ parameter determination. Similarly, as per Naylor et al. (2010), the correct way to fit the data in the G-R recurrence relation is to use MLM. In MLM, data are assumed to be exponentially distributed, and it also considers EQ magnitudes larger than the observed maximum magnitude. In the case of LSM, EQs larger than the maximum magnitude are not considered.

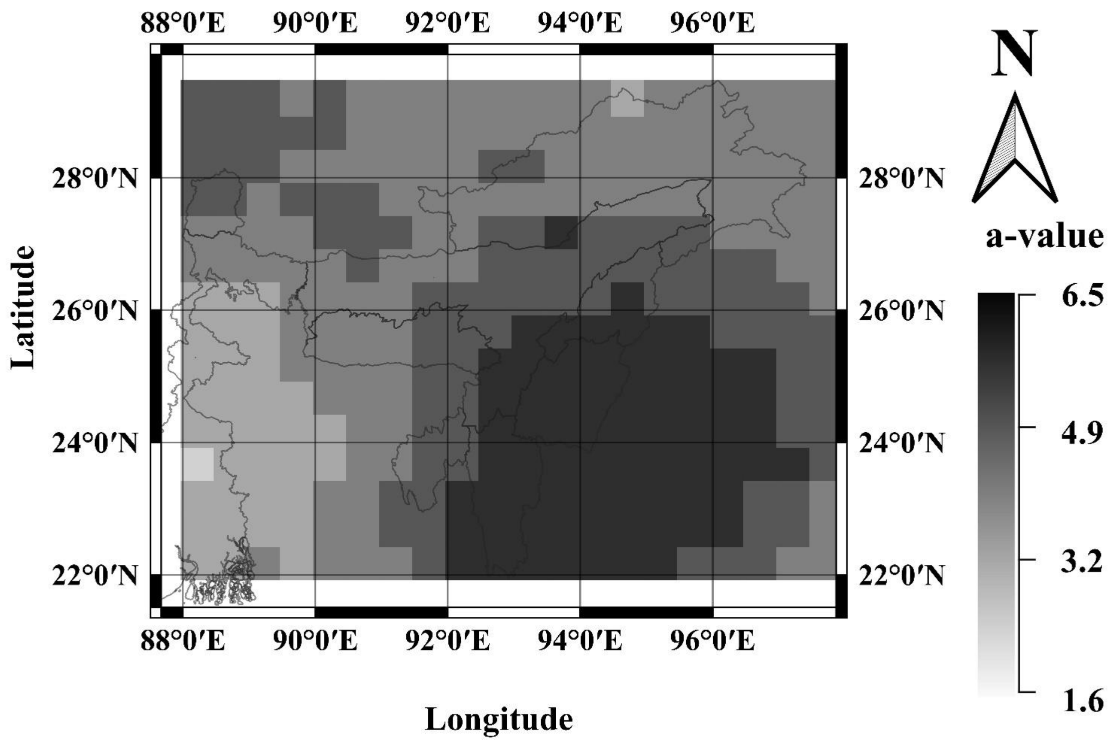

Figure 4.8 Spatial variation of ‘a’ values in and around NE India region obtained from LSM

In this work, both MLM and LSM methods are used for the determination of seismic activity in the region. For LSM, the whole study area is divided into 0.5° × 0.5° size grids. For each grid, EQs which are located within 300km from the centres of the grids, are considered to estimate the 'a' and 'b' parameters using Eq. 4.2. In addition, parameters for each grid are found out by least square linear regression analysis between the annual rates of exceedance vs. magnitude plot for the EQs. Figure 4.8 and Figure 4.9 show the spatial variation of 'a' parameter and 'b' parameter, respectively, for the NE India region based on LSM. It can be observed from Figure 4.8 and Figure 4.9 that 'a' value varies from 2.98-5.73 and 'b' value varies from 0.64-1.06, respectively, within the region.

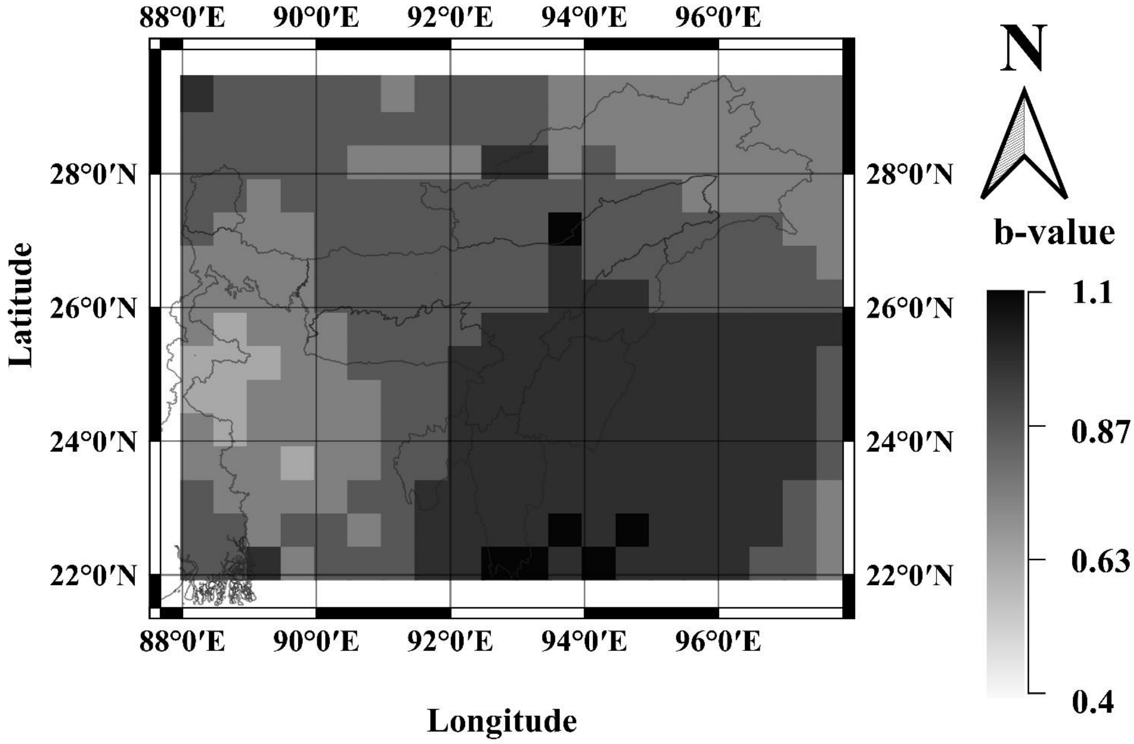

Figure 4.9 Spatial variation of 'b' values in and around NE India region obtained from LSM

While applying MLM (Weichart 1980), the whole study area is divided into 0.1° × 0.1° size grids. Further, for each grid, a 3° × 3° window (following Thingbaijam et al. 2008) is considered to keep the centre of the grid to be the centre of the window. Seismicity parameters are obtained for each grid considering the EQs located within the window. This way, considering the EQ catalogue corresponding to the duration of completeness for each magnitude class, seismicity parameters are determined using MLM. Then the seismicity parameter 'β' (=2.303b) is obtained by solving for the following maximum-likelihood condition Eq. 4.3 (Weichart 1980);

$$\frac{\sum_i t_i m_i e^{-\beta m_i}}{\sum_i t_i e^{-\beta m_i}} = \bar{m} \qquad 4.3$$

Where,

m_i = Value of the central magnitude in i^{th} interval.

t_i = Periods of completeness in years for the i^{th} interval.

$\bar{m}$= Average magnitude of the events in the catalogue.

In the present work, Eq. 4.3 is solved by the numerical iterative scheme based on Newton's method (Weichart 1980). After getting the β value, the total numbers of EQs per year greater than Mc (which is 4.0 in this work) are estimated by using Eq. 4.4 (Weichart 1980) as;

$$N_{Mc} = N_T \frac{\sum_i e^{-\beta m_i}}{\sum_i t_i e^{-\beta m_i}} \qquad 4.4$$

Where, N_T is the total number of EQ events. Once N_{Mc} is known using Eq. 4.4, 'a' parameter can be obtained by solving Eq. 4.5 below;

$$\mathrm{a} = \log N_{Mmin} + \mathrm{b} \times M_{min} \qquad 4.5$$

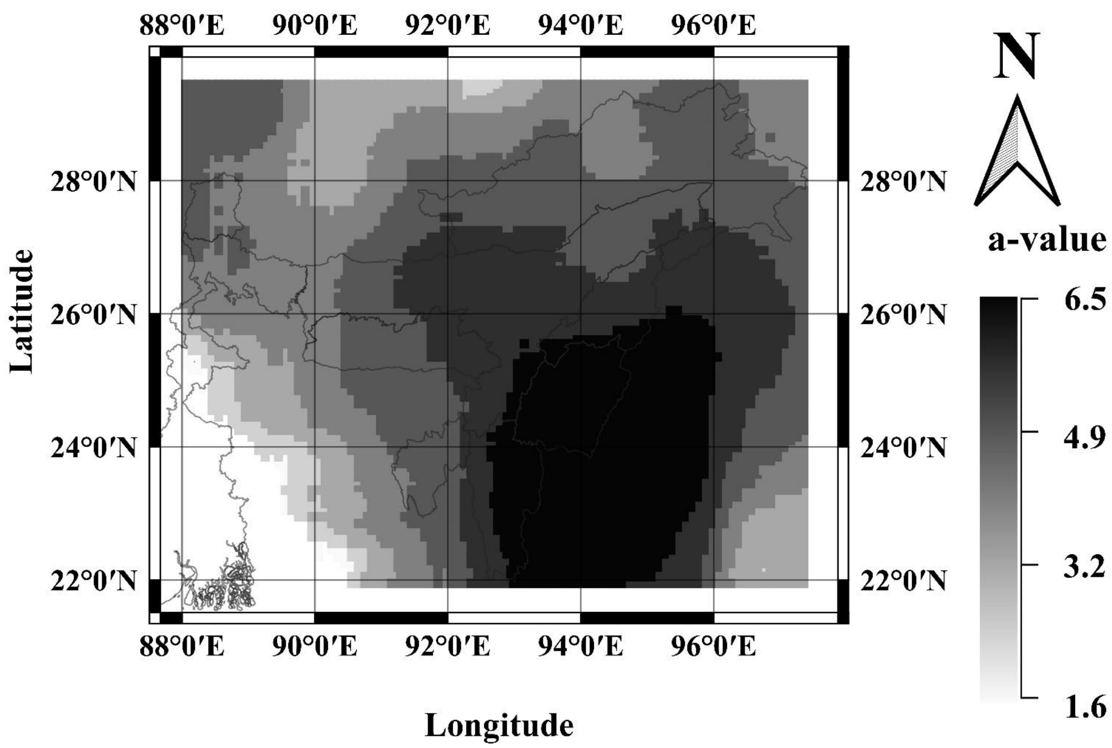

Figure 4.10 Spatial variation of 'a' values in and around NE India region obtained from MLM

The above-discussed methodology is applied to each grid of the study area, and seismicity parameters are determined. Spatial variation of obtained ‘a’ and ‘b’ values are shown in Figure 4.10 and Figure 4.11, respectively. It can be observed from Figure 4.10 and Figure 4.11 that ‘a’ value varies from 1.66 to 6.48 while the ‘b’ value varies from 0.43 to 0.85.

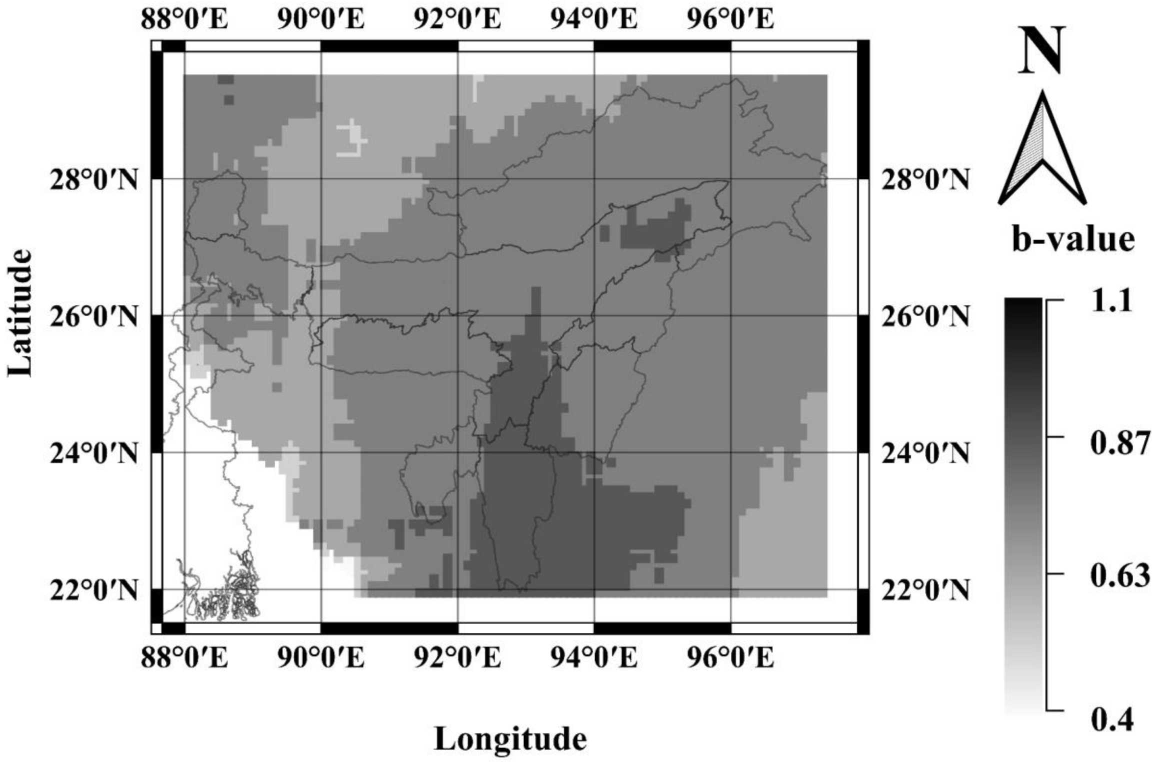

Figure 4.11 Spatial variation of ‘b’ values in and around NE India region obtained from MLM

4.4.4 Source zonation

Based on the obtained seismicity parameters variation for NE from Figure 4.8, Figure 4.9, Figure 4.10 and Figure 4.11, it can be concluded that the entire NE India does not show the same seismic activity throughout. Instead, the region can be divided into separate seismic zones based on the spatial distribution of seismicity parameters. To do so, zonation should be performed while taking into consideration variations in the 'a' and 'b' parameters. However, for such zonation, combined ‘a’ values and ‘b’ values are required. As a result, obtained ‘a’ and ‘b’ values based on LSM (Figure 4.8 and Figure 4.9) and ‘a’ and ‘b’ values based on MLM (Figure 4.10 and Figure 4.11) are used in the hierarchical clustering approach with KL index to identify source zones. Before applying the clustering method, obtained a and b parameters (from each of the two methods) are normalised as per Eq. 4.6 below (Hashemi and Karimi, 2016);

$$x_N = \frac{x-\bar{x}}{\sigma} \qquad 4.6$$

Where,

x = Raw value of a parameter.

$\bar{x}$ = Mean of the value of the variables.

σ = Standard deviation of data of the variables.

x_N = Normalised value of the parameter.

Firstly, normalisation of ‘a’ values and ‘b’ values obtained from LSM is done, and then hierarchical clustering with KL index is used. The optimum number of clusters obtained in this case is 4 (refer to Figure 4.12). These 4 seismic clusters are shown in Figure 4.13. The properties of these clusters are summarised in Table 4.4. In a similar way, considering normalised ‘a’ values and ‘b’ values obtained from MLM, hierarchical clustering with KL index is done. In this case, the optimum number of clusters obtained is 5 (Figure 4.14). These 5 seismic clusters are shown in Figure 4.15. The properties of the clusters are summarised in Table 4.5.

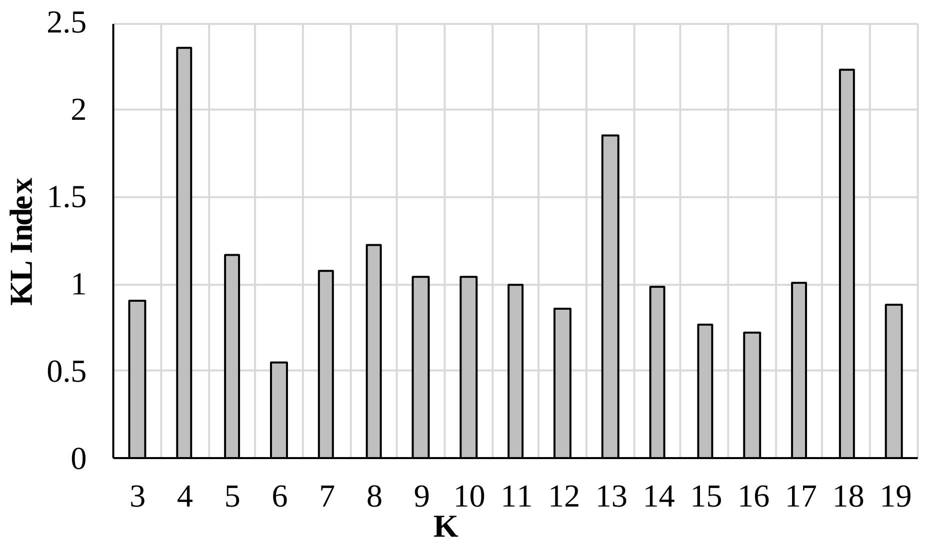

Figure 4.12 Variation in KL index with K for clustering of seismicity obtained from LSM

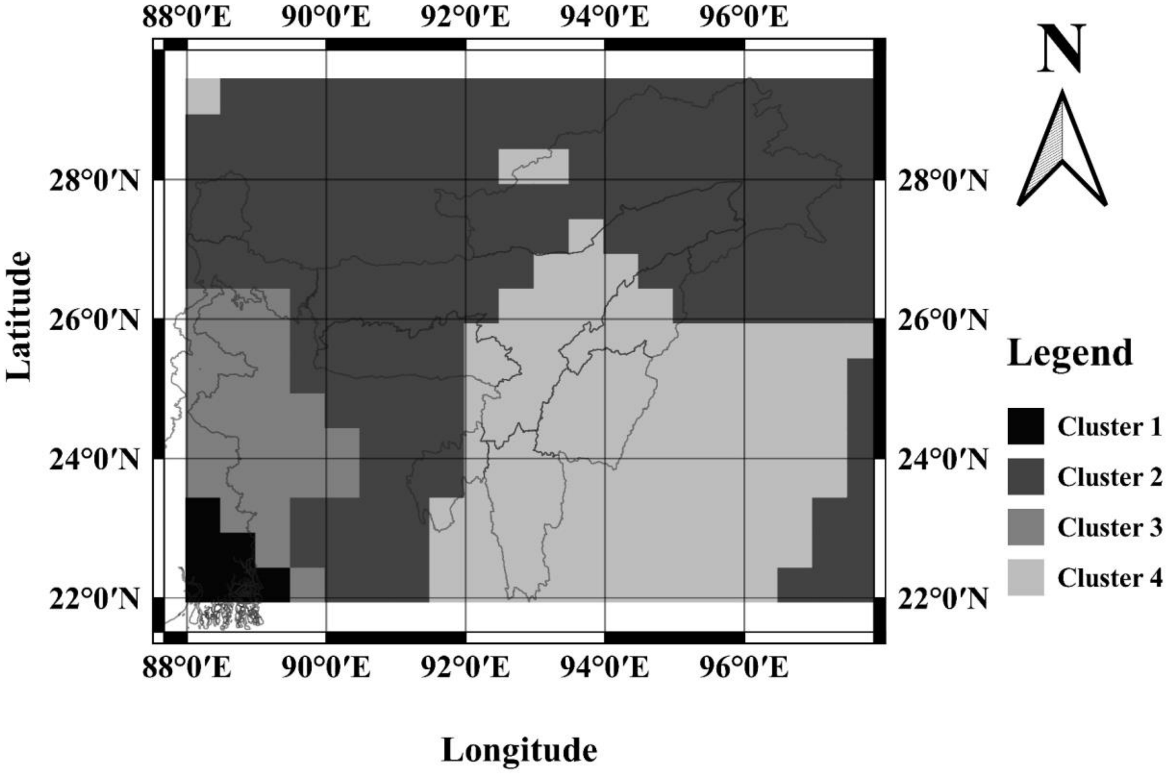

Figure 4.13 Seismic clustering obtained from spatial variation of seismicity parameters based on LSM

Table 4.4 Detail of the clusters obtained from the spatial distribution of seismicity parameter from LSM

Cluster number	%-age of grid within the cluster	mean		Standard deviation	
		b	a	b	a
1	2.0	0.870	3.537	0.027	0.173
2	55.7	0.817	4.232	0.039	0.283
3	9.0	0.718	3.368	0.033	0.198
4	33.3	0.946	5.194	0.035	0.203

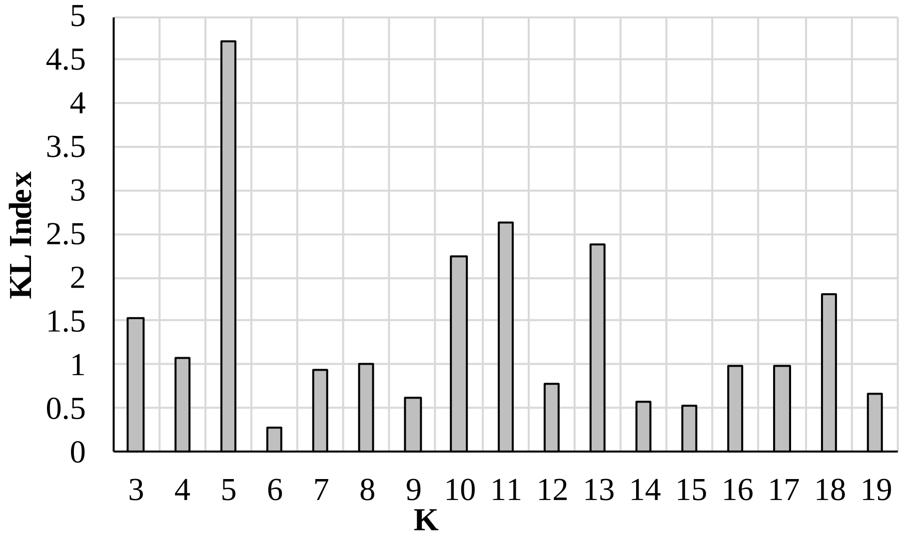

Figure 4.14 Variation in KL index with K for clustering of seismicity obtained from MLM

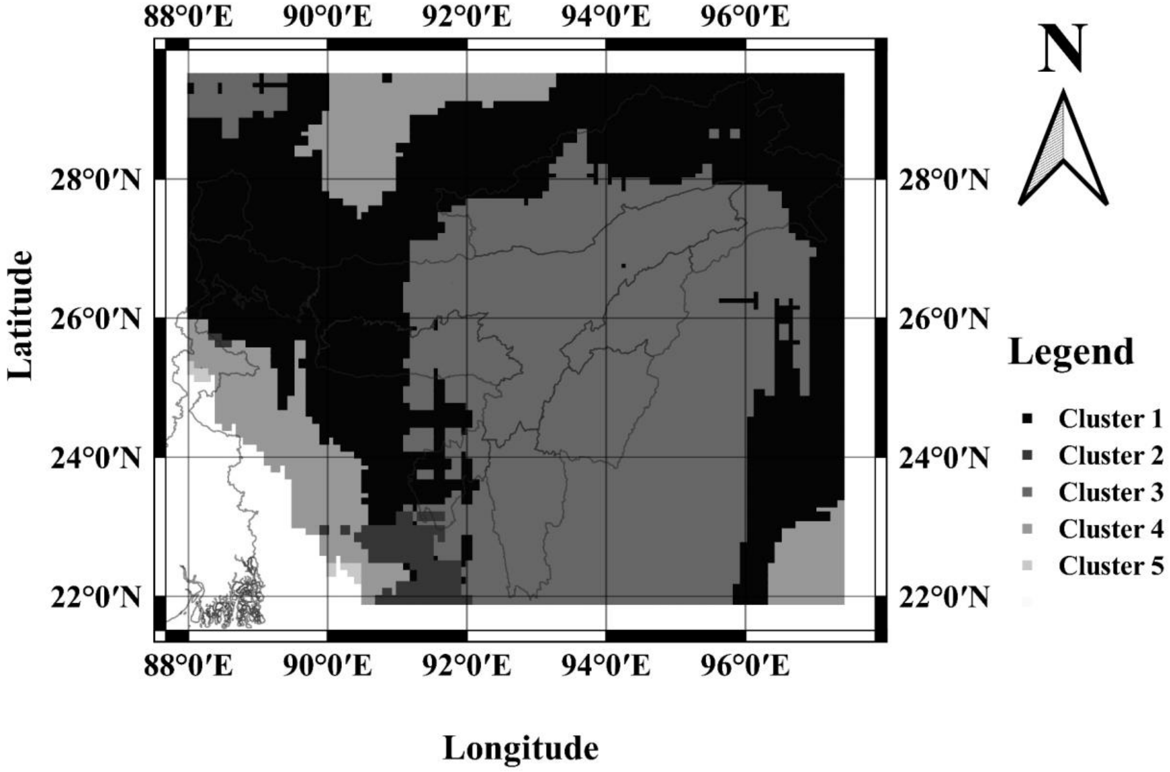

Figure 4.15 Seismic clustering obtained from spatial variation of seismicity parameters obtained from MLM

Table 4.5 Detail of the clusters obtained from spatial distribution of seismicity parameter from MLM

Cluster number	%-age of grid within the cluster	mean		Standard deviation	
		b	a	b	a
1	41.82	0.72	4.42	0.03	0.39
2	1.76	0.78	4.00	0.03	0.38
3	43.65	0.78	5.57	0.03	0.53
4	12.56	0.64	3.22	0.03	0.36
5	0.21	0.49	1.91	0.03	0.12

4.5 Source Zonation of NE India Collectively Based on Seismicity Parameters and Past EQ Locations

In the first attempt, source zones based on past EQ locations are done in this work, as discussed earlier. Based on past EQs, a total 9 out of 12 zones are found in the NE India region (refer to Figure 4.4). Later, based on seismicity parameters ('a' and 'b' collectively), two source zonation maps are found based on LSM and MLM each. Two maps based on EQ data and seismicity parameters (based on LSM and MLM) are developed in this section. For the first map, a source zonation map based on past EQ (refer to Figure 4.4) and seismicity parameters based on LSM (refer to Figure 4.13) are used, and a source map is developed, as shown in Figure 4.16. Similarly, another source zonation map based on past EQ (refer to Figure 4.4) and seismicity parameter based on MLM (refer to Figure 4.15) are used in the hierarchical clustering approach (discussed above), and a source map is developed as shown in Figure 4.17.

It can be observed from Figure 4.16 and Figure 4.17 that in both maps, the entire NE region is divided into 14 source zones. However, the demarcations of many source zones in both maps are different. This difference in the demarcation of source zones in the two maps can be attributed to entirely different approaches followed in LSM and MLM while estimating seismicity parameters. As a result, a final map of source zonation of the NE India region is developed based on visual identification and comparing 14 zones shown in Figure 4.16 and Figure 4.17 is done as shown in Figure 4.18. It can be observed from Figure 4.18 that the entire NE India region has been divided into 11 zones based on the current study.

Visual perception cannot guarantee accuracy. Hence, taking only visible account for the analysis may incorporate significant bias in the study result. However, as it was unavoidable in this study area, extensive calculations and estimation of seismicity parameters were carried out prior to the visual identification of final source zones. Figure 4.16 and Figure 4.17 are already developed based on clustering techniques; thus, zonation is not based on individual assessment alone. In case of source zonation maps based on LSM or MLM are required, the user can refer to Figure 4.16 and Figure 4.17. However, for general understanding, the source zonation map shown in Figure 4.18 is proposed from the present work.

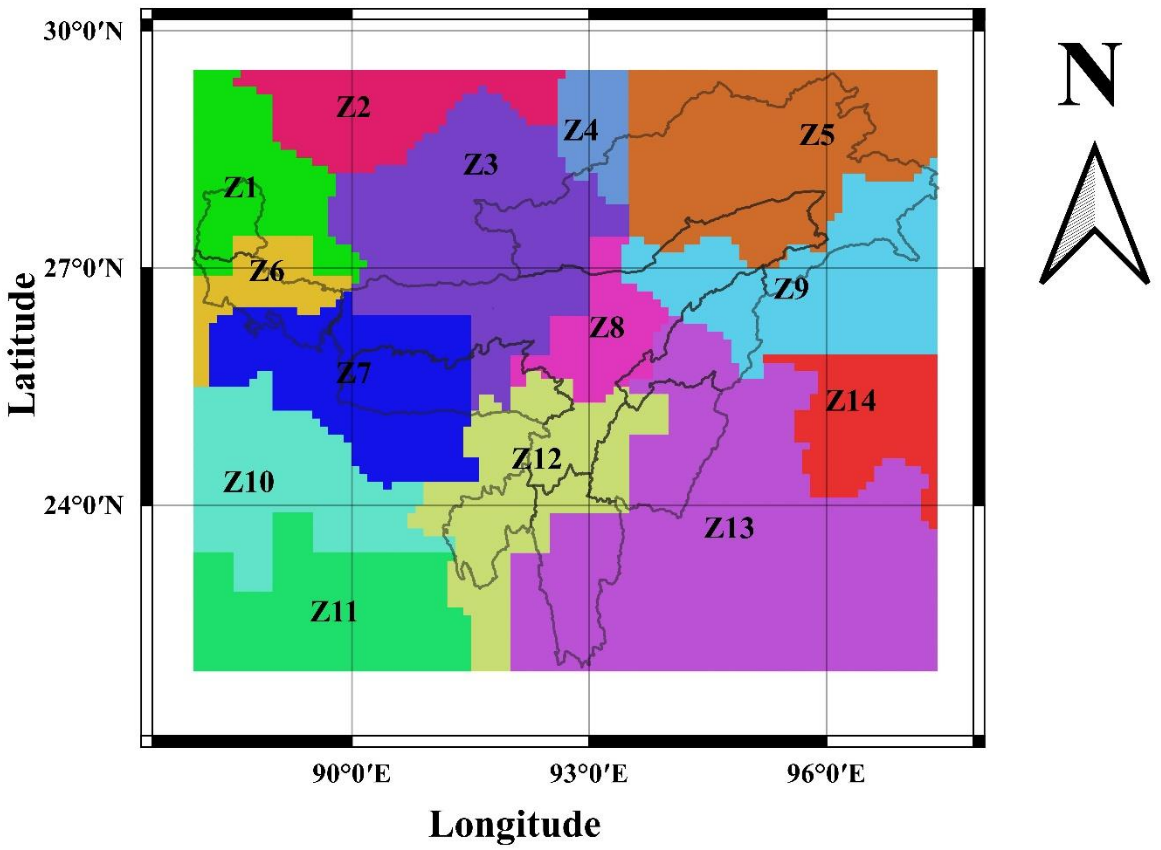

Figure 4.16 Zones based on the combination of past EQ location and spatial distribution of seismicity parameter obtained from LSM

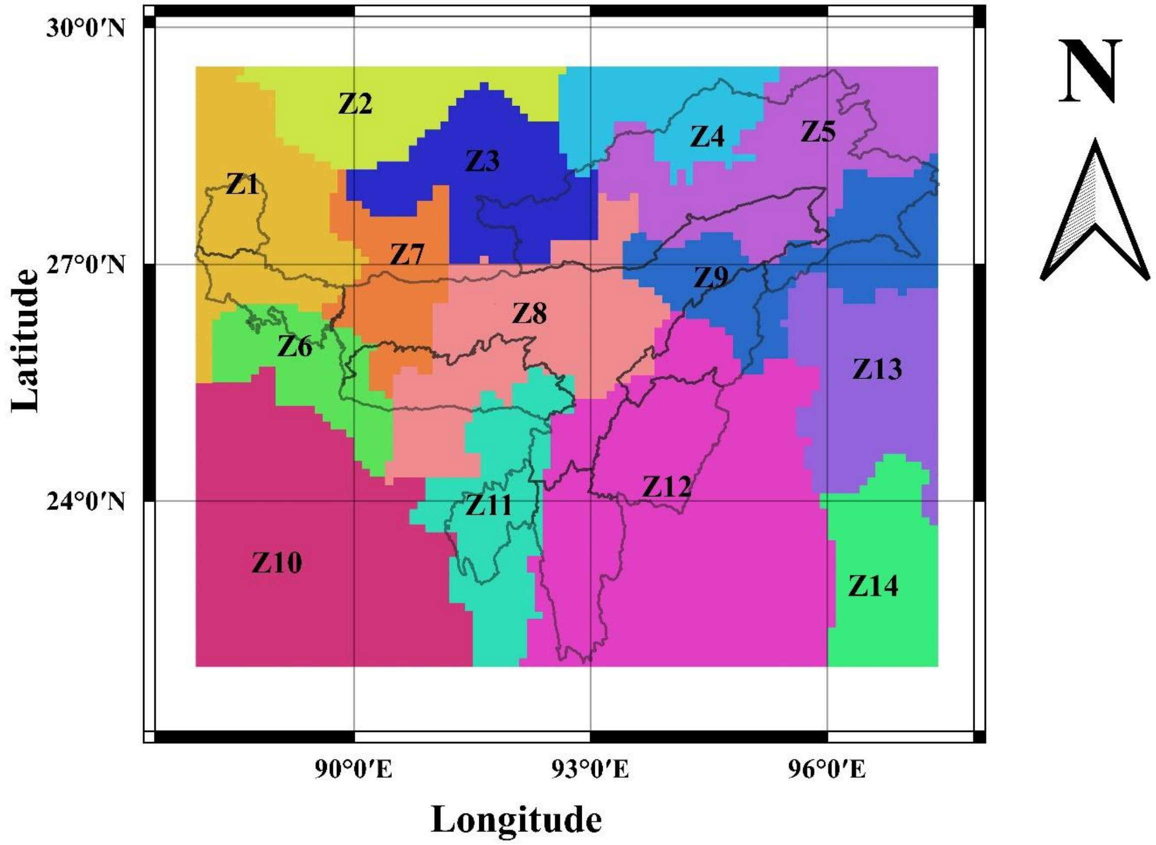

Figure 4.17 Zones based on combination of past EQ location and spatial distribution of seismicity parameter obtained from MLM

4.6 Fault Geometry and Geology in Each Source Zone

The source zonation map developed in Figure 4.18 is solely based on EQ information in the region. As per Sawires et al. (2016), geology and fault alignments both have crucial information for understanding the current active faults as well as identifying the dominant tectonic and structural trends in a region. Figure 4.18 shows identified source zones (Z1 to Z11) overlying the shaded relief map of NE India and the faults present in the region. The shaded relief map of NE India is developed collectively from SRTM based Digital Elevation Model (DEM) collected from Jarvis et al. (2008) and faults information obtained from Bhukosh, Chen (2020), and Kundu and Gahalaut (2013) (Chapter 3). Figure 4.18 shows that zones obtained in the present study, Z1, Z2, Z3, and Z4, are different parts of the Eastern Himalayan region. Z1, Z3, and Z4 each contain four significant thrusts, namely MCT, MBT, MFT, and IST. Similarly, Z2 contains IST. Besides these, Z1 has Chalsa Thrust, Thaljhora Thrust, Matiali Thrust, and Gorubathan Fault, while Z3 contains Ultapani Fault. Coming to the Z4, this zone contains some of the important faults such as Mishmi Thrust, Tidding Suture, Lohit Thrust, Walong Thrust, Bame-Tuting fault, and Miri Thrust. Z5, situated to the west of the Shillong Plateau, contains the Dhubri Fault, Chedrang Fault, and Katihar-Nailphama Fault. Z6, which covers the Shillong Plateau, the Mikir Hill, and a part of the Assam Valley, has some significant faults such as Barapani shear, Kulsi fault, Dudhnoi fault, Oldham fault, Atherkhet fault, Dhansiri-Kopili fault, Dighalpani-Kakijan fault, Kalyani shear, and Dauki fault. Enclosing a part of the Assam Valley and a part of the Indo-Burma Ranges, Z7 has a segment of Naga Thrust, Mishmi Thrust, Tidding Suture, and Lohit Thrust. Z8 zone is entirely in the Bengal Basin region containing Eocene Hinge Zone, Debagram-Bogra fault, Gaibandha fault, Jangipur fault, Malda Kishanganj fault, and Jangipur fault. Z9, which is situated east of Z8, contains a part of the Bengal Basin and a part of the Indo-Burma Range. The important faults present in Z9 are the Sylhet fault and a segment of the Dauki fault. Z10 entirely consists of the Indo-Burma Ranges where CMF, a segment of Naga Thrust, a segment of Dauki Wanding fault, Nanting fault, and Mat fault are present. To the east of Z10 is Z11, where the Shan-Sagaing fault is present. Based on the above discussion, it can be said here that the governing fault geometry and geology of each of the identified source zones are different from each other.

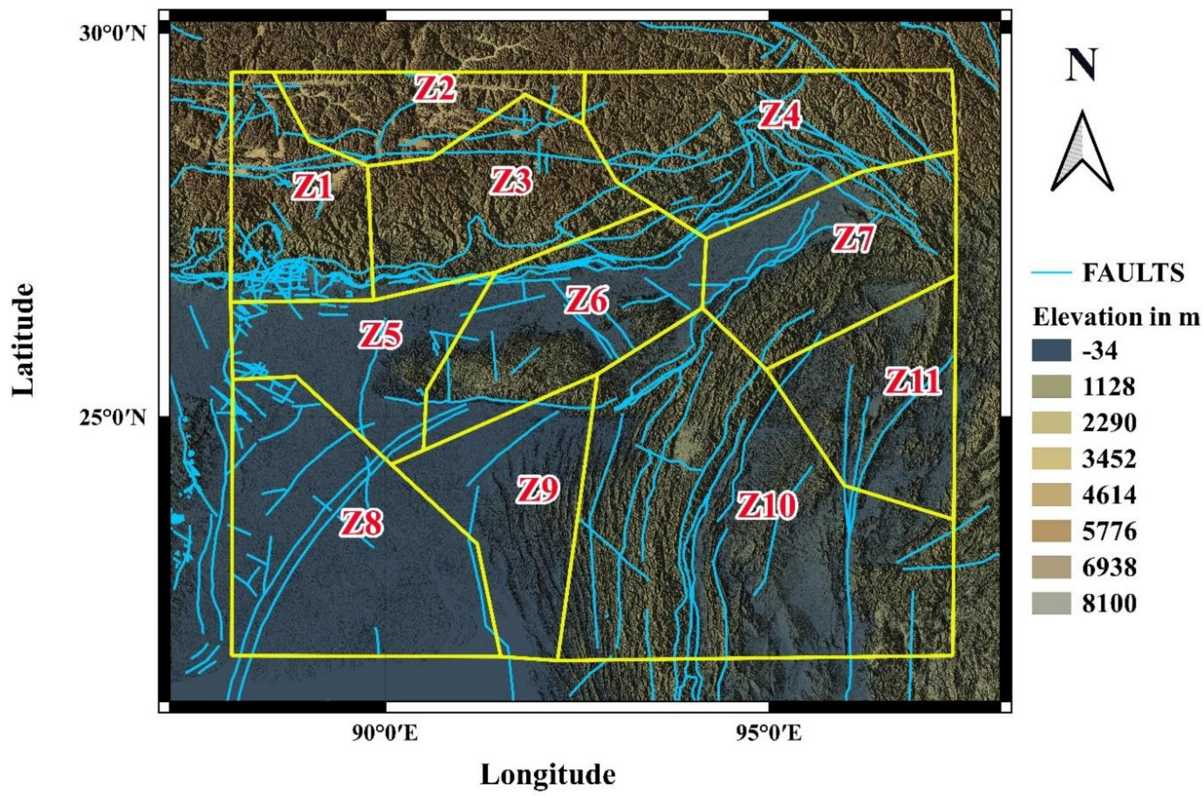

Figure 4.18 Source zones overlying with the shaded relief map of NE India and the Faults present in NE India

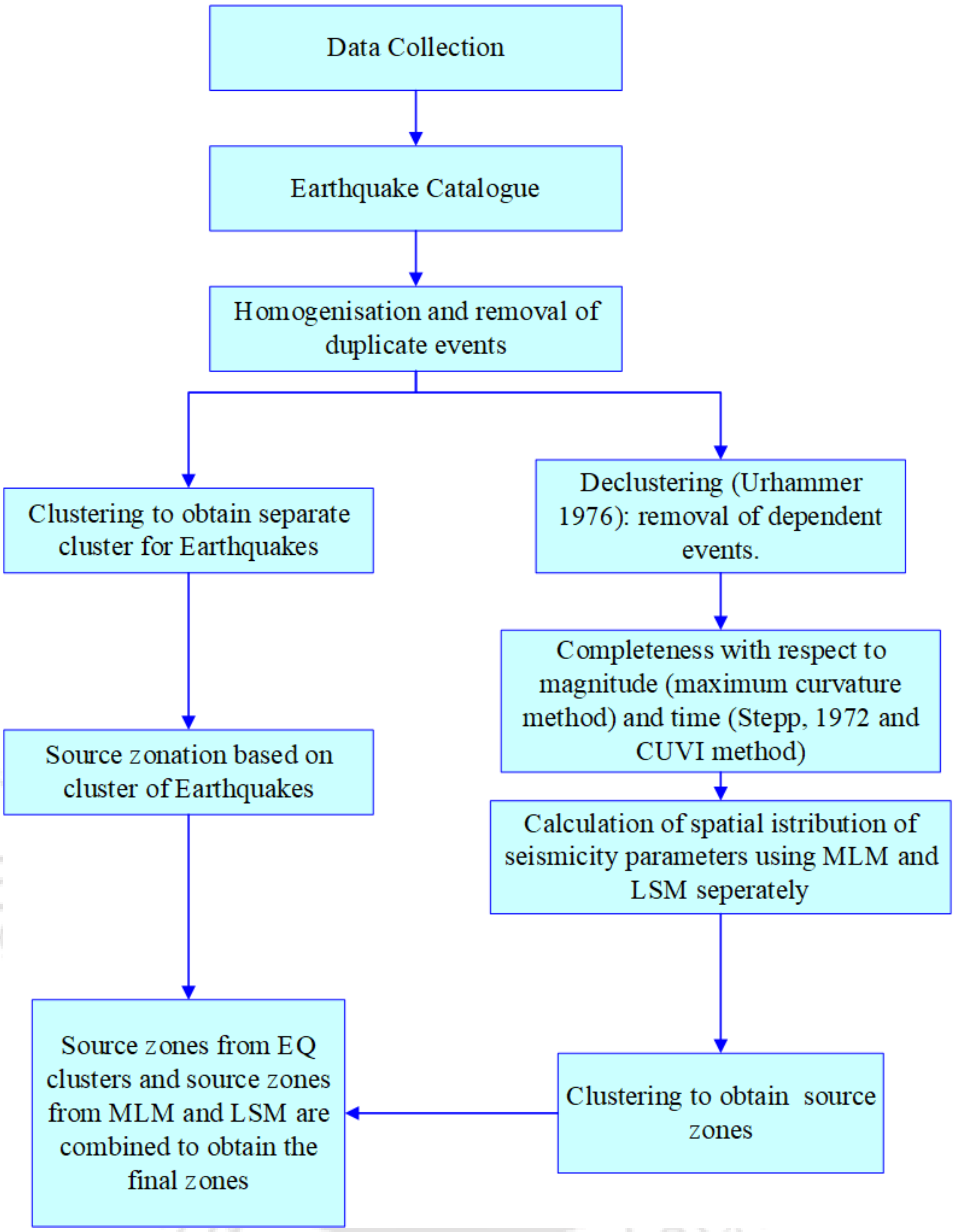

Figure 4.19 Flow chart of various activities followed in the present work to obtain source zones of the NE India region

4.7 Summary

Figure 4.19 depicts the whole analysis process performed in this chapter as a flow chart. Identification of source zones will be helpful in understanding the relative contribution of each source zone to the seismic hazard of the region/ site. Overcoming the limitation of previous studies where source zonations were done for the region primarily based on visual identification (which can vary from person to person) alone, the present work follows a hierarchical clustering approach with the KL index. Firstly, source zones are identified based on past EQ locations.

Later, a declustered complete EQ catalogue is used for the identification of source zones based on variation in seismicity parameters following LSM and MLM approaches. Once source maps based on past EQ locations and seismicity parameters are developed, these are used together to develop source zonation maps.

The final zonation improves the uniform distribution of past EQs and seismicity parameters within the zones, which is the primary objective of seismic source zonation. It is only after the detailed analyses that a final map of source zonation based on visual observation of two separate maps is also proposed for general understanding. This map has a combined effect of both LSM and MLM. The final map divides the entire NE into 11 source zones. Besides this final zonation map, the zonation map based on the location of past EQ is also essential as it covers the surrounding area of NE India. For SHA, one can choose between these zonation maps per the requirement.

Chapter 5. Probabilistic Seismic Hazard Analysis for NE India

5.1 General

EQs are natural events, thus associated with uncertainty in their location, size, and ground motion. PSHA serves as a tool to deal with these uncertainties by providing a well-defined prediction of future EQ scenarios at specific sites. In the present work, PSHA is being carried out for NE India. The steps related to PSHA are mentioned in Chapter 2, Section 2.5. The entire work comprises the utilisation of the EQ catalogue, selection of source models, and determination of seismicity parameters. This Chapter presents the entire procedure adopted for PSHA of NE India and the results obtained from the analyses.

5.2 EQ Catalogue Used for the Analysis

For SHA, information about past EQs in terms of magnitude, epicentre coordinates, time of occurrence, focal depth etc. is crucial. For the present work, the declustered EQ catalogue developed developed in Chapter 4 is considered for the analyses. The entire procedure of preparing the EQ catalogue for NE India is mentioned in Section 4.4.1, Chapter 4. The declustered catalogue consists of a total of 5622 independent EQs. Out of these 5622 events, numbers of EQ events having Mw < 4.0, 4.0 ≤ Mw < 5.0, 5.0 ≤ Mw < 6.0, 6.0 ≤ Mw < 7.0, 7.0 ≤ Mw < 8.0 and Mw≥ 8.0 are 214, 4452, 754, 159, 36 and 7 respectively. The oldest great EQ present in this EQ catalogue occurred in the years 825 AD and 1100AD (Rajendran et al. 2004; NDMA 2010; Borah et al. 2021). This catalogue is used in this work for the determination of seismicity parameters, as mentioned in the later section.

5.3 Seismic Source Model

Chapters 2 and **4** highlight that different parts of NE show variable tectonics and dissimilar seismic activities. Thus, the entire NE cannot be considered as a single seismic source zone. In this work, two models of seismic source zones are considered to account for epistemic uncertainties associated with the position and geometry of seismic sources (Grünthal et al. 2018). The subsections below provide information on each source zone model considered.

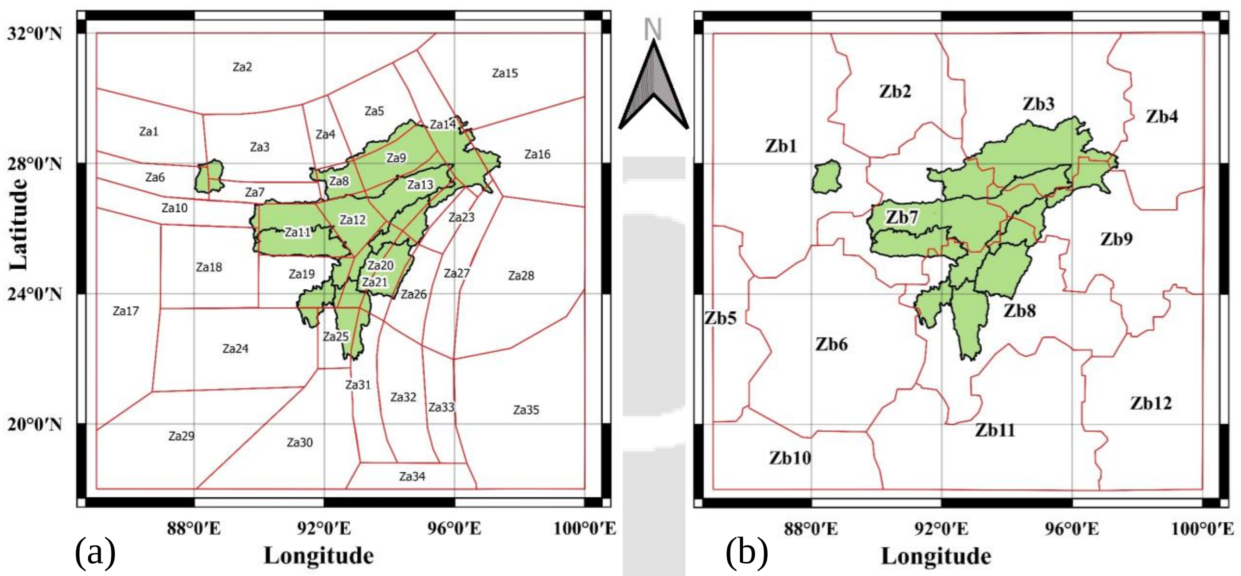

Figure 5.1 Areal sources of the region for (a) SM1 and (b) SM2

5.3.1 Source Model 1 (SM1) based on Seismicity and Tectonic Features

Gupta (2006) and Kolathayar and Sitharam (2012) did seismic source delineation of NE based on visual observations of EQ events and tectonics. As a result, a total of 33 areal sources were identified by Gupta (2006) and Kolathayar and Sitharam (2012). Within these 33 areal sources, the Shillong Plateau, the Bengal Basin and a part of the Assam Valley are present in a single areal source. As per Angelier and Baruah (2009), however, all these three source zones (the Shillong Plateau, the Bengal Basin and one part of the Assam Valley) can be separated from each other on the basis of tectonics, regional geology, focal mechanism and spatial distribution of seismicity. Hence, in the present work, while the other 32 sources identified by Gupta (2006) and Kolathayar and Sitharam (2012) are considered the same, the 33rd source zone (which contains the Shillong Plateau, the Bengal Basin and one part of Assam Valley) is further divided into three areal sources. This way, a total of 35 source zones (as shown in Figure 5.1 a) are considered in SM1 for the present analyses.

5.3.2 Source Model 2 (SM2) based on Distribution of Past EQ Events

Chapters 2 and 4 highlighted that source zonations of NE done by prior studies were primarily based on tectonic setting, geology etc. As a result, such zonations do not ensure a uniform distribution of past EQs and seismicity parameters within each identified source zone. Highlighting these limitations of previous studies, the present study used the hierarchical clustering method on past EQs for the NE region and identified 12 source zones, ensuring uniform distribution of past EQs. Further details on these identified source zones can be found in Chapter 4. Figure 5.1 (b) presents 12 source zones for NE India as per the present study. For the current work, this source model is referred to as SM2.

5.4 Determination of Seismicity Parameters for NE India

Seismicity parameters determine the frequency of various magnitude EQs to occur in the study region. In order to estimate these seismicity parameters, EQs belonging to each source zone in the declustered EQ catalogue are to be separated first from the combined catalogue. While doing so, it is observed that some source zones have very few numbers for EQs required to estimate seismicity parameters. For such source zones, EQs from adjoining areal sources (which individually have a smaller number of EQs) are combined to estimate seismicity parameters for the combined source zone. For SM1 (Figure 5.1 a), the source zones combinations considered are (Za1 +Za3 +Za6 +Za7 +Za10), (Za4 +Za5 +Za8 +Za9 +Za13 +Za14), (Za17 +Za18 +Za24 +Za29 +Za30), (Za11 +Za12 +Za19 +Za20 +Za25), (Za21 +Za22 +Za23 +Za26 +Za27 +Za31 +Za32 +Za33 +Za34), (Za15 +Za16) and (Za28 +Za35). Similarly, for SM2 (Figure 5.1 b), Zb5, Zb6, and Zb10 are source zones with very few EQ events. These source zones are combined (Zb5 + Zb6 + Zb10) to calculate the seismicity parameters for SM2. Once zones with sufficient EQs are ready, such catalogues have to be checked for completeness with respect to magnitude and time as discussed in the next subsection.

5.4.1 Completeness Analyses

Before assessing seismicity parameters from the EQ catalogue, the completeness of the catalogue with respect to magnitude and time must be checked. The magnitudes of completeness (Mc) for each of the earlier-mentioned zones (individual as well as combined) are determined using the MAXC method (Wiemer and McNutt 1997; Wiemer and Katsumata 1999; Wiemer and Wyss 2000). The plot of the total number of EQs greater than or equal to different magnitudes versus magnitude values (known as Frequency Magnitude Distribution or FMD) for each source zone is developed, and the magnitude corresponding to maximum curvature is identified as Mc. Typical FMDs for source zones Zb1 and Zb4 are shown in Figure 5.2 (a) and (b), respectively, suggesting Mc values for both Zb1 and Zb4 as 4.3 (Mw). Similarly, Mc for other (combined and individual) source zones (of both source models) are determined. For all the cases, the Mc value is found to be 4.3 (Mw).

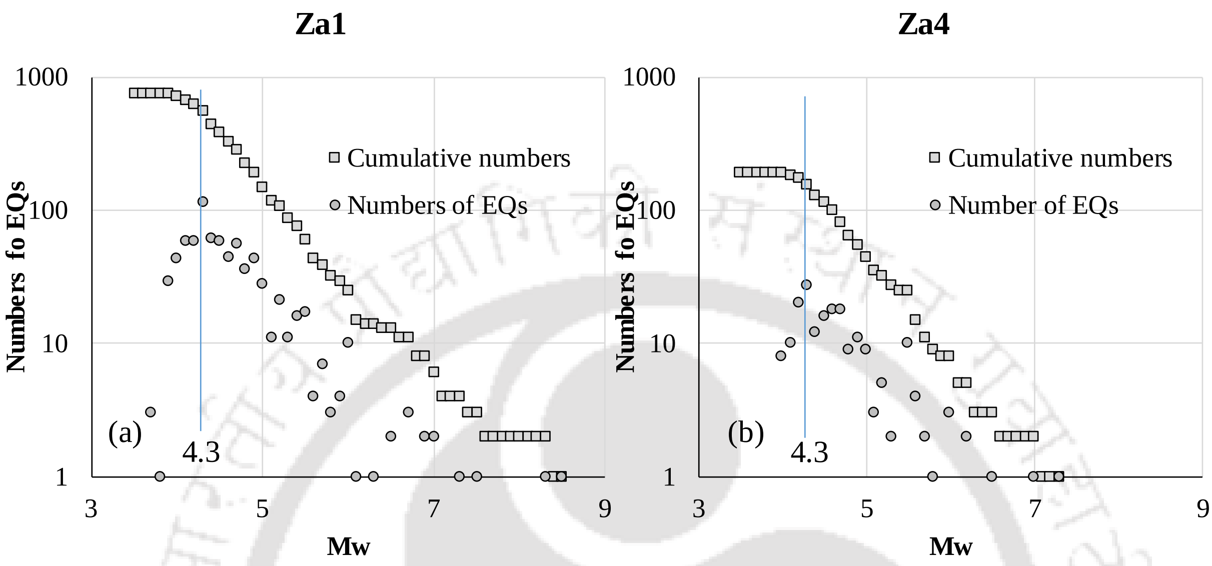

Figure 5.2 Completeness with respect to magnitude for (a) Zb1 of SM2 (b) Zb4 of SM2

Table 5.1 Year of completeness for different ranges of Magnitude (Mw) for SM1

Magnitude ranges	Za2	Za1+Za3+Za6+Za7+Za10	Za4+Za5+Za8+Za9+Za13+Za14	Za17+Za18+Za24+Za29+Za30	Za11+Za12+Za19+Za20+Za25	Za21+Za22+Za23+Za26+Za27+Za31+Za32+Za33+Za34	Za15+Za16	Za28+Za35
4.3-4.7	1984	1964	1962	1977	1966	1964	1975	1970
4.8-5.2	1973	1963	1958	1964	1958	1959	1962	1961
5.3-5.7	1967	1958	1958	1982	1959	1958	1925	1921
5.8-6.2	1937	1950	1940	1935	1924	1956	1923	1917
6.3-6.7	1902	1833	1941	1834	1842	1926	1870	1929
6.8-7.2	1904	1852	1905	1737	1885	1906	1870	1914
7.3-7.7	1411	1833	1806	1787	1762	1858	1870	1839

7.8-8.2	1411	1100	No EQ	1787	1762	No EQ	1870	1839
≥8.3	No EQ	1100	No EQ	No EQ	No EQ	No EQ	No EQ	No EQ

Table 5.2 Year of completeness for different ranges of Magnitude (Mw) for SM2

Magnitude ranges	Zb1	Zb2	Zb3	Zb4	Zb7	Zb8	Zb9	Zb11	Zb12	Zb5 +Zb6 +Zb10
4.3-4.7	1963	1987	1968	1969	1965	1967	1962	1970	1986	1979
4.8-5.2	1963	1959	1962	1964	1958	1967	1961	1964	1986	1976
5.3-5.7	1923	1932	1925	1930	1932	1959	1924	1964	1984	1982
5.8-6.2	1899	1924	1938	1923	1933	1930	1925	1942	1917	1935
6.3-6.7	1833	1902	1905	1870	1915	1927	1926	1963	1988	No EQ
6.8-7.2	1852	1915	1905	1870	1923	1938	1906	1963	1941	1737
7.3-7.7	1762	No EQ	1762	1762	1806	1762	1762	1762	1762	No EQ
7.8-8.2	No EQ	1897	No EQ	No EQ	1806	No EQ	No EQ	1762	1762	1737
≥8.3	No EQ	No EQ	1762	No EQ	No EQ	No EQ	No EQ	No EQ	No EQ	No EQ

Once Mc values are known, completeness analysis with respect to time is performed using the CUVI method (Mulargia and Tinti 1985). For this, EQs in each source zone are divided into different magnitude classes (keeping ΔM=0.5Mw) as: i) 4.3≤Mw≤4.7, ii) 4.8≤Mw≤5.2, iii) 5.3≤Mw≤5.7, iv) 5.8≤Mw≤6.2, v) 6.3≤Mw≤6.7, vi) 6.8≤Mw≤7.2, vii) 7.3≤Mw≤7.7, viii) 7.8≤Mw≤8.2 and ix) Mw≥8.3. Figure 5.3 (a to f) and Figure 5.4 (a to f) show typical plots of the cumulative number of EQs versus year for Zb1 and Zb4. From Figure 5.3 (a), it can be observed that EQ occurrence is stable after the year 1963. Before 1963, the cumulative frequency of EQ occurrence did not follow any definite pattern. Thus, for magnitude class 4.3≤Mw≤4.7, the EQ catalogue for Zb1 is complete after 1963. Similar observations can be drawn from other plots of Figure 5.3 (a-f) and Figure 5.4 (a-f) for Zb1 and Zb4, respectively.

Summary of the duration of completeness for SM1 and SM2 are presented in Table 5.1 and Table 5.2, respectively.

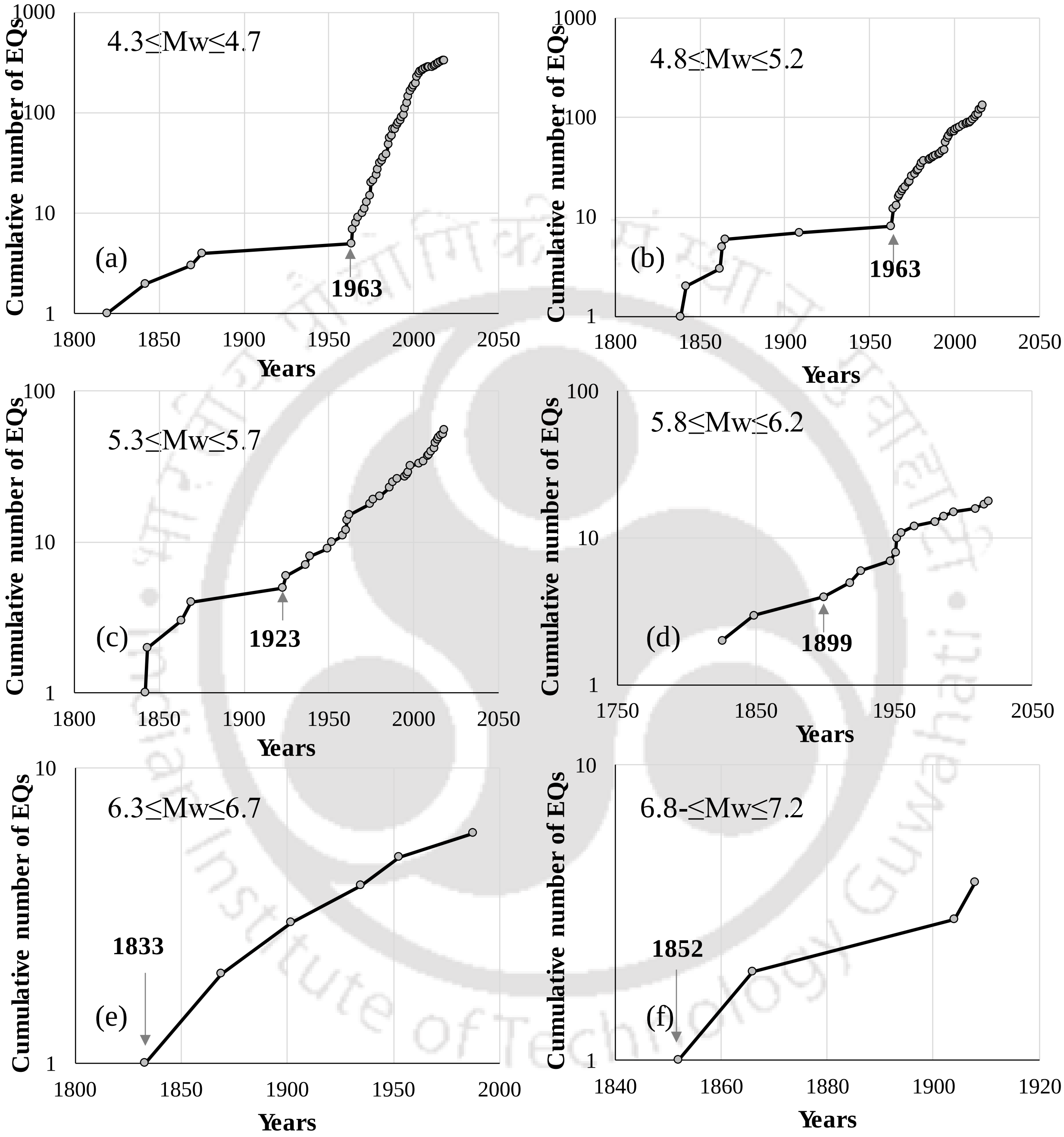

Figure 5.3 Plots of the cumulative number of EQs versus year to obtain years of completeness for Zb1 of SM2

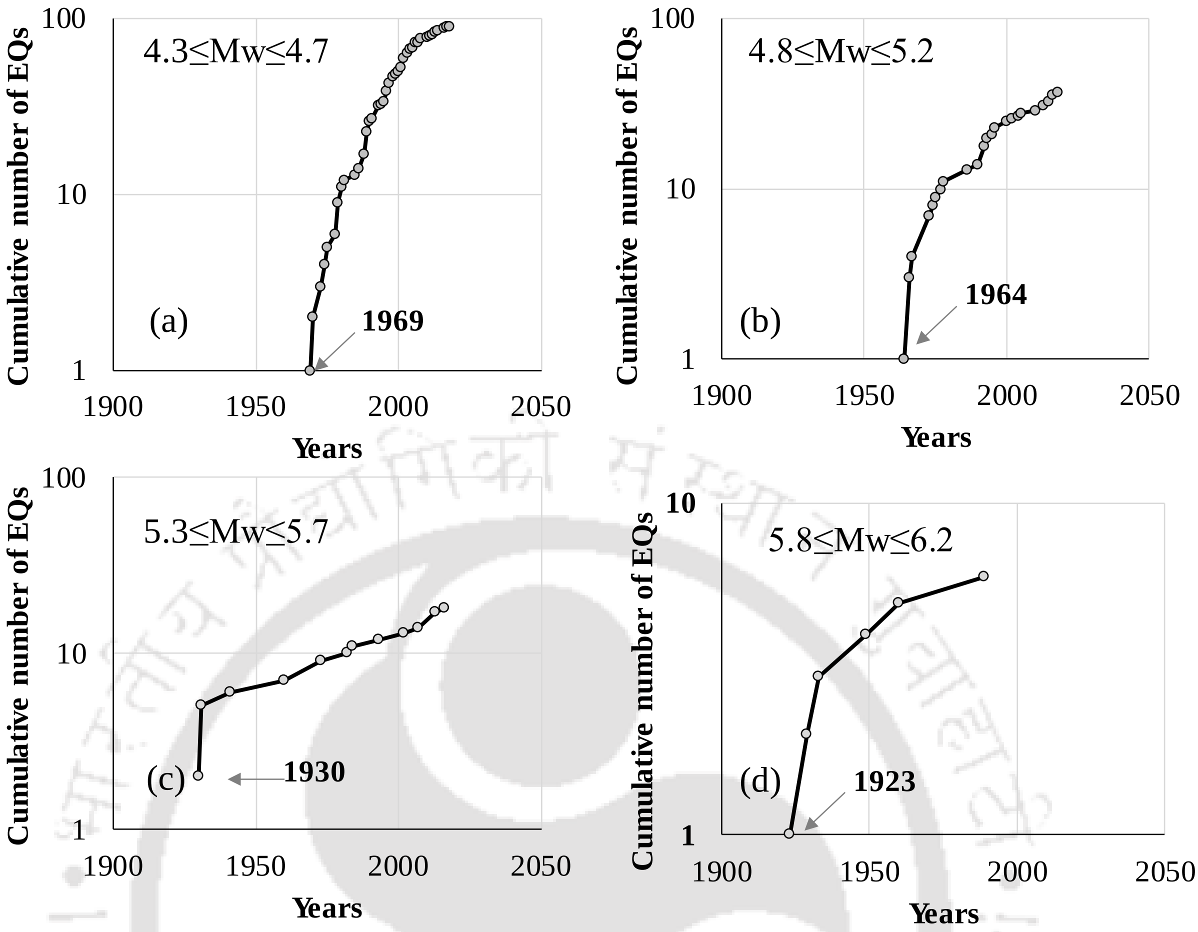

Figure 5.4 Plots of cumulative number of EQs versus year to obtain years of completeness for Zb4 of SM2

Only the complete portion of the EQ catalogues are used for assessing seismicity parameters, as discussed in the next section.

5.4.2 Estimation of Seismicity Parameters

The number (N) of EQs having a magnitude equal to or greater than a particular magnitude (M) can be related to M by the G-R recurrence relation as (Gutenberg and Richter 1944):

$$\log N(M) = a - bM \quad 5.1$$

In the present work, seismicity parameters ‘a’ and ‘b’ are estimated by using MLM as LSM provides biased estimates of the b-value.

In MLM, "β” (=2.303b) is determined by solving the following maximum-likelihood condition (Weichart 1980);

$$\frac{\sum_i t_i m_i e^{-\beta m_i}}{\sum_i t_i e^{-\beta m_i}} = \bar{m} \qquad 5.2$$

Where, m_i = value of the central magnitude of the ith interval, t_i = periods of completeness in number of years for the ith interval, $\bar{m}$ = average value of magnitude in the catalogue. Once β (=2.303b) is known, the total number of EQs per year greater than Mc (N_{MC}) can be determined as (as per Weichart 1980);

$$N_{Mc} = N_T \frac{\sum_i e^{-\beta m_i}}{\sum_i t_i e^{-\beta m_i}} \qquad 5.3$$

Where, N_T stands for the total number of events. Once N_{MC} is known, "a" parameter can be determined as;

$$\mathrm{a} = \log N_{Mc} + \mathrm{b} M_{min} \qquad 5.4$$

This way, the values of "a" and "b" can be obtained. Uncertainties in "a" and "b" values have also been estimated by using Weichart (1980) method. Table 5.3 and Table 5.4 summarise the values of "a" and "b" and their standard deviation obtained from MLM for SM1 and SM2, respectively.

Table 5.3 Seismicity parameters obtained for areal sources of SM1

SM1 areal sources	MLM				Mc
	a	b	σ(a)	σ(b)	
Za2	5.35	1.05	0.22	0.05	4.3
Za1+ Za3+Za6+Za7+Za10	4.84	0.94	0.17	0.04	4.3
Za4+Za5+Za8+Za9+Za13+Za14	4.78	0.93	0.20	0.05	4.3
Za17+Za18+Za24+Za29+Za30	4.32	0.93	0.3	0.07	4.3
Za11+Za12+Za19+Za20+Za25	5.03	0.97	0.17	0.04	4.3
Za21+Za22+Za23+Za26+Za27+Za31+Za32 +Za33+Za34	5.80	1.02	0.11	0.03	4.3
Za15+Za16	5.07	1.03	0.24	0.06	4.3
Za28+Za35	5.01	0.96	0.17	0.04	4.3

Table 5.4 Seismicity parameters obtained for areal sources of SM2

SM2 areal sources	MLM				Mc
	a	b	σ(a)	σ(b)	
Zb1	5.41	1.04	0.16	0.04	4.3
Zb2	4.42	0.93	0.30	0.07	4.3
Zb3	5.25	1.05	0.22	0.05	4.3
Zb4	4.68	0.98	0.28	0.07	4.3
Zb7	5.00	0.96	0.18	0.04	4.3
Zb8	5.91	1.06	0.12	0.03	4.3
Zb9	5.19	0.97	0.15	0.03	4.3
Zb11	5.05	0.97	0.19	0.04	4.3
Zb12	4.42	0.89	0.25	0.06	4.3
Zb5+Zb6+Zab10	4.25	0.95	0.35	0.08	4.3

5.4.3 Source-wise Distribution of Seismicity Parameters

In the last section, values of seismicity parameters have been determined for the individual as well as combined source zones. In case a combined zone was considered for estimating seismicity parameters, such seismicity parameters must be redistributed to individual zones. Iyengar and Ghosh (2004) utilised the "principle of superposition" to redistribute "a" parameter from a region to individual faults within that region on the basis of fault length and numbers of EQs. Iyengar and Ghosh (2004) considered the b value of all the faults within a region to be the same as the "b" value of that region.

In the present study, referring to the work by Iyengar and Ghosh (2004), the "b" value of all individual source zones (which were combined to estimate seismicity parameters) is considered the same as the "b" value of the combined source zones. For the determination of "a" value, however, the number of EQs that had occurred on an individual source zone '*k*' is solely considered by defining the weight, W_k in redistribution as;

$$W_k = \frac{Number\ of\ EQs\ within\ \ individual\ source\ zone\ k}{Number\ of\ EQs\ in\ the\ combined\ source\ zone} \qquad 5.5$$

Once W_k is known, the activity, "N_k" for source zone, "k" can be obtained using Eq. 5.6 as;

$$N_k = 10^{(a-bM_{min})} \times W_k \qquad 5.6$$

Parameter N_k in Eq. 5.6 is equal to $\alpha(m^0)$ (which is used in Eq. 5.10 later) that defines the frequency of m^0 (=Mc) and above magnitude to occur in individual source zone 'k'. Following the above procedure, $\alpha(m^0)$ and b values for all source zones of SM1 and SM2 are obtained. Table 5.5 and Table 5.6 summarise $\alpha(m^0)$ and b values for each of the source zones of SM1 and SM2, respectively.

Table 5.5 Seismicity parameters of areal sources of SM1 used for PSHA

SM1 areal sources	MLM		SM1 areal sources	MLM	
	$\alpha(m_0)$	b		$\alpha(m_0)$	b
Za1	2.11	0.94	Za19	1.44	0.97
Za2	6.90	1.05	Za20	0.97	0.97
Za3	0.74	0.94	Za21	2.26	1.02
Za4	0.55	0.93	Za22	0.53	1.02
Za5	0.40	0.93	Za23	1.66	1.02
Za6	2.07	0.94	Za24	0.55	0.93
Za7	0.77	0.94	Za25	0.86	0.97
Za8	1.44	0.93	Za26	8.78	1.02
Za9	0.86	0.93	Za27	4.22	1.02
Za10	0.52	0.94	Za28	4.84	0.96
Za11	1.98	0.97	Za29	0.19	0.93
Za12	1.91	0.97	Za30	0.46	0.93
Za13	0.20	0.93	Za31	1.78	1.02
Za14	2.62	0.93	Za32	4.76	1.02
Za15	2.84	1.03	Za33	0.93	1.02
Za16	1.51	1.03	Za34	0.73	1.02
Za17	0.18	0.93	Za35	2.82	0.96
Za18	0.73	0.93			

Table 5.6 Seismicity parameters of areal sources of SM2 used for PSHA

SM2 areal sources	MLM	
	$\alpha(m_0)$	b
Zb1	8.87	1.04
Zb2	2.69	0.93

Zb3	5.73	1.05
Zb4	2.77	0.98
Zb5	0.14	0.95
Zb6	1.20	0.95
Zb7	7.39	0.96
Zb8	22.63	1.06
Zb9	10.68	0.97
Zb10	0.12	0.95
Zb11	7.83	0.97
Zb12	4.13	0.89

5.5 Maximum Possible Earthquake Magnitude (Mp) for Each Areal Source

EQ catalogue sometimes may not indicate the true potential of the seismic source which is very important for assessing the seismic hazard values. As a result, the maximum possible magnitude (Mp) of a source can be larger than the largest observed magnitude (M_{max}^{obs}) known on the source. Many statistical methods of Mp assessments (Kijko and Sellevoll 1989; Kijko and Graham 1998; Kijko 2004; Kijko and Singh 2011) are available, which require information on past EQs. Besides this, Mp also depends on other properties (length, rupture characteristics etc.) of the seismic source. Thus, in this work, two methods are used for calculating Mp. In the first method, based on fault length, Mmax (called as Mmax-fault) is calculated using the relation of Wells and Coppersmith (1994). For the calculation of Mmax-fault, the fault length is considered as the subsurface rupture length, and the slip type is considered as all (Table 2A in Wells and Coppersmith 1994)

In the second method, Mmax is calculated using the Robson–Whitlock–Cooke procedure (Kijko and Singh 2011). In this method, Mp is calculated by increasing the Mmax-obs by a suitable margin like, the method suggested by Gupta (2002), which was widely used by various researchers (NDMA 2010; Bahuguna and Sil 2018) for NE India. However, in the Robson–Whitlock–Cooke procedure, the maximum possible magnitude (MmaxRWC) is estimated by using the following formula for truncated magnitude distribution (Kijko and Singh 2011).

$$\text{MmaxRWC} = M_{max}^{obs} + 0.5(M_{max}^{obs} - M_{n-1}) \quad 5.7$$

Where;

M_{max}^{obs}=maximum observed magnitude.

M_{n-1}=the second largest observed magnitude.

Once values of Mmax from both the above methods are known for each source zone, the following four criteria are considered for choosing one Mp value for each source zone.

1. If Mmax-obs < Mmax-fault < MmaxRWC, then Mmax-fault is considered as Mp.
2. If Mmax-fault < Mmax-obs, Mmax-obs is considered as Mp.
3. If Mmax-fault > MmaxRWC by a small amount (up to 0.2Mw), Mmax-fault is considered as Mp.
4. If Mmax-fault > MmaxRWC by a larger amount (more than 0.2Mw), MmaxRWC is considered as Mp.

Values of Mp for each source zone for SM1 and SM2, following the above-mentioned criteria, are summarised in Table 5.7 and Table 5.8, respectively.

Table 5.7 Mp estimation for areal sources of SM1

Zones	Mmax obs	Mmax RWC	Fault length (km)	Mmax-fault	Mp	Zones	Mmax obs	Mmax RWC	Fault length (km)	Mmax-fault	Mp
Za1	6.7	7.1	324	8.1	7.1	Za19	7.6	7.7	160	7.7	7.7
Za2	8	8.2	409	8.3	8.3	Za20	6	6.1	173	7.7	6.1
Za3	6	6.2	321	8.1	6.2	Za21	7.2	7.5	292	8.1	7.5
Za4	7.7	8	119	7.5	7.7	Za22	5.9	6.1	117	7.5	6.1
Za5	7.7	9	172	7.7	7.7	Za23	6.7	7.1	93	7.3	7.3
Za6	7.5	7.6	335	8.1	7.6	Za24	7.2	7.8	307	8.1	7.8
Za7	6.8	7.2	427	8.3	7.2	Za25	7.5	8.3	173	7.7	7.7
Za8	6.8	7.2	157	7.7	7.2	Za26	7.2	7.3	228	7.9	7.3
Za9	7.2	7.8	213	7.8	7.8	Za27	7.8	8	330	8.1	8.1
Za10	8.5	8.6	253	8.0	8.5	Za28	7.8	7.9	372	8.2	7.9
Za11	8.1	8.2	272	8.0	8.1	Za29	5.9	6.1	143	7.6	6.1
Za12	7.2	7.7	136	7.6	7.6	Za30	6.7	7.1	244	7.9	7.1
Za13	6.5	6.8	283	8.0	6.8	Za31	6.1	6.2	276	8.0	6.2
Za14	7.1	7.5	216	7.9	7.5	Za32	7.5	8	492	8.4	8
Za15	6.5	6.7	320	8.1	6.7	Za33	6.6	6.7	62	7.1	6.7

Za16	8.5	9.1	162	7.7	8.5	Za34	6	6.1	90	7.3	6.1
Za17	5	5.1	143	7.6	5.1	Za35	8	8.4	296	8.1	8.1
Za18	7.8	8.6	196	7.8	7.8						

Table 5.8 Mp estimation for areal sources of SM2

Zones	Mmax obs	MmaxRWC	Fault length (km)	Mmax-fault	Mp
Zb1	8.5	8.6	890	8.8	8.8
Zb2	8	8.2	138	7.6	8
Zb3	8.5	8.9	493	8.4	8.5
Zb4	7.3	7.5	249	8	7.5
Zb5	5	5.2	123	7.5	5.2
Zb6	7.8	8.1	463	8.4	8.1
Zb7	8.1	8.6	410	8.3	8.3
Zb8	7.8	8.0	392	8.2	8.2
Zb9	7.6	7.7	245	7.9	7.9
Zb10	5.9	6.2	77	7.2	6.2
Zb11	7.8	8.1	517	8.4	8.1
Zb12	8	8.4	212	7.8	8

5.6 Dealing With the Uncertainties of PSHA

PSHA was developed primarily to account for the effects of magnitude uncertainty, hypocentral distance uncertainty, and uncertainty in ground motion exceedance (Kramer 1996). Eq. 5.8 can be used to summarise all of these uncertainties (EM-1110-2-6050, 1999):

$$v_z = \sum_{n=1}^{N} \sum_{m_i=m^0}^{m_i=m^u} \lambda_n(m_i) \left[\sum_{r_j=r_{min}}^{r_j=r_{max}} P_n(R = r_j|m_i) P(Z > z|m_i r_j)\right] \quad 5.8$$

Where,

v_z = annual frequency of exceedance of ground motion level "z".

$\sum_N$ = summation of results from N numbers of seismic sources.

$\lambda_n(m_i)$ = annual frequency of EQ occurrence of magnitude m_i

on n^{th} seismic source.

$P_n(R = r_j|m_i)$ = probability of occurrence of an EQ having magnitude m_i on nth source at a distance r_j away from the site of interest.

$P(Z > z|m_i r_j)$ = probability of exceedance of ground motion level z for an EQ of magnitude m_i located at a distance r_j from the site of interest.

In PSHA, v_z is utilised to find the probability of exceedance $P_E(z)$ of ground motion of level z at a site, in a given exposure period (t) by following Eq. 5.9:

$$P_E(z) = 1 - e^{-v_z.t} \qquad 5.9$$

5.6.1 Magnitude uncertainty

The seismicity parameters are the primary source of information for estimating magnitude uncertainty. By utilising Eq. 5.10, the cumulative number of EQs of magnitude m on an areal source can be determined (EM-1110-2-6050, 1999).

$$N(M > m) = \alpha(m^0)\frac{10^{-b(m-m^0)} - 10^{-b(m^u-m^0)}}{1 - 10^{-b(m^u-m^0)}} \qquad 5.10$$

Where b and $\alpha(m^0)$ for all sources obtained in previous sections, and Mc and Mp are considered as m^0 and m^u, respectively. The frequency, $\lambda_n(m_i)$ of magnitude m_i to occur on an areal source can be calculated as per EM-1110-2-6050, (1999) by using the following Eq. 5.11:

$$\lambda_n(m_i) = N\left(m_i - \frac{\Delta m}{2}\right) - N\left(m_i + \frac{\Delta m}{2}\right) \qquad 5.11$$

In this study, Δm for each areal source is considered equal to $\frac{1}{6}(m^u - m^0)$.

5.6.2 Probability of EQ occurrence in areal sources

For the sake of this study's analysis, it is assumed that the probability of occurrence of an EQ is constant throughout an entire areal source. Thus, the probability of occurrence of an EQ at a given distance R may be calculated by dividing the area between R+ΔR/2 and R-ΔR/2 by the total area of the areal source. ΔR signifies the difference between two consecutive distances required to determine the probability. Calculating the area between two successive radial distances, on the other hand, is exceedingly complex and time-consuming for an irregularly shaped areal source. Additionally, for different places, each time, the size and shape of the area within specific distances varies, making the process more challenging. Another method for finding the probability of EQ occurrence in areal sources is to convert the whole area to small grids. But in this present study, a novel method is used. Rather than treating the source as an area, the entire source is assumed to be filled with a huge number of points spaced at extremely close distances. Therefore, the chance of EQ occurring at a specific distance R may be estimated by dividing the number of points between R+ΔR/2 and R-ΔR/2 by the total number of points of the areal source. Calculating the distances between different locations and the points will be simple and convenient after the points are generated using this new way. The shape of the source will not affect the distance calculation in this new methodology, making it an easy and simple procedure.

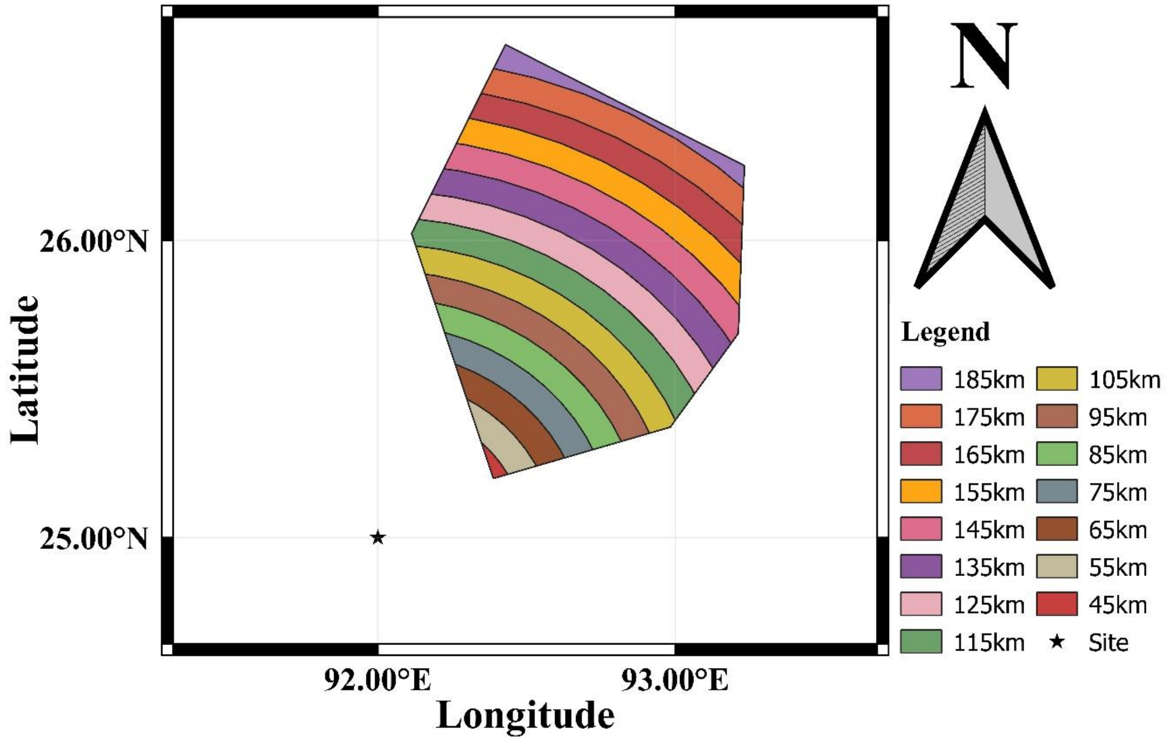

Figure 5.5 Area considered for each distance to calculate probabilities

Prior to using this novel strategy, it is vital to compare the result obtained from points consideration and areas consideration. A random irregular areal source and a site are considered for this purpose, as seen in Figure 5.5. Distance probabilities are estimated for distances of 5km, 15km, 25km, and so on, using R=10km. Figure 5.5 also illustrates the areas considered when calculating probability for distances of 5km, 15km, 25km, and so on. Areas are calculated for all distances using the QGIS software. Then, using the aforementioned procedure, the probabilities of occurrences are calculated for each distance considering areas. The alternative method is developing a Matlab code to generate the same area source and the points included within it (a total of almost 10^5 points). The number of points for each distance is determined using the new method, and the probabilities are calculated. The results of both methods are compared in Figure 5.6, and a satisfactory match can be seen. As a result, this analysis employs this new approach for evaluating the distance uncertainty ($P_n(R = r_j|m_i)$ which will be used in Eq. 5.8 for areal sources.

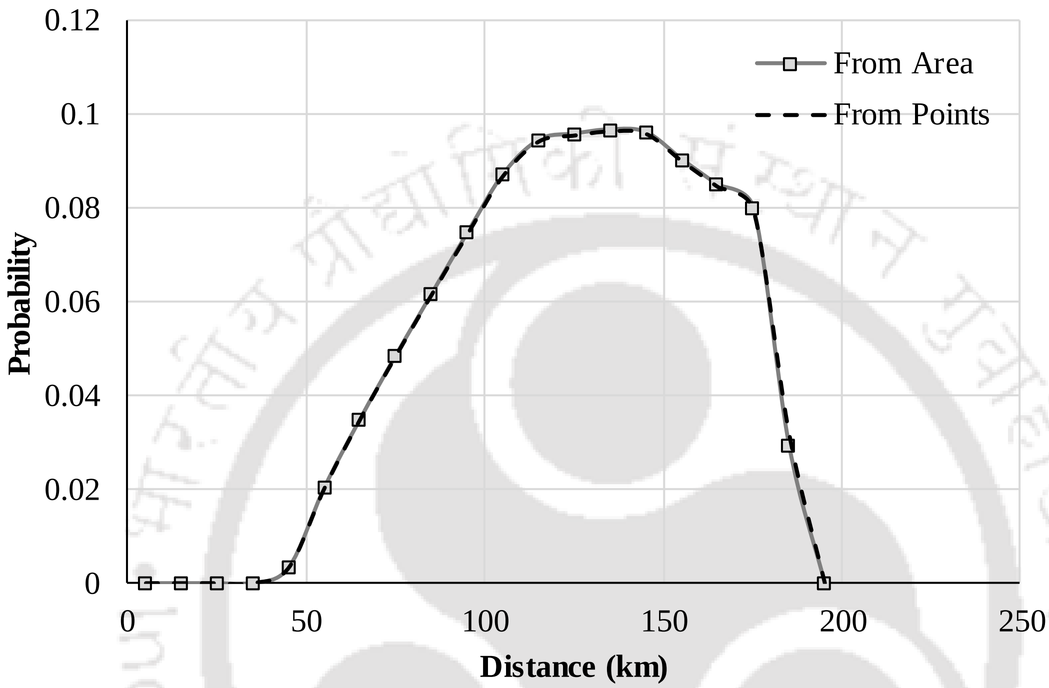

Figure 5.6 Comparison of the distance probabilities obtained from considering area and considering points.

5.6.3 Uncertainty in Ground Motion Exceedance

Knowing the probable ground motion level at a certain location is crucial for designing EQ-resistant structures. GMPEs can be used to calculate the probable level of ground motion at a specific site based on the probable EQ magnitude and hypocentral distance. These GMPEs are built using either existing EQ ground motion records or synthetic ground motion records, depending on the site's properties or a combination of the two. While generating the GMPE, some data scattering is observed due to the rupture mechanics, path, site conditions, source variability and heterogeneity. This type of scattering is quantified by including the projected GMPE's standard deviation. This standard deviation reflects some degree of error in the ground motion level computation using the GMPEs (Kramer 1996). PSHA accounts for this uncertainty by estimating the probability of the ground motion parameter exceeding a threshold value for a particular mi and r_j. The conditional probability distribution $P(Z > z|m_i r_j)$ can be estimated using the following equation (Kramer 1996):

$$P(Z > z|m_i r_j) = 1 - F_z(Z^*) \tag{5.12}$$

Where, $F_z(Z^*)$is the value of the cumulative distribution function (CDF) of Z at m_i and r_j. In the case of GMPE, $F_z(Z^*)$ is primarily considered a lognormal distribution (EM-1110-2-6050, 1999) and can be expressed as follows:

$$F_z(Z^*) = F\left\{\frac{\ln(z) - \overline{\ln(z)}}{S[\ln(z)]}\right\} \tag{5.13}$$

Where,

$\ln(z)$ = logarithm of the ground motion level,

$\overline{\ln(z)}$ = mean natural logarithm of the ground motion level,

$S[\ln(z)]$ = standard deviation of the natural logarithm of the ground motion level.

F {} = cumulative distribution function of a unit normal variable.

For certain GMPEs, 10-based logarithms are also considered instead of natural logarithms. For the present study, values of z are taken in increments of 0.05g, starting from a minimum value of 0.01 for the calculation of $F_z(Z^*)$. The values of m_i and r_j considered are taken from section 5.6.1 and section 5.6.2. For each of the sources, the corresponding $F_z(Z^*)$ are determined for the region.

5.7 GMPEs Considered

In the declustered EQ catalogue developed earlier, *M* value varies from 4.3 to 8.8 (Mw). Further, based on the position of source zones, the epicentral distances vary up to 400km. Referring to the works by Baro et al. (2018) and Baro et al. (2020) and taking the above range of magnitude and epicentral distance into account, three GMPEs, namely: Toro (2002), NDMA (2010), and Anbazhagan et al. (2013) are used for the current study. It should be noted here that GMPEs by NDMA (2010) and Anbazhagan et al. (2013) were developed for NE India and the Himalayan region, and GMPE by Toro (2002) was developed for the subduction zone. All three GMPEs are valid for site class A conditions. However, to apply the GMPEs, the ranking of these GMPEs is required to be found out. Based on the ranking, the weights of the GMPEs used in the logic tree for the determination of seismic hazard are found. In the present work, the log-likelihood (LLH) ranking method is used to determine the weights of the GMPES. In the LLH ranking method, the LLH factor is found, which is defined by

$$\mathrm{LLH} = \frac{-1}{N}\sum_{i=1}^{N}\log_2[g(x_i)] \quad 5.14$$

Here,

$g(x_i)$ = likelihood of the model for observation x_i,

N = number of observations

From LLH factors, the weights for each GMPE can be obtained by the following formula.

$$w_k = \frac{2^{-LLH_k}}{\sum_{k=1}^{nG} 2^{-LLH_k}} \quad 5.15$$

Where,

LLH_k=LLH of the k-th GMPE

nG= number of the GMPEs

For the present work, a total of 98 Accelerograms recorded at rock sites for 17 EQs that occurred in NE India are considered to estimate the *LLH* factor for all the GMPEs. The LLH factors and the weights obtained for the three GMPEs are shown in Table 5.9. Further, these weights are used in logic tree to carry out the PSHA for the NE India region as described in later sections.

Table 5.9 Log Likelihood (LLH) values and weight obtained for the GMPEs

GMPE	LLH	Ranking	weights
Anbazhagan et al. (2013)	4.14	2	0.22
NDMA (2010)	6.38	3	0.05
Toro (2002)	2.41	1	0.73

5.8 Logic Tree Considered and Overall Uncertainty Determination in PSHA

Variations in seismotectonics, methods adopted and models used in SHA lead to variations in the estimation of the seismic hazard values at a particular location (Cramer 2001; Cramer et al. 2002). This variability is caused by both random natural processes (aleatory uncertainty) and uncertainties in knowledge and measurements (epistemic uncertainty). As discussed earlier, in order to incorporate uncertainty in predicted seismic hazard values, two source models (SM1 and SM2), and three GMPEs are considered in the present analysis using a logic tree (see Figure 5.7). Weights assigned to each component of the logic tree are shown in Figure 5.7. For the two source models, equal weights have been given for the analyses. For GMPEs, the weights obtained in section 5.7 are assigned, keeping the sum of the total weights as 1.0.

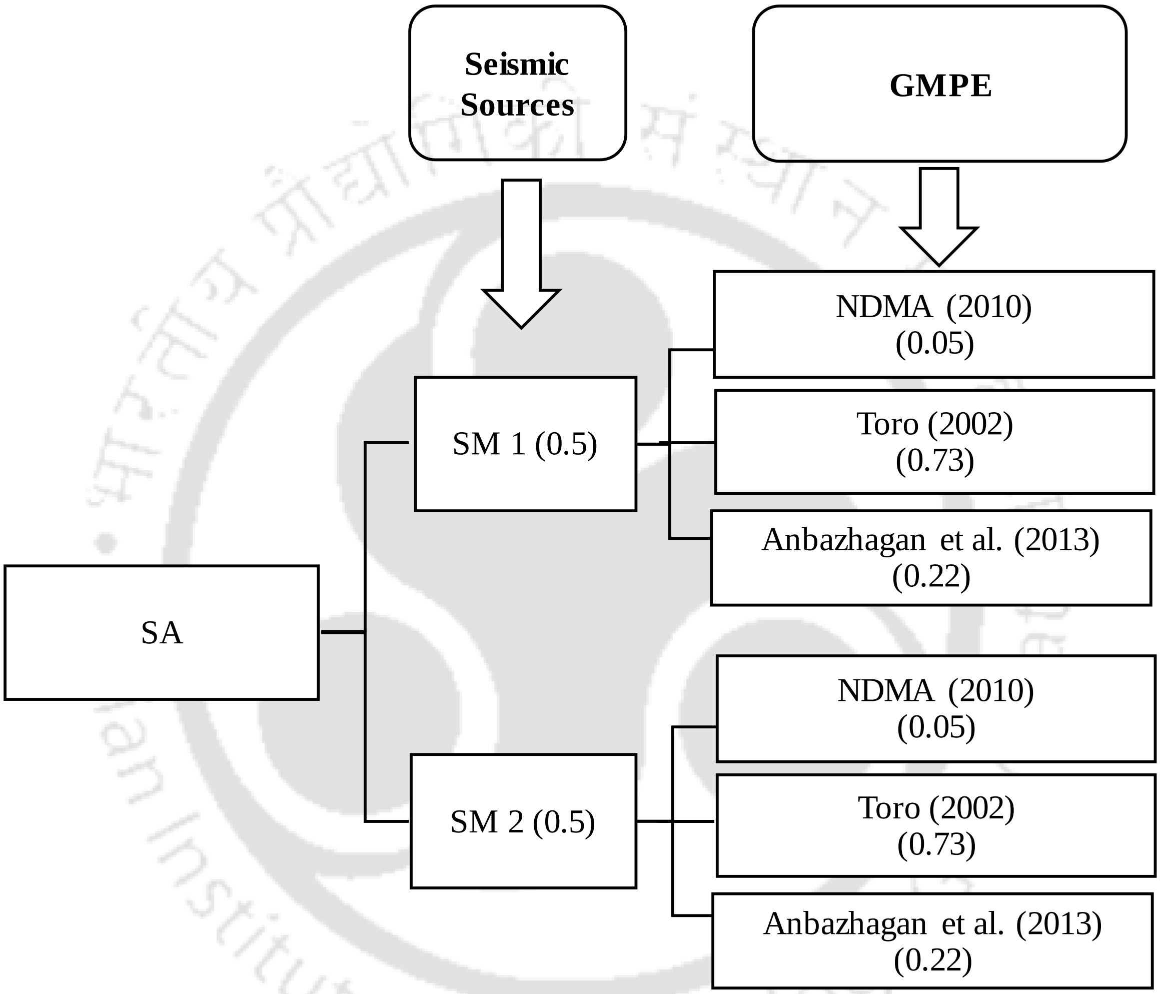

Figure 5.7 Logic-tree considered for PSHA in this work

In the present study, by using the declustered EQ catalogue developed in Chapter 4, areal sources identified in Section 5.3 , values of parameters estimated in Section 5.4 , GMPEs selected discussed in Section 5.7 , and the methodology adopted described in Section 5.6 , PSHA is performed considering each logic tree branch. There are two methods for estimating the final probable seismic hazard level from a logic tree. In the first method, the hazard level is first estimated for each alternative model, which is then weighted in accordance with the weights provided in the logic tree. In the second method, the logic tree is sampled by using a sampling technique to build a distribution of alternative models, and then hazard levels are estimated from the same (Cramer 2001), which in turn is used to estimate the uncertainty/ variability of the final hazard levels. Generally, uncertainty/variability is expressed in terms of either standard deviation or coefficient of variation (CoV) along with the mean hazard value. The CoV can be defined as the ratio of standard deviation and mean of any distribution. Though the mean hazard level represents the best estimation of seismic hazard for the region (Cramer 2001), the CoV determination is also important as it describes the confidence in the mean hazard level estimation (Das et al. 2016). In other words, a CoV map gives a visual idea of the effect of the lack of knowledge in the estimated seismic hazard levels (Cramer 2001; Cramer et al. 2002). In the present work, the uncertainties associated with limited knowledge are addressed by using the Monte Carlo approach to sample the logic tree used for PSHA. The method is described in the following section.

5.9 Monte Carlo Sampling of the Logic Tree

In order to account for uncertainty in estimated seismic hazard values, two source models (SM1 and SM2), and three GMPEs are used for the PSHA utilising a logic tree (see Figure 5.7). Figure 5.7 depicts the weights allocated to each component of the logic tree. In the Monte Carlo approach, multiple values of possible seismic hazard levels are determined by generating Monte Carlo samples of the logic tree. The steps followed to generate the Monte Carlo samples of the logic tree are given below:

1. Based on the weight assigned to each component of the logic tree, the probability of occurrence of each branch is calculated (Table 5.10).
2. Based on the probabilities of occurrence, the cumulative probability of occurrence is calculated for each branch. (Table 5.10)
3. Then, based on the cumulative probability of occurrence, random number intervals are set for each of the logic tree branches (Table 5.10)
4. A random number is generated by using the “rand” command in Matlab. This command will generate a random number between 0 to 1 (Table 5.10).
5. Based on the value of the random number, the logic tree branch is selected, and the seismic hazard level is calculated for that branch (Table 5.10).
6. Steps 4 and 5 are repeated until the required seismic hazard level numbers are calculated from the required numbers of sample logic tree branches.

7. Finally, mean hazard level values and CoVs are calculated using the seismic hazard level values obtained from following the above steps of the seismic hazard levels.

A total of 10,000 Monte Carlo samples of the logic tree are generated to obtain the multiple values of possible seismic hazard levels. Based on these values, mean hazard level values and CoVs are calculated. Besides, hazard levels are also obtained for four percentiles, 5^{th}, 16^{th}, 95^{th} and 84^{th}. Hazard levels at different percentiles help in comparing the variability of the estimated hazard values.

Table 5.10 Table showing the probability of occurrence of logic tree branches (Weight of the components of the logic tree are in the brackets)

Branch number	Seismic source model	GMPE	Probability	Cumulative Probability	Range of the random number (Rn)
BN 1	SM1 (0.5)	NDMA (2010) (0.05)	0.025	0.025	0<Rn≤0.025
BN 2	SM1 (0.5)	Toro (2002) (0.73)	0.365	0.39	0.025<Rn≤0.39
BN 3	SM1 (0.5)	Anbazhgan et al. (2013) (0.22)	0.11	0.50	0.39<Rn≤0.50
BN 4	SM2 (0.5)	NDMA (2010) (0.05)	0.025	0.525	0.5<Rn≤0.525
BN 5	SM2 (0.5)	Toro (2002) (0.73)	0.365	0.89	0.525<Rn≤0.89
BN 6	SM2 (0.5)	Anbazhgan et al. (2013) (0.22)	0.11	1	0.89<Rn≤1

5.10 Deaggregation of EQ Hazard for the Important Cities in NE India

The deaggregation of PSHA for PGA corresponding to 10% probabilities in 50 years and 2% probabilities in 50 years are carried out in the present study for 8 important cities of NE India (see Figure 5.8, Table 5.11). For the deaggregation, the magnitude and distance bins considered in the present study are listed in Table 5.12. The selection of these magnitude and distance bins is inspired by the bins suggested by USNRC RG-1.208 (2007).

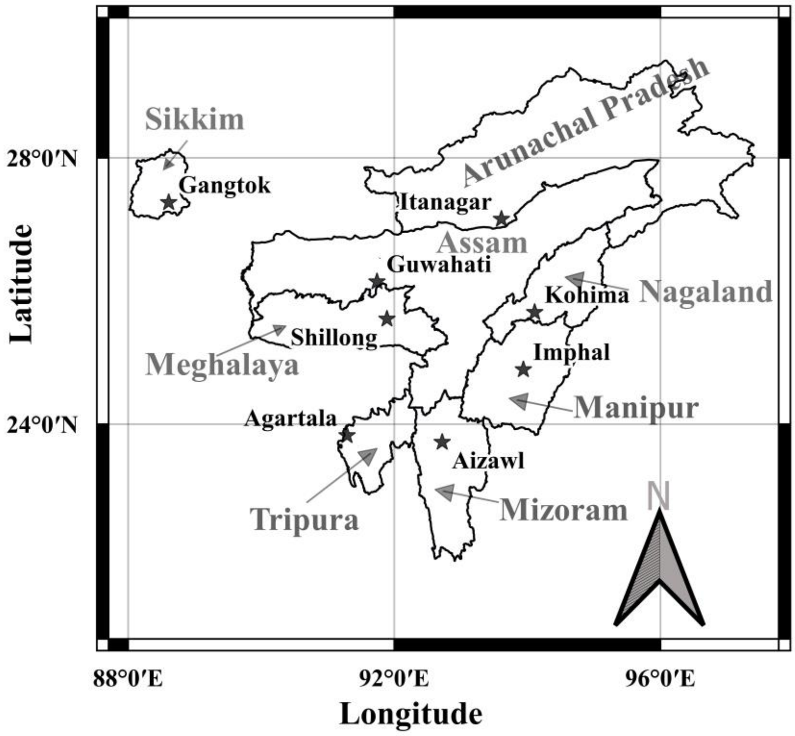

Figure 5.8 Important cities in each of the eight states of NE India (Stars show the location of the cities) considered for deaggregation

Table 5.11 Most important cities in all the states of NE India with centre coordinate

City	State	Long	Lat
Guwahati	Assam	91.74	26.14
Agartala	Tripura	91.29	23.83
Imphal	Manipur	93.94	24.82
Shillong	Meghalaya	91.89	25.58
Aizawl	Mizoram	92.72	23.73
Kohima	Nagaland	94.11	25.68
Gangtok	Sikkim	88.61	27.33
Itanagar	Arunachal	93.61	27.08

Table 5.12 Magnitude and distance bins for Deaggregation of hazard

Magnitude Bins (Mw)	Epicentral distance Bins (km)
4-4.5	0-15
4.5-5	15-25
5-5.5	25-50
5.5-6	50-100
6-6.5	100-200
6.5-7	200-300
7-7.5	300-400
7.5-8	
8≤	

To perform the deaggregation of the seismic hazard, first, PGA values for each of the cities are found by considering the logic tree branches separately by performing PSHA. Then the contributions of the magnitude and epicentral distance bin for each branch of the logic tree are found by using Eq. 5.16. After finding the contribution of the magnitude and epicentral distance bin for different branches of the logic tree, the mean contribution of each magnitude and hypocentral distance bin is obtained by considering the weight assigned to the component of the logic tree (Figure 5.7).

$$v_{PGA} = \sum_{n=1}^{N} \lambda_n(m_{mb}) P_n(R = r_{mb}|m_{mb}) P(Z_{pga} > PGA|m_{mb} r_{mb}) \quad 5.16$$

Where,

v_{PGA} = annual frequency of exceedance of PGA

m_{mb}= the mean magnitude of the magnitude bin

r_{mb}= the mean distance of the distance bin

$\sum_N$ = summation of results from N numbers of seismic sources

$\lambda_n(m_{mb})$ = annual frequency of EQ occurrence of magnitude m_{mb}

on n^{th} seismic source.

$P_n(R = r_{mb}|m_{mb})$ = probability of occurrence of an EQ having magnitude m_{mb} on n^{th} source at a distance r_{mb} away from the site of interest.

$P(Z_{pga} > PGA|m_{mb}r_{mb})$= probability of exceedance of PGA for an EQ of magnitude m_{mb} and located at a distance r_{mb} away the site of interest.

Z_{pga}=the value of PGA obtainded by utilising GMPE for magnitude m_{mb} and epicentral distance r_{mb}..

5.11 Results and Discussion

In the present study, the entire NE is divided into grids of 0.5° × 0.5°. Then, for each grid, the mean hazard level and CoV of the hazard level are estimated for 10 % and 2 % probability of exceedance in 50 years by following the method described in section 10.

Further, considering the magnitude and distance bins listed in Table 5.12, the deaggregations of PSHA for mean PGA corresponding to 10% probabilities in 50 years and 2% probabilities in 50 years are carried out for eight cities of NE India. In the present study, exceedance approach is used for deaggregation of the seismic hazard. To do so, firstly, PGA values corresponding to 10% probabilities in 50 years and 2% probabilities in 50 years are estimated for the eight cities. Then, the annual rate of exceedance of the estimated PGA values is obtained for the different combinations of the selected magnitude and distance bins using Eq. 5.16. This way, the contribution of each combination of the magnitude and distance bins is obtained in the deaggregation of the seismic hazard. Besides this, the deaggregation of seismic hazards in latitude and longitude is also performed in the present study. The entire area is divided into 0.1° × 0.1° grid, and contribution toward PGA values from each grid is estimated. Latitude and longitude base deaggregation is performed to find out the sources that significantly contributed toward the hazard of the cities.

Detailed discussions on the results obtained are given in different subsections below.

5.12 Spatial Variations in Spectral Acceleration for NE

Figure 5.9 (a-d) presents contours of mean Spectral Accelerations (Sa) at 0s, 0.1s, 0.2s and 1s periods corresponding to 10% probability of exceedance in 50 years (also known as design basis earthquake or DBE) for NE India. It can be observed from Figure 5.9 (a) that the present study shows variation in PGA from 0.23g to 0.51g within NE India. The highest seismic hazard level is observed in the eastern part of Manipur and in the southern part of Nagaland (Figure 5.9 a). For Manipur, PGA varies from 0.35g to 0.51g. Similarly, for Nagaland, PGA varies from 0.27g to 0.43g. Considering seismic history, it can be observed that the density of past EQs and the magnitudes are very high near Manipur and Nagaland. Further, both these states lie in the

Indo–Burma ranges that have high seismic activity. The lowest seismic hazard value is observed in Sikkim and parts of Arunachal Pradesh and Assam. For Sikkim, the PGA varies from 0.23g to 0.31g. Though some major and great EQs have occurred near Sikkim in the past, regional seismic activity is relatively low, which can be understood from the fact that the number of past EQs in the region is lower in comparison to adjoining regions.

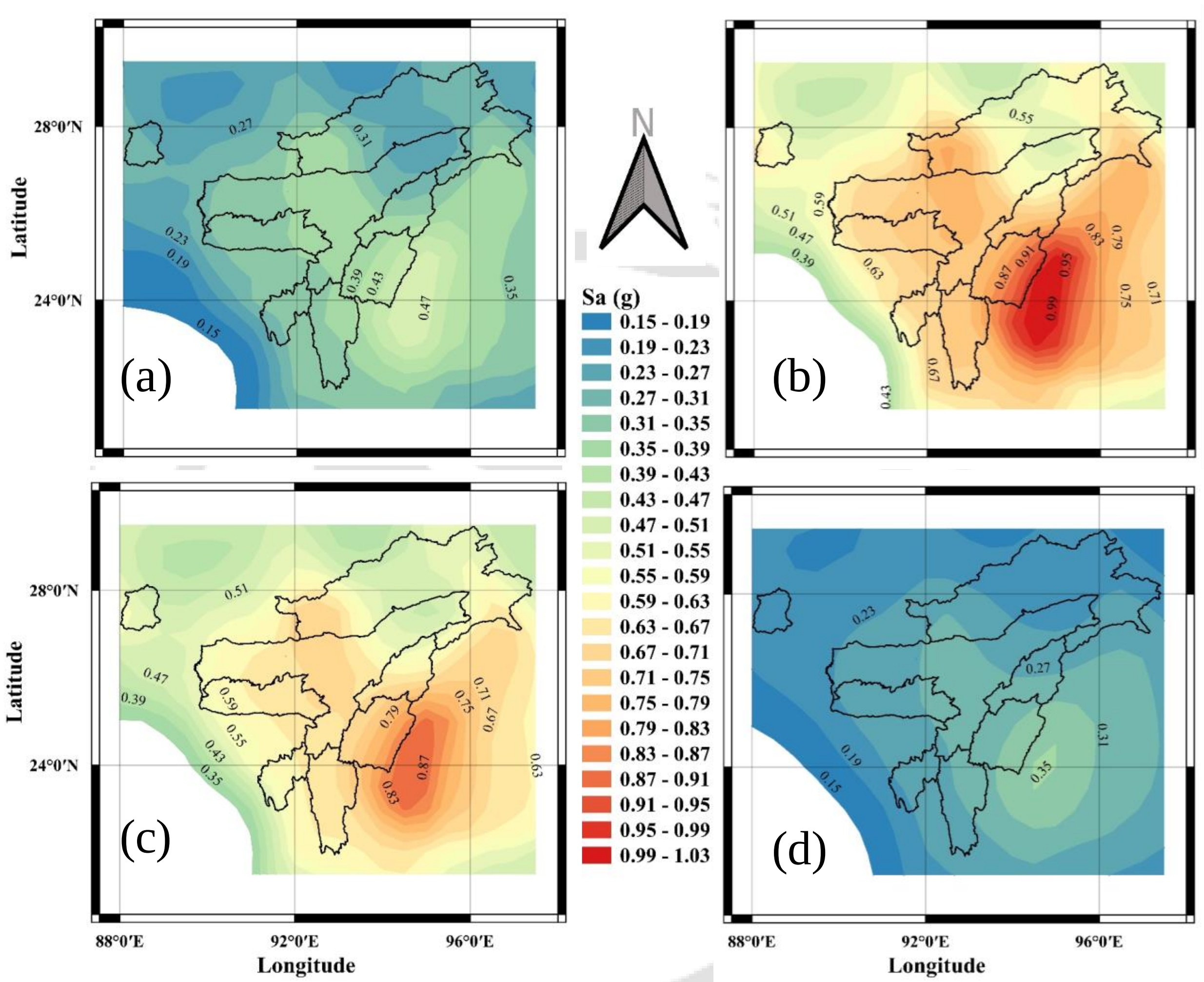

Figure 5.9 Spatial variation in mean seismic hazard level for 10% probability in 50 years for SC A. (a) PGA (b) Sa at 0.1s (c) Sa at 0.2s (d) Sa at 1s

A small segment in the eastern part of Assam and the south-central part of Arunachal Pradesh show PGA ranging from 0.23g to 0.27g. Observing past EQs indicates that this region has a significantly lesser number of past EQs, with the presence of major to great EQs. Further, a very high PGA value (0.35g to 0.39g) is observed in the middle parts of Assam. It is observed that this region falls close to the Bhutan Himalaya and the Shillong Plateau, where the 1897 Assam EQ occurred. Part of Assam (near the Indo-Burma ranges) shows higher PGA (> 0.35g). Similar observations can also be made from western and eastern parts of Arunachal Pradesh, where PGA varies from 0.35g to 0.39g. These regions witnessed a high density of past EQs, and higher magnitude EQs. Based on the above discussion, it can be concluded that both Assam and Arunachal Pradesh are spread over large areas, consisting of regions that have experienced rare to very frequent EQs. As a result, wide variation in PGA values can be witnessed in the above two states (from 0.23g to 0.39g).

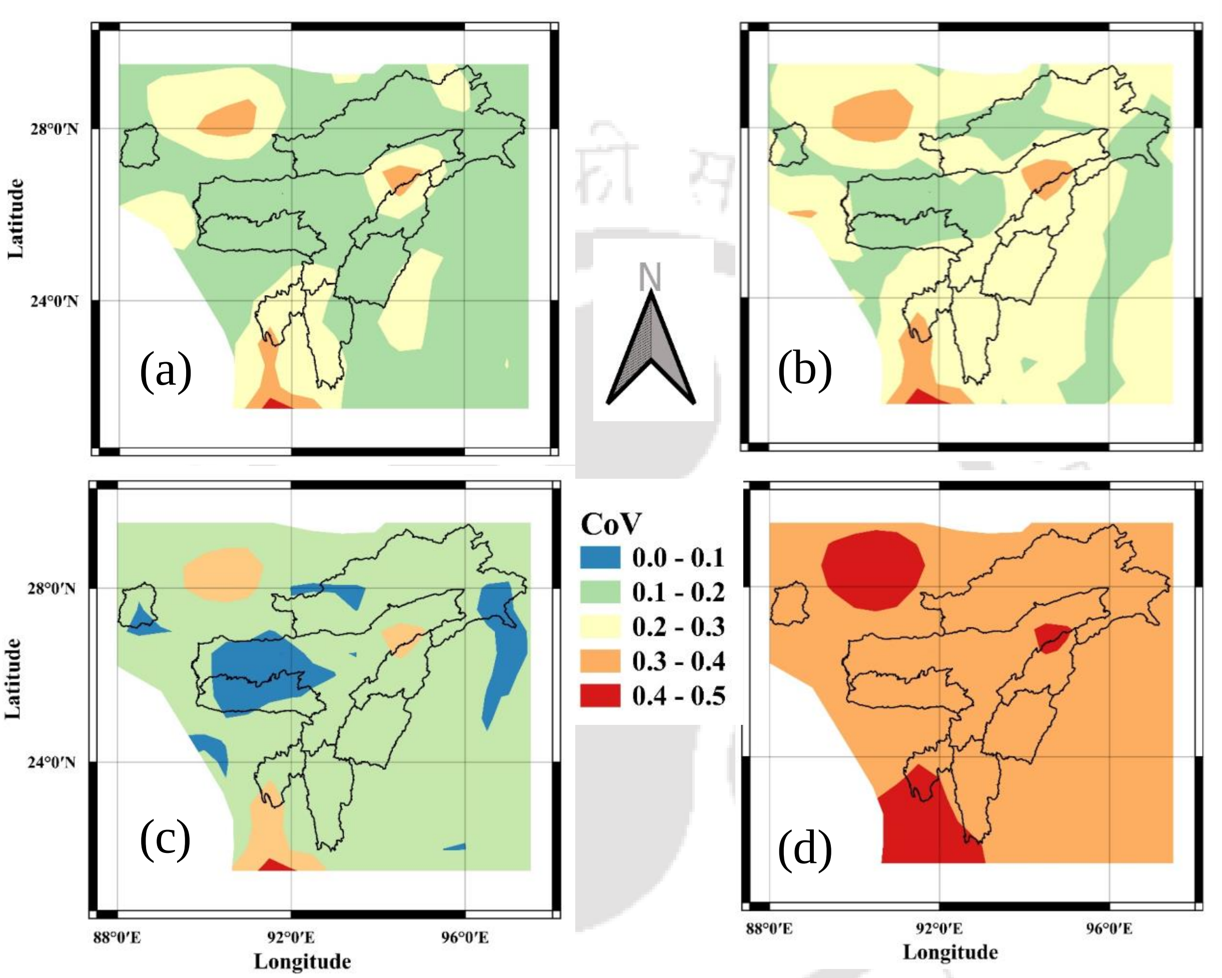

Figure 5.10 Spatial variation in CoV of seismic hazard level for 10% probability in 50 years for SC A. (a) PGA (b) Sa at 0.1s (c) Sa at 0.2s (d) Sa at 1s

Other states of NE do not show such wide variation in PGA values. Meghalaya has PGA varying from 0.27g to 0.39g; Tripura has PGA variation from 0.27g to 0.35g, and Mizoram has PGA variation from 0.31g to 0.39g (see Figure 5.9 a). It should be noted here that Meghalaya has a history of significant numbers of strong, major and great EQs. Similarly, many strong to major EQs had occurred near Tripura and Mizoram. However, based on the distribution of PGA for NE, it can be stated that the highest PGA values are observed near the Indo-Burma ranges, followed by the Bhutan-Himalaya and then the Shillong Plateau. In other words, though NE has experienced a significant number of EQs on regional faults, seismic sources governing the regional seismic hazard values are plate boundaries. Similar observations can be made from Figure 5.9 (b-d) for 0.1s, 0.2s and 1s.

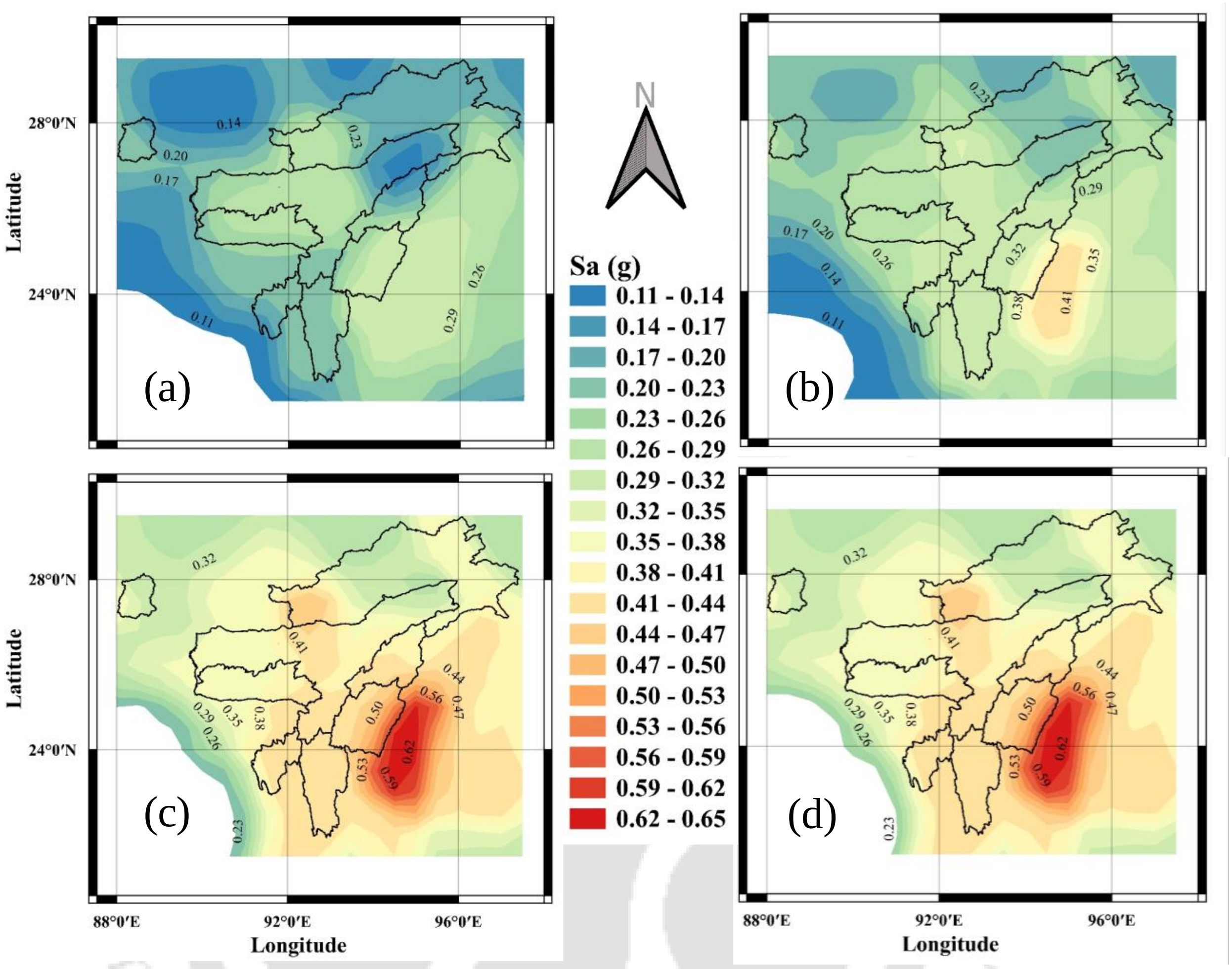

Figure 5.11 Spatial variation in PGA for 10% probability in 50 years for SC A. (a) 5th percentile (b) 16th percentile (c) 84th percentile (d) 95th percentile

As mentioned earlier, to evaluate the variability/uncertainty of the estimated seismic hazard levels, CoVs are calculated. The CoV map of hazard level for 10% probability in 50 years is depicted in Figure 5.10 (a-d) as contours corresponding to Sa at the time periods of 0s, 0.1s, 0.2s and 1s. The CoV of PGA within NE India varies from 0.1 to 0.4, which indicates that the variability in the estimated PGA value is low. However, for Sa at 1s, location-wise CoV values are higher as compared to PGA or other time periods. Figure 5.11 (a-d) shows the spatial variation in PGA for 10% probability in 50 years for SC A. at 5th, 16th, 84th, and 95th percentile. The mean PGA value is closer to the PGA value at the higher percentile.

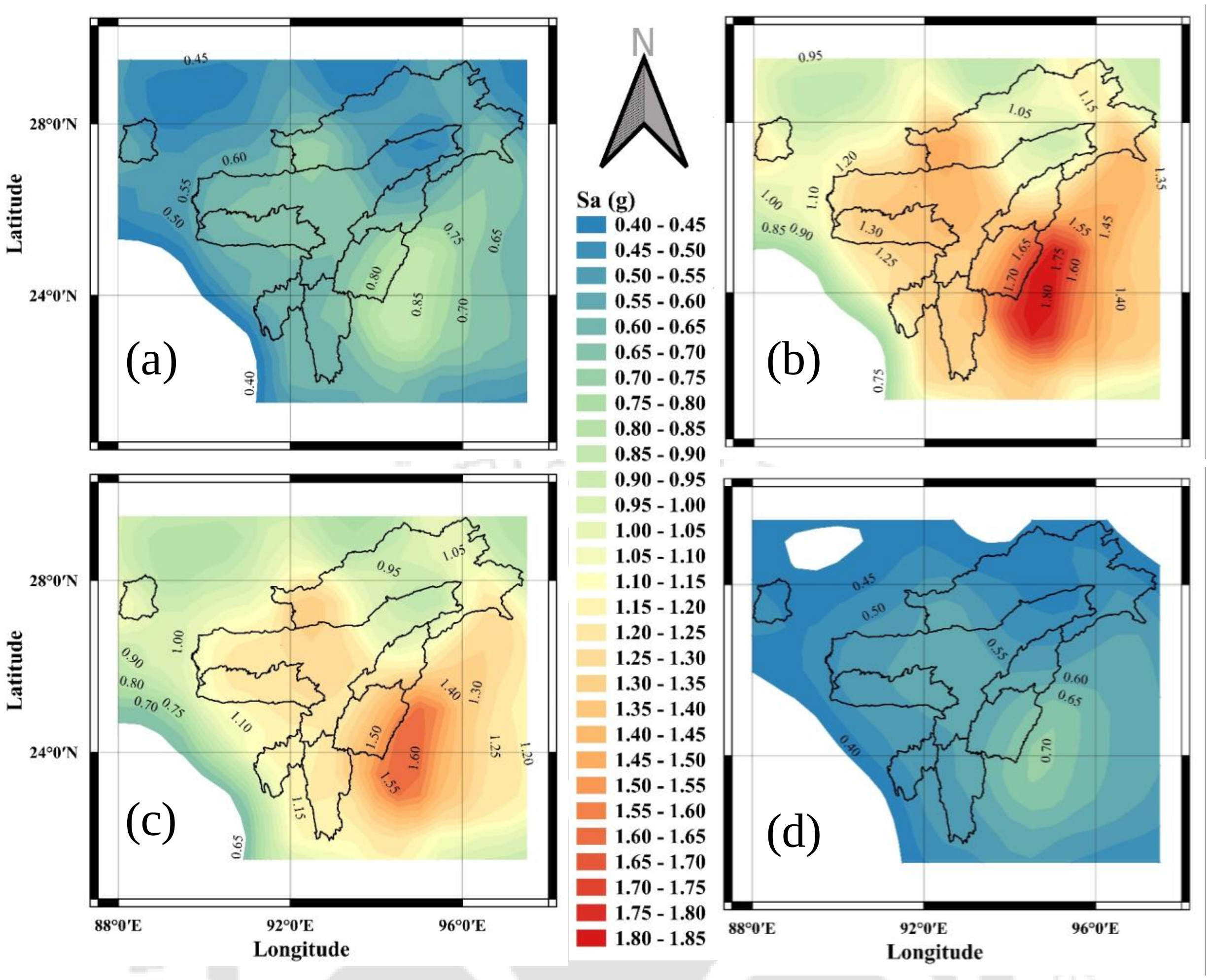

Figure 5.12 Spatial variation in mean seismic hazard level for 2% probability in 50 years for SC A for (a) PGA (b) Sa at 0.1s (c) Sa at 0.2s (d) Sa at 1s

Similarly, Figure 5.12 (a-d) presents PSHA outcomes for 2% probability of exceedance in 50 years (also called Maximum Credible Earthquake or MCE) at 0s, 0.1s, 0.2s and 1s periods, respectively. It is observed that for MCE, PGA values in the NE region vary from 0.40g to 0.85g as per the present analysis. Further, similar to DBE, even MCE-based assessment shows that the dominating seismic sources include the Indo-Burmese range, followed by the Bhutan Himalayas and the Shillong Plateau NE. As mentioned earlier, the CoV values are calculated to find out the variability/uncertainty corresponding to MCE-based hazard level estimation. The CoV map of hazard levels for Sa at 0s, 0.1s, 0.2s and 1s periods corresponding to 2% probability of exceedance in 50 years is depicted in Figure 5.13 (a-d). The CoV of PGA within NE India varies from 0.1 to 0.4. The location-wise CoV values corresponding to Sa at 01s, 02s and 1s are higher than the one estimated for PGA. It is notable that a similar phenomenon was observed for DBE as well. Figure 5.14 (a-d) shows the spatial variation in PGA for 2% probability in 50 years for SC A. at 5th, 16th, 84th, and 95th percentile. Similar to the DBE, in the case of MCE also, the mean PGA value is closer to the PGA value at the higher percentile.

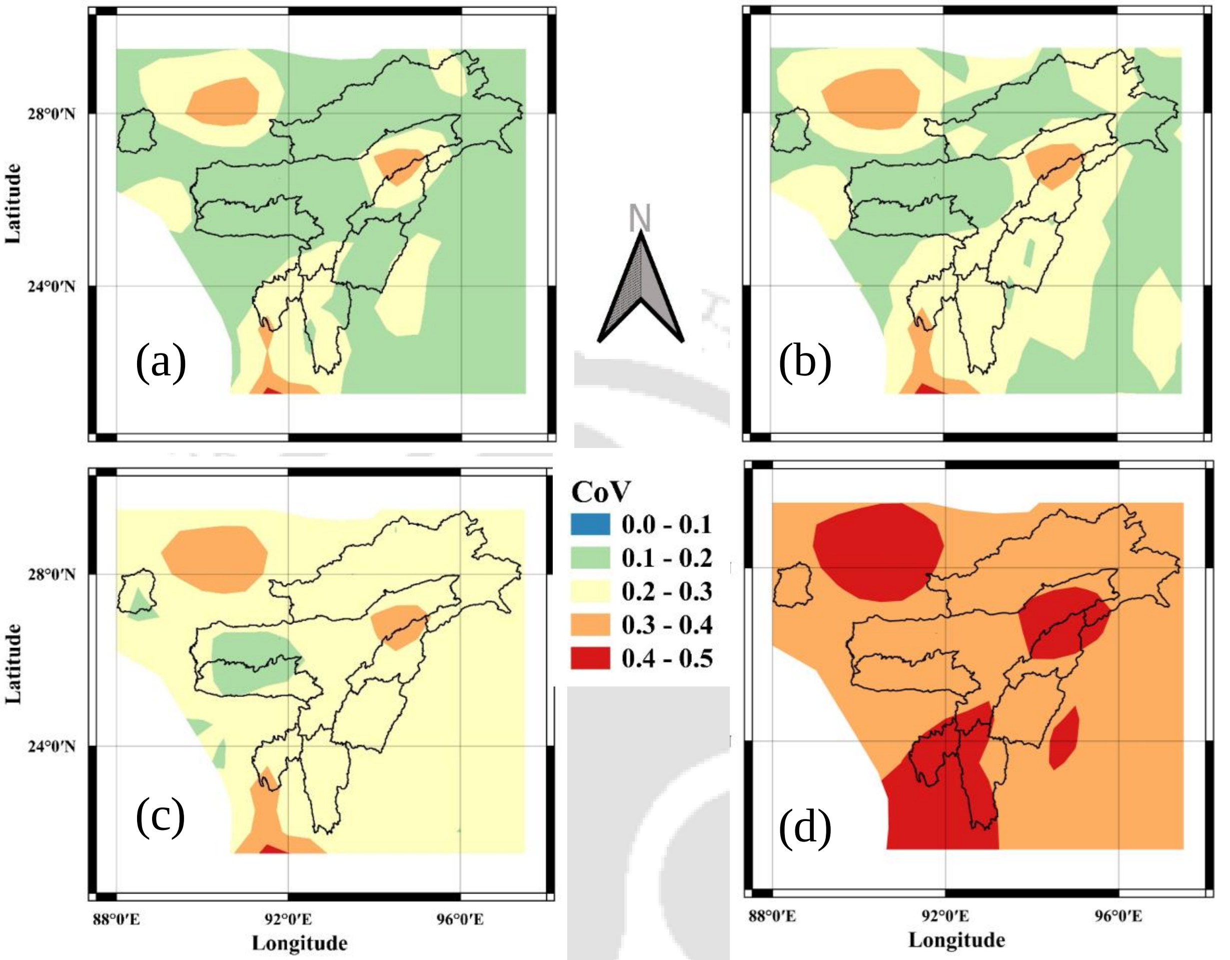

Figure 5.13 Spatial variation in CoV of seismic hazard level for 2% probability in 50 years for SC A. (a) PGA (b) Sa at 0.1s (c) Sa at 0.2s (d) Sa at 1s

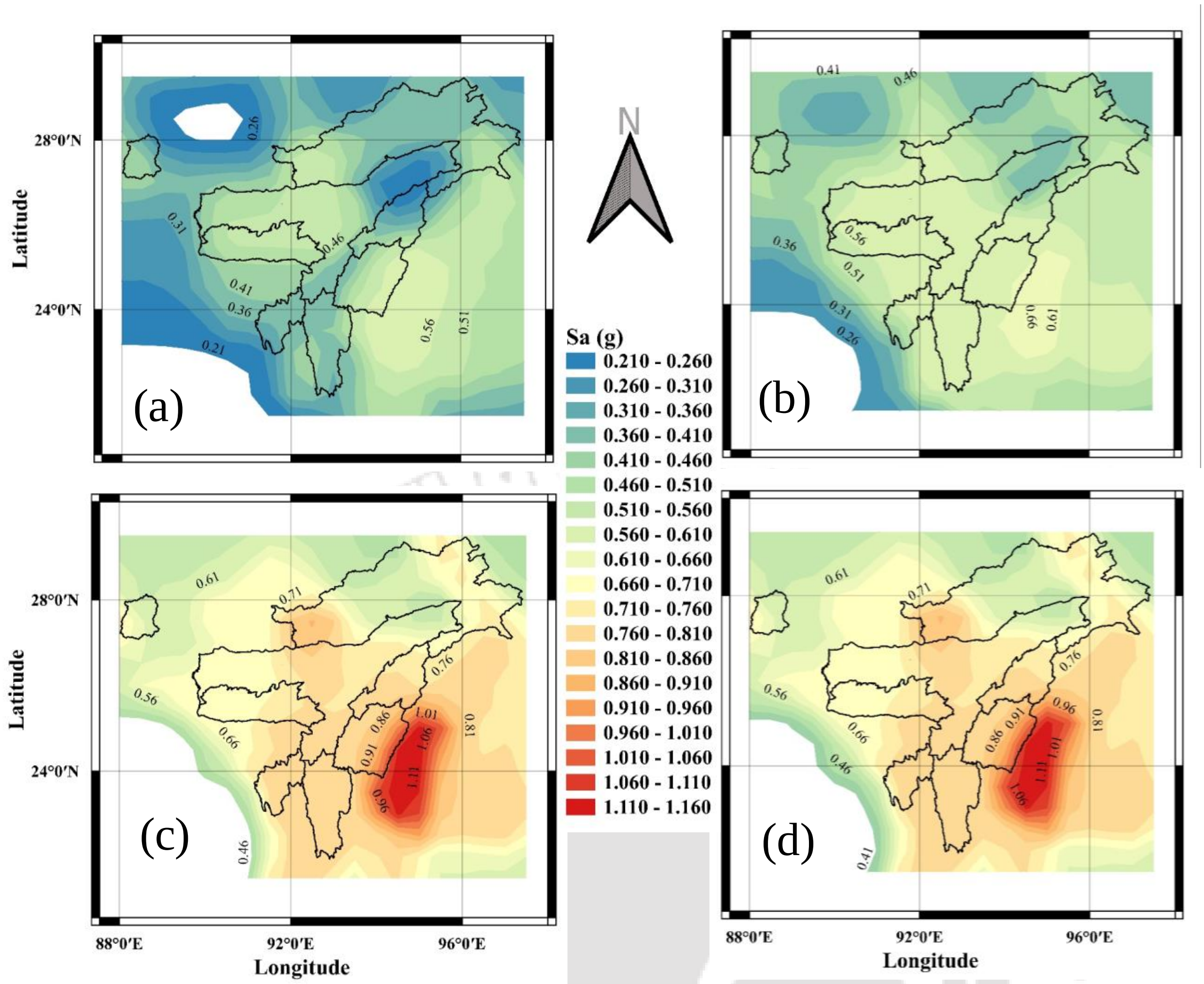

Figure 5.14 Spatial variation in PGA for 2% probability in 50 years for SC A. (a) 5th percentile (b) 16th percentile (c) 84th percentile (d) 95th percentile

Table 5.13 compares DBE-based PGA values for eight cities of NE India obtained based on the present work and previous studies. The hazard curves for PGA values obtained in various Monte Carlo realisations for all eight cities are also shown in Figure 5.15. It is observed that the present findings are slightly different from the findings by NDMA (2010) and Pallav et al. (2012). Further, present values are significantly higher than those works by Das et al. (2016), Baro et al. (2020), Sil et al. (2013) and lower than the studies by Sharma and Malik (2006), Nath and Thingbaijam (2012), Bahuguna and Sil (2018) and Ghione et al. (2021). The spatial mean hazard level distribution of the present study shows comparable results with Sitharam and Kolathayar (2013) and Nath and Thingbaijam (2012). The primary reasons for variations in the seismic hazard values between any two studies are changes in EQ catalogues and the seismicity parameters. The SM2 used in this work, ensured uniform distribution of seismic activity in each source zone. This can be one significant attribute that makes the present findings more reliable than earlier studies.

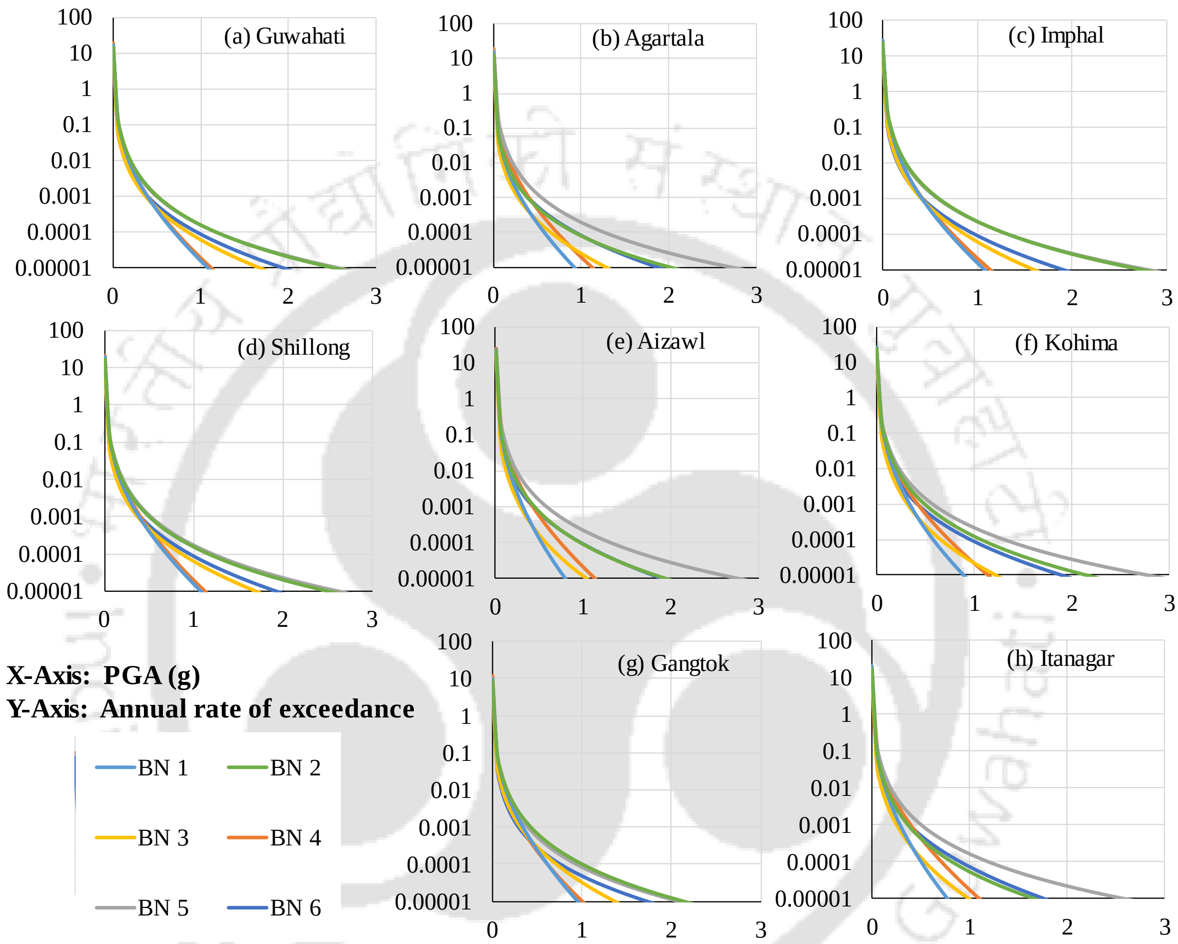

Figure 5.15 Hazard curves for PGA obtained in various MonteCarlo realization / considering different logic tree branches (BN 1 -BN 6) for the eight cities of NE India

Table 5.13 Comparison of DBE based PGA (in g) obtained in the present study with existing study

Cities	Sharma and Malik (2006)	NDMA (2010)	Pallav et al. (2012)	Nath and Thingbaijam (2012)	Sil et al. (2013)	Das et al. (2016)	IS-1893 (2016) part 1	Bahuguna and Sil (2018)	Baro et al. (2020)	Ghione et al. (2021)	Present study (SC A)
Guwahati	0.50	0.23	-	0.66	-	0.24	0.18	0.46	-	0.35	0.31-0.35
Agartala	0.30	0.18	-	0.25-0.30	0.11-0.20	0.22	0.18	-	-	-	0.31-0.35
Imphal	0.40	0.35	0.35-0.5	0.68	-	0.14	0.18	-	-	-	0.39-0.43
Shillong	0.48	0.25	-	0.72	-	0.32	0.18	-	0.19	0.55	0.35-0.39
Aizawl	0.30	0.18	-	0.50-0.55	0.10-0.17	0.11	0.18	-	-	-	0.35-0.39
Kohima	0.50	0.20	-	0.60-0.70	-	0.15	0.18	-	-	-	0.35-0.39
Gangtok	0.35	0.15	-	0.30-0.35	-	0.15	0.12	-	-	-	0.27-0.31
Itanagar	0.44	0.30	-	0.30-0.35	-	0.18	0.18	-	-	-	0.27-0.31

5.13 Deaggregation of EQ Hazard for the Important Cities in NE India

This section explains the observations based on separate deaggregation plots (Figures 16–19) for 8 important cities of NE. Results obtained for all the cities in deaggregation of EQ Hazard is described in the following subsections.

5.13.1 Guwahati

Figure 5.16 (a) and (b) show the 3D deaggregation plot for Guwahati City corresponding to DBE and MCE, respectively. However, to compare the contribution from different magnitudes and distances, 2D deaggregation plots are also developed [Figure 5.17 (a) and Figure 5.18 (a)]. Figure 5.17 (a) and Figure 5.18 (a) show the deaggregation plot (for all the seismic sources combined) for Guwahati city centre (91.74°E, 26.14°N) corresponding to DBE and MCE, respectively. For DBE, EQs in magnitude bin of 6.5-7Mw and distance bin of 25-50km have the highest contribution towards PGA, as can be observed from Figure 5.17 (a). In addition, moderate EQs (Mw-5 to 5.9) with distances up to 100km and major EQs (Mw-7 to 7.9) with distances from 50km to 100km are also contributing towards seismic hazard significantly. Based on the above distance range, the areal sources with higher contributions are Za7, Za8, Za11, Za12, and Zb7. Each of these seismic sources has experienced a large number of strong and major EQs that occurred in the past. Further, results from the deaggregation of seismic hazard in latitude and longitude (Figure 5.19 a), it can be found that faults such as the Atherkhet Fault, Kulsi Fault, Dudhonoi Fault, Kapili Fault, Dighalpani Kakijan Fault, MCT, MBT, MFT, Dauki fault, Oldham Fault, and Barapani are contributing significantly toward hazard level. The epicentre of 1897 Shillong great EQ lies within 200km of Guwahati city. While observing Figure 5.17 (a), it can be concluded that though the contribution from great EQ towards the seismic hazard of Guwahati city centre is significant, it is comparatively lesser than the contribution from strong to major EQs. In other words, EQs with magnitude in the range of moderate, strong and major EQs, happening within 100 km radial distance, are contributing more to the seismic hazard of Guwahati in comparison to great EQs located within 200 km distance. This is primarily due to the fact that the larger is the EQ magnitude, the lesser is the probability of its occurrence within a defined time period.

Figure 5.18 (a) presents deaggregation plots for MCE for Guwahati city centre. It can be observed from Figure 5.18 (a) that EQs in the magnitude range of 7 to 7.5 shows the highest contribution for MCE. In other words, with an increase in the return period, there is an increase in percentage contribution from higher magnitude EQs.

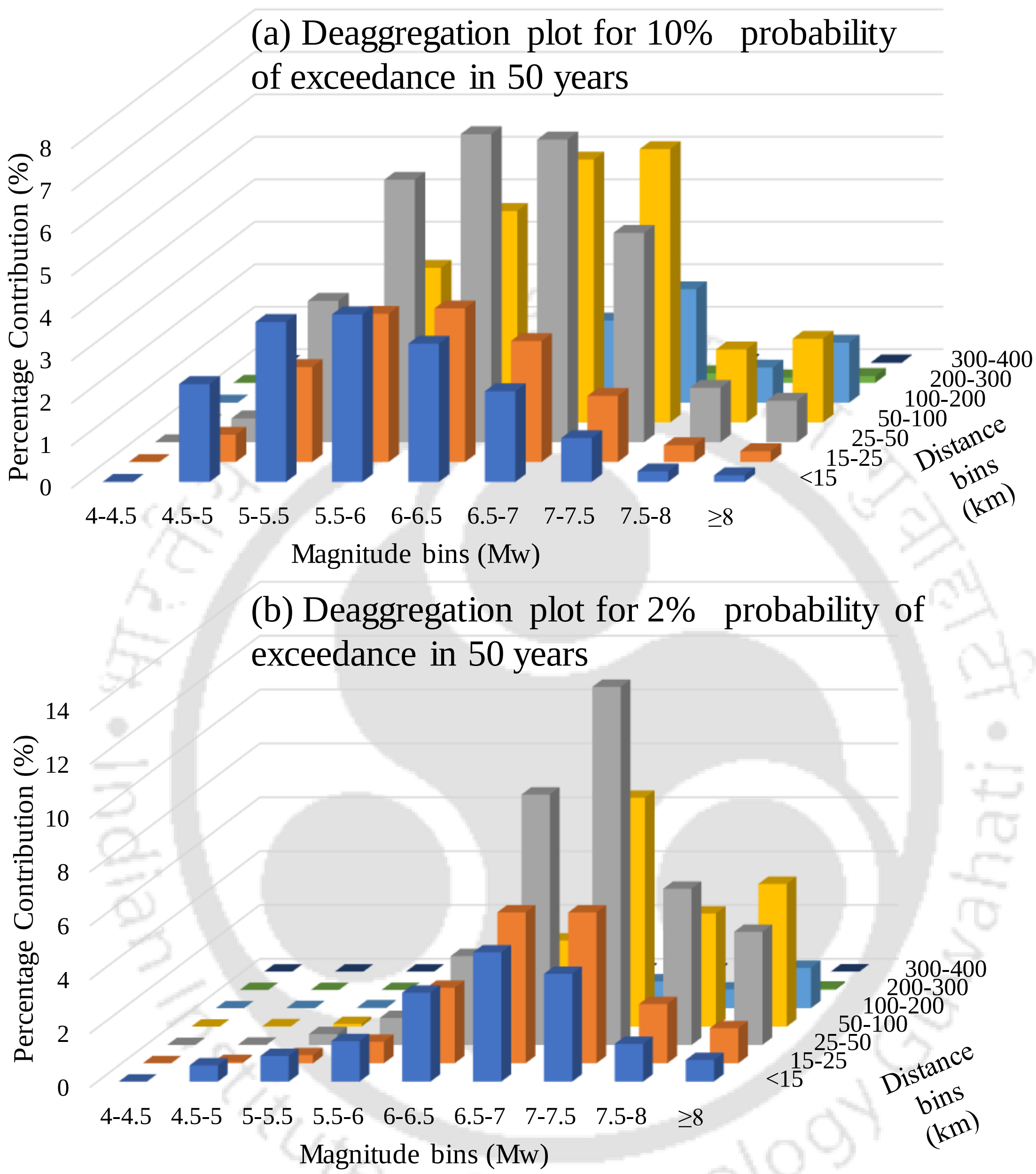

Figure 5.16 3-D Deaggregation plot for Guwahati city considering seismic hazard level for a. 10% probability of exceedance in 50 years b. 2% probability of exceedance in 50 years

5.13.2 Agartala

For Agartala city centre (91.29°E, 23.83°N), Figure 5.17 (b) shows that the highest contribution toward PGA of DBE is from the EQs in the magnitude bin of 6-6.5Mw and distance bin of 25-50km. Contributions from Moderate EQ with distance up to 50 km. Strong EQs with distances up to 100 km. Major EQs with distances from 50 km to 100 km are also significant. Sources with higher contributions are Za11, Za18, Za19, Za20, Za21, Za24, Za25, Zb6, Zb8, and Zb7. As per Figure 5.19 (b), faults such as the Dauki Fault and Sylhet Fault significantly influence the hazard level of the location. These sources have been responsible for several major EQs in the past. However, as per current findings, though the above sources are contributing more to seismic hazard, the role of major EQs on such sources is relatively less in regional seismic hazard. This is primarily due to the longer return period of major EQ.

In the case of MCE (which has a higher return period than DBE), EQs in magnitude bin of 6.5-7Mw and distance bin of 25-50km contribute maximum towards PGA (as shown in Figure 5.18 b). Further, it is collectively observed that the contribution of great EQ is negligible for DBE and MCE for Agartala (see Figure 5.17 b and Figure 5.18 b). It is primarily due to the fact that no great EQ has happened within 200km radial distance of Agartala in the past.

5.13.3 Imphal

Figure 5.17 (c) and Figure 5.18 (c) show the deaggregation plot (for all the seismic sources combined) for Imphal city centre (93.94°E, 24.82°N) corresponding to DBE and MCE, respectively. For DBE, EQs in the magnitude bin of 6-6.5Mw and 25-50km distance bin have the highest contribution (see Figure 5.17 c). In addition, Moderate EQ with distances up to 50 km. Strong EQs with distances from 15 km to 100 km and Major EQs with distances from 25 km to 100 km also significantly contribute to the seismic hazard of Imphal city centre. Figure 5.19 (c) clearly highlights that the contribution from EQs in the Indo-Burma Ranges is maximum. Primarily areal sources Za21, Za20, Za26, and Zb8 contribute the highest to the seismic hazard of Imphal city. Some of the significant faults influencing the hazard level are the Naga Thrust, Disang Fault, and CMF. Past EQ data suggest numbers of strong to major EQ occurrences on these sources. However, the current results suggest that the seismic hazard for DBE is mainly due to strong EQs.

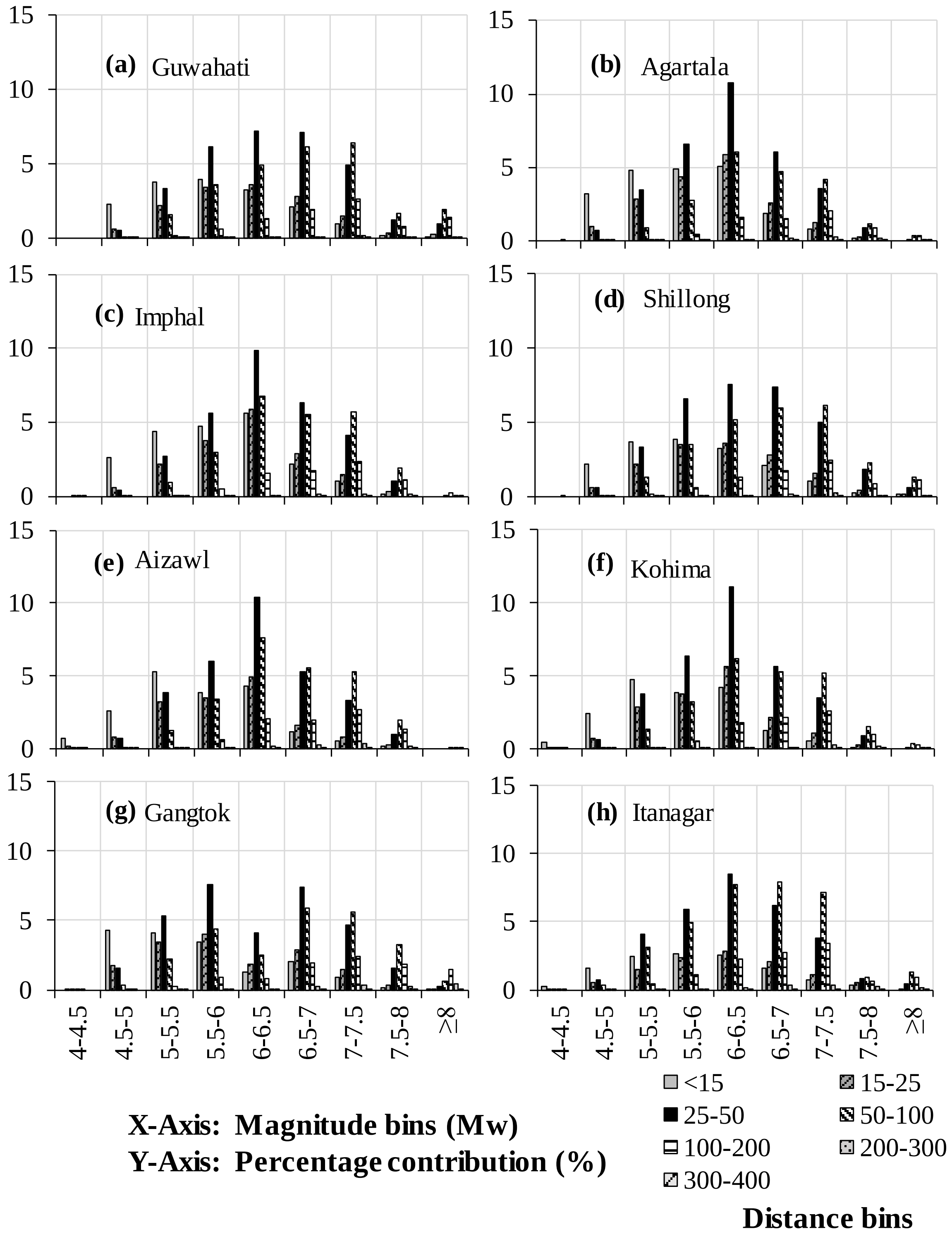

Figure 5.17 Deaggregation plot for eight important cities considering seismic hazard level for 10% probability of exceedance in 50 years

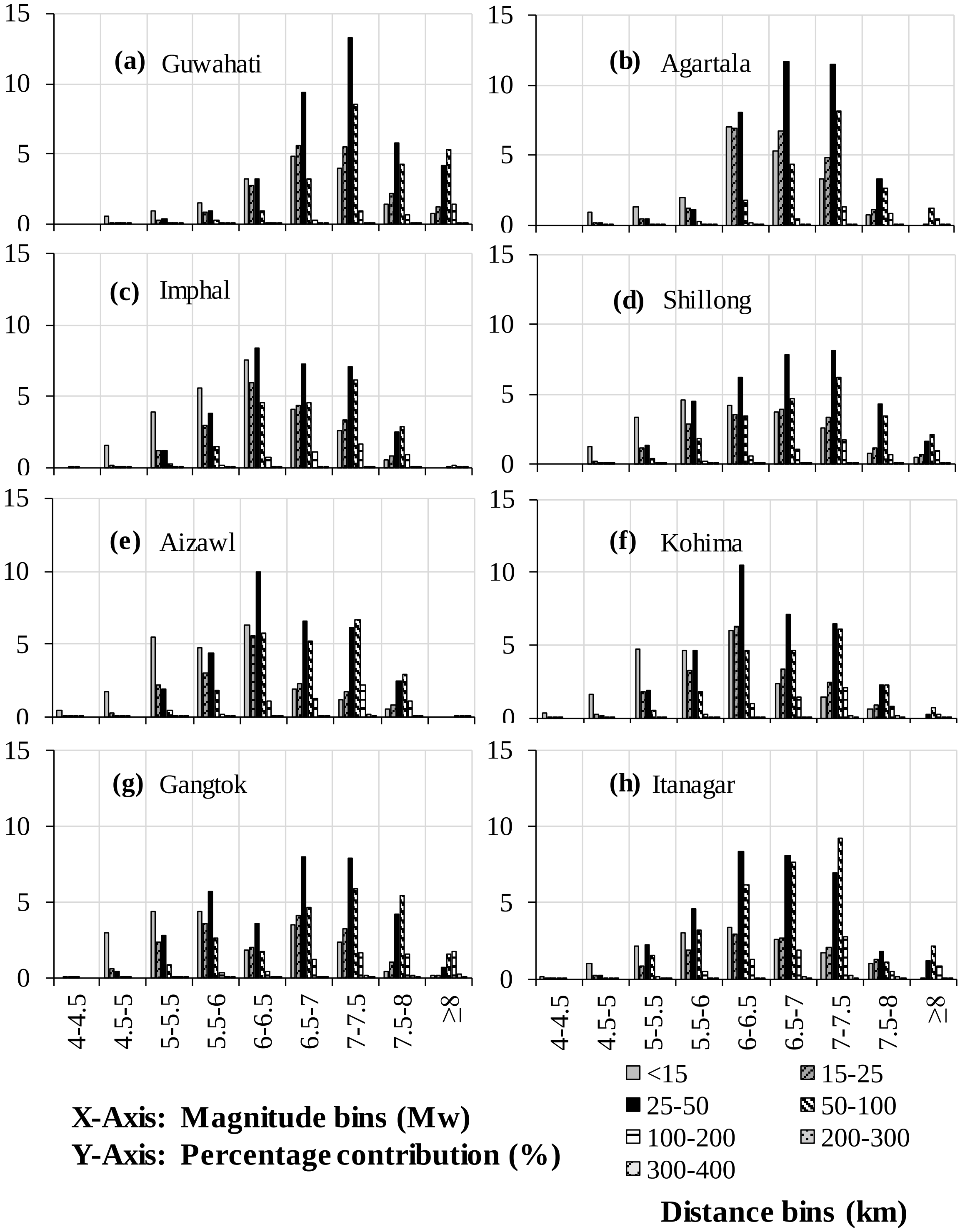

Figure 5.18 Deaggregation plot for eight important cities considering seismic hazard level for 2% probability of exceedance in 50 years

For MCE, EQs in the magnitude bin of 6–6.5 Mw and distance bin of 25–50 km have the highest contribution (Figure 5.18 c). Similar to Guwahati and Agartala, the contribution of major EQs is becoming more significant for MCE for Imphal as well. Further, Moderate EQ with a distance of less than 15 km, Strong EQs with a distance of less than 50 km, and Major EQs with a distance from 25 to 100 km have significant contributions to the PGA of MCE (see Figure 5.18c). The spatial distribution of PGA (as discussed in the last section) suggests that Imphal City has the highest PGA for DBE and MCE. However, there is no contribution from great EQ in the highest PGA value. It should be noted that the closest great EQ occurred more than 250km away from the city. As a result, such great EQ has negligible contribution in controlling the seismic hazard of Imphal city.

5.13.4 Shillong

For Shillong city (91.89°E, 25.58°N), Figure 5.17 (d) suggests that the highest contribution towards PGA of DBE is from EQs in the magnitude bin of 6–6.5Mw and distance bin of 25-50km. In addition, moderate EQ with distances upto 50 km. Strong EQs with distances from 25 km to 100 km and Major EQs with distances from 25 km to 100 km also have significant contributions. Collectively, it can be concluded that strong to major EQs significantly influence the seismic hazard level of the city. Figure 5.19 (d) shows that the areal sources Za11, Za19, Za12, and Zb7 are influencing the seismic hazard level. Based on the available records of past EQs, significant numbers of strong, major EQs and two great EQs have occurred within the above areal sources in the past. However, the contribution of great EQ for PGA of DBE is comparatively less. Based on the deagrregation of hazard in terms of latitude and longitude, it is found that Kulsi Fault, Dudhonoi Fault, Kapili Fault, Barapani Shear, Dauki Fault, Oldham Fault, Eocene Hinge Zone, and Sylhet Fault has significant contribution toward the seismic hazard Figure 5.19 (d).

Based on the deaggregation plots for Shillong city centre corresponding to the PGA of MCE (shown in Figure 5.18 d), it is clear that the highest contribution is from EQs in the magnitude bins of 7-7.5Mw and distance bin of 25-50km. Here, the contribution of major EQs is relatively more than other EQs, which is attributed to a longer return period for MCE in comparison to DBE. Further, the contribution of great EQs is relatively much lesser than major and strong EQs.

5.13.5 Aizawl

Figure 5.17 (e) and Figure 5.18 (e) show the deaggregation plots (for all the seismic sources combined) for the Aizawl city centre (92.72°E, 23.73°N) corresponding to DBE and MCE, respectively. For DBE, EQs in magnitude bin of 6-6.5Mw and distance bin of 25-50km show the highest contribution towards PGA (Figure 5.17e). Besides this, moderate EQs with distances up to 50 km, strong EQs with distances up to 100 km and major EQs with distances from 25 km to 100 km also contribute towards the regional seismic hazard. However, these contributions are comparatively lesser. Further, areal sources Za19, Za20, Za21, Za25, Za31, and Zb8 significantly contribute to Aizawl city's seismic hazard. Figure 5.20 (a) shows that the CMF and the Mat Fault, which were responsible for significant numbers of strong to major past EQs, are mainly contributing to the seismic hazard level of Aizawl city. The above information on magnitudes, distances, and sources suggest that EQs from the Indo-Burma Ranges are mainly contributing to the PGA of DBE. Similarly, Figure 5.18 (e) indicates that for MCE, the maximum contribution is observed from magnitude bin of 6-6.5Mw within distance bin 25-50km. Similar to the other cities discussed earlier, contributions from higher EQs increase significantly for the case of MCE here as well (Figure 5.17 e and Figure 5.18 e).

5.13.6 Kohima

For Kohima city centre (94.11°E, 25.68°N), EQs in magnitude bin of 6-6.5Mw happening within distance bin of 25–50 km have the highest contribution towards PGA of DBE (Figure 5.17 f). In addition, significant but comparatively lesser contributions are also observed from moderate EQs with distance of less than 50 km, strong EQs with distance from <15 km to 100 km and major EQs with a distance from 25 km to 100 km. Based on the above distance ranges, the areal sources responsible are Za12, Za13, Za20, Za21, Za22, Za23, Za26, Zb3, Zb7, and Zb8. Figure 5.20 (b) shows that CMF, Kapili Fault, Disang Thrust, Dighalpani Kakijan Fault and Naga Thrust are the significant faults influencing the hazard level of the region significantly.

The deaggregation plot for Kohima city centre corresponding to the PGA of MCE is shown in Figure 5.18 (f). It can be observed from this figure that EQs in the magnitude bin of 6-6.5 and distance bin of 25-50km have the highest contribution for MCE. Significantly high contributions are also observed from EQs in the magnitudes range of 6.5-7 and 7-7.5, happening in distance bins of <15km, 15-25km, 25-50, and 50-100km.

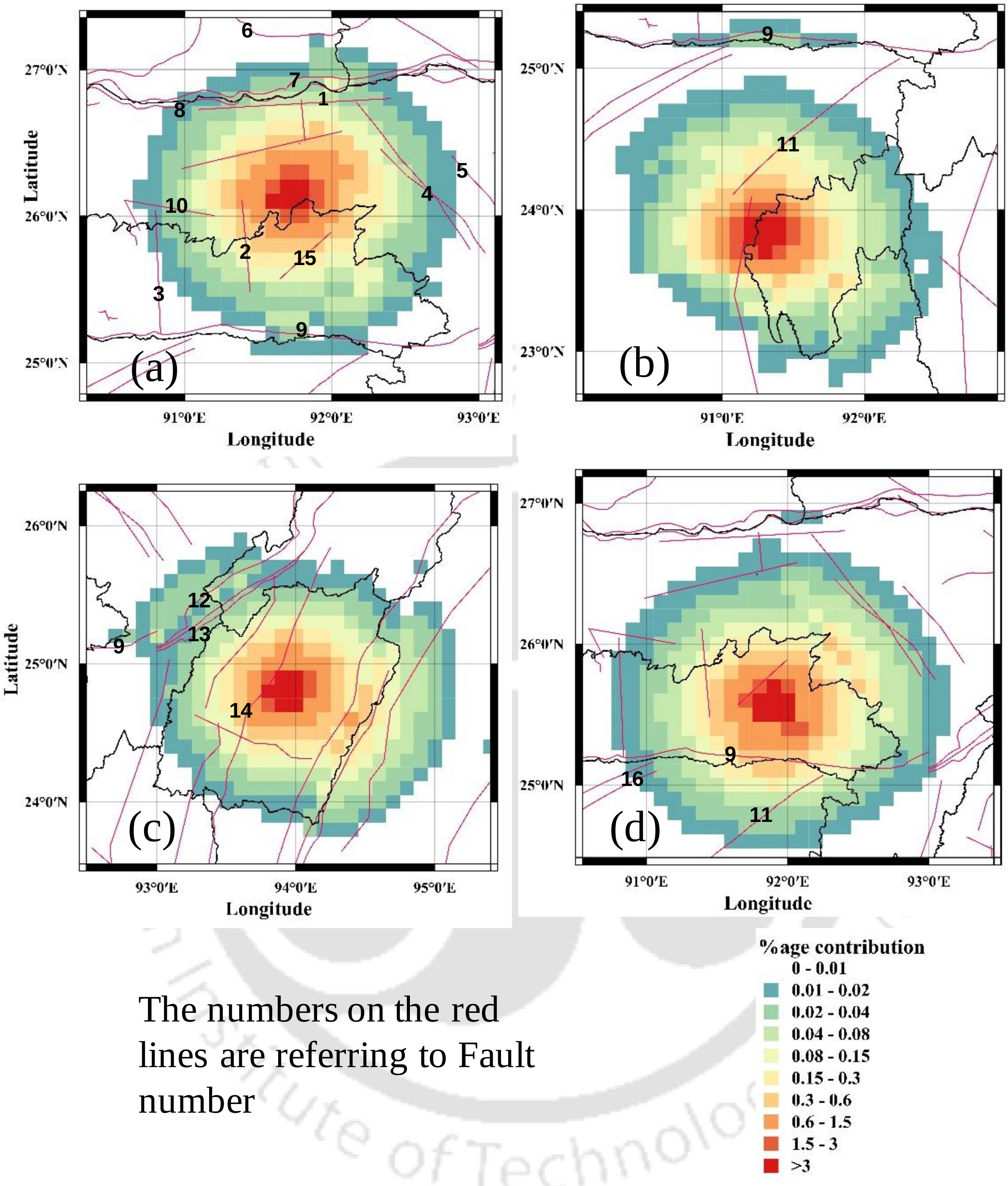

Figure 5.19 Deaggregation in Latitude and Longitude on exceedance of PGA for DBE for (a) Guwahati, (b) Agartala, (c) Imphal, (d) Shillong

5.13.7 Gangtok

The deaggregation plot for Gangtok city centre (88.61°E, 27.33°N) corresponding to DBE and MCE is presented in Figure 5.17 (g) and Figure 5.18 (g), respectively. From Figure 5.17 (g), it is observed that for DBE, EQs in the magnitude bin of 5.5–6 Mw and distance bin of 25-50km have the highest contribution towards the seismic hazard. Further, Moderate EQs with distances up to 50 km, strong EQs with distances from 25 km to 100 km and major EQs with distances from 25 km to 100 km also have significant contributions. This means moderate, strong and major EQs contribute significantly to the seismic hazard. Though great EQs show a significant contribution, but it is comparatively less. Based on the distance bins, EQs, mainly in the Himalayan regions, are responsible for seismic hazards in Gangtok. Areal sources contributing most are; Za1, Za3, Za6, Za7, Za10, and Zb1. Figure 5.20 (c) shows that the contributions from MCT, MBT and MFT are significant.

Figure 5.18 (g) clearly indicates that EQs in the magnitude bin of 6.5–7Mw and distance bin of 25-50km have the highest contribution towards PGA of MCE, followed by EQs Moderate EQ with distance less than 15 km, strong EQs with distances from 15 km to 100 km and Major EQs with distance from 25 km to 100 km. Thus, the contribution of major EQs is comparatively more to the PGA of MCE in the case of Gangtok city. Contributions from higher EQs increase significantly for the case of MCE here as well (Figure 5.16 g and Figure 5.17 g).

5.13.8 Itanagar

For Itanagar city centre (93.61°E, 27.08°N), the maximum contribution towards PGA of DBE is coming from EQs in magnitude bin of 6-6.5Mw, and distance bin of 25–50 km (Figure 5.17 h). Contribution from moderate EQ with distances up to 100 km, strong EQs with distances from 25 km to 100 km and major EQs with distances from 50 km to 100 km is also significant. Out of all the cities mentioned above, only in the case of Itanagar city, EQs in the distance bin of 50-100km have the highest contribution. It is primarily because EQs near Itanagar city are significantly less. In addition, sources nearby have lesser seismic activity. As a result, sources from larger distances influence the seismic hazard of Itanagar significantly. Contributing areal sources are Za8, Za9, Za12, Za13, Zb3, Zb7, and Zb9. Figure 5.20 (d) shows that the Kopili fault, Dighalpani-Kakijan Fault, Naga Thrust, MCT, MBT, and MFT influence the seismic hazard level of the region significantly.

The deaggregation plot for Itanagar city centre for PGA of MCE is shown in Figure 5.18 (h). This figure shows that EQs in the magnitude bin of 7-7.5 and distance bin of 50-100km have the highest contribution to the seismic hazard. Other significant contributions are due to Strong EQs with distance from 25 km to 100 km and Major EQs with distance from 25 km to 100 km.

The results obtained from the deaggregation of seismic hazard for all 8 cities are summarised in Table 5.14

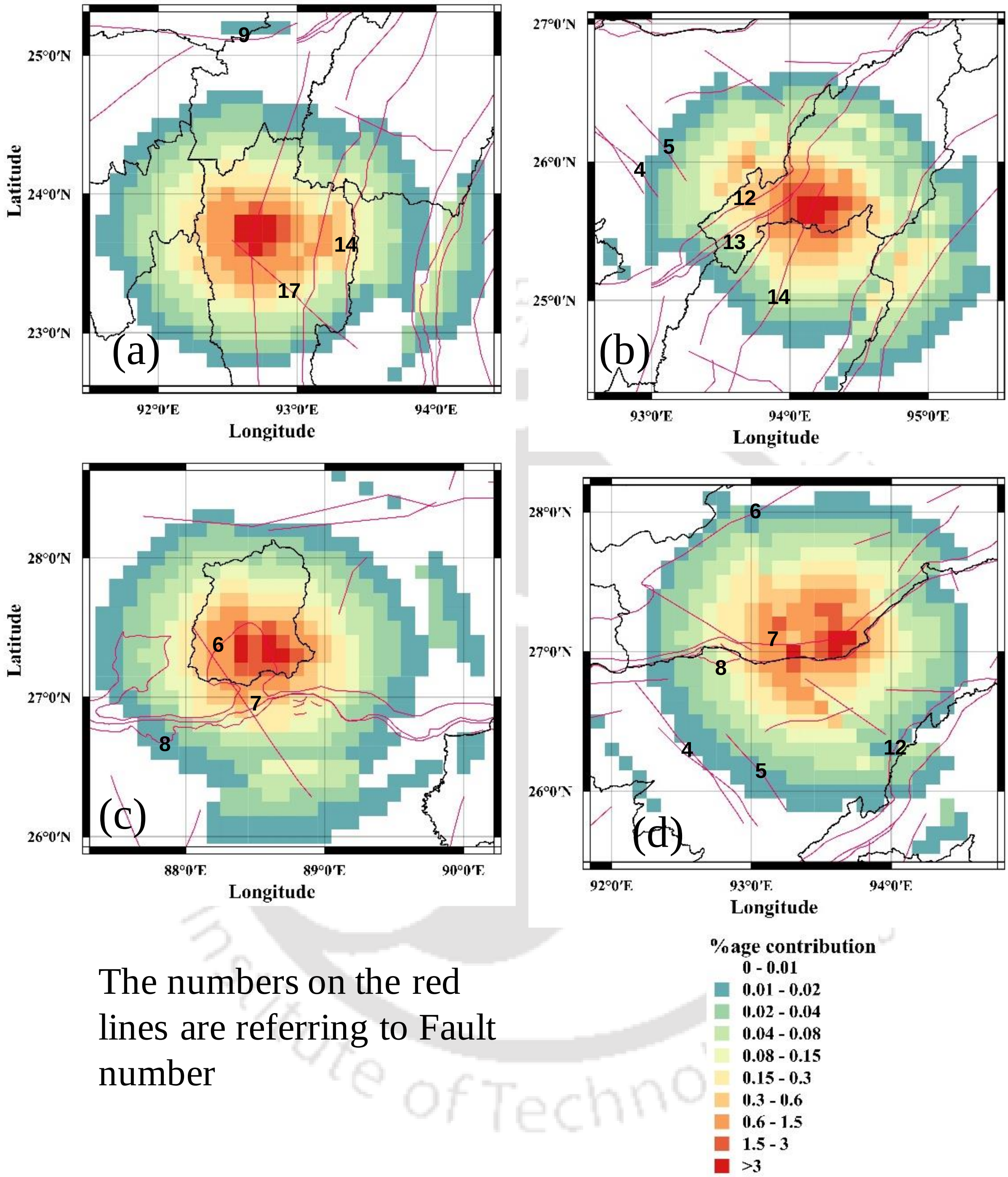

Figure 5.20 Deaggregation in Latitude and Longitude on exceedance of PGA for DBE for (a) Aizawl, (b) Kohima, (c) Gangtok, (d) Itanagar

Table 5.14 Results obtained from seismic hazard deaggregation for the 8 cities

Cities	Hazard level	Combination having highest contribution	EQs having Significant contribution	Areal sources having significant contribution	Linear sources having significant contribution (brackets after the fault name are showing the fault numbers referring to Fig. 18 and 19)	Great EQ's contribution
Guwahati	DBE	(6-6.5Mw) + (25-50km)	Moderate EQ with distance upto 100 km. Major EQs with distance from 50 km to 100 km	Za7, Za8, Za11, Za12, and Zb7	Atherkhet Fault (1), Kulsi Fault (2), Dudhonoi fault (3), Kapili fault (4), Dighalpani Kakijan fault (5), MCT (6), MBT (7), MFT (8), Dauki fault (9), Oldham Fault (10), and Barapani shear (15)	Significant
	MCE	(7-7.5Mw) + (25-50km)	Strong EQs with distance up to 50km. Major EQs with distance upto 100km	same as above	same as above	Significant
Agartala	DBE	(6-6.5Mw) + (25-50km)	Moderate EQ with distance upto 50km. Strong EQs with distance upto 100km. Major EQs with distance from 50km to 100km	Za11, Za18, Za19, Za20, Za21, Za24, Za25, Zb6, Zb8, and Zb7	Dauki fault, and Sylhet fault (11)	Negligible
	MCE	(6.5-7Mw) + (25-50km)	Strong EQs upto 25km. Major EQs with distance upto to 100km	same as above	same as above	Negligible

Imphal	DBE	(6-6.5Mw) + (25-50km)	Moderate EQ with distance upto 50km. Strong EQs with distance from 15km to 100km and Major EQs with distance from 25km to 100km	Za21, Za20, Za26, and Zb8	Naga thrust (12), Disang fault (13), and CMF (14)	Negligible
	MCE	(6-6.5Mw) + (25-50km)	Moderate EQ with distance less than 15km. Strong EQs with distance less than 50 km. Major EQs with distance from 25 to 100km	same as above	same as above	Negligible
Shillong	DBE	(6-6.5Mw) + (25-50km)	Moderate EQ with distance upto 50km. Strong EQs with distance from 25km to 100km and Major EQs with distance from 25km to 100km	Za11, Za19, Za12, and Zb7	Kulsi Fault, Dudhonoi fault, Kapili fault, Barapani shear, Dauki fault, Oldham Fault, Eocene Hinge Zone (16), and Sylhet fault	Significant
	MCE	(7-7.5Mw) + (25-50km)	Moderate EQ with distance less than 15km. Strong EQs with distance from 15km to 100km and Major EQs with distance from 25km to 100km	same as above	same as above	Significant
Aizawl	DBE	(6-6.5Mw) + (25-50km)	Moderate EQ with distance upto 50km. Strong EQs with distance upto 100km and Major EQs with distance from 25km to 100km	Za19, Za20, Za21, Za25, Za31, and Zb8	CMF, Mat Fault (17)	Negligible

	MCE	(6-6.5Mw) + (25-50km)	Moderate EQ with distance less than 15km. Strong EQs with distance up to 100km and Major EQs with distance from 25km to 100km	same as above	same as above	Negligible
Kohima	DBE	(6-6.5Mw) + (25-50km)	Moderate EQ with distance less than 50km. Strong EQs with distance from <15km to 100km and Major EQs with distance from 25km to 100km	Za12, Za13, Za20, Za21, Za22, Za23, Za26, Zb3 Zb7, and Zb8	CMF, Kapili faults, Disang Thrust, Dighalpani Kakijan fault and Naga Thrust	Negligible
	MCE	(6-6.5Mw) + (25-50km)	Moderate EQ with distance less than 15km. Strong EQs with distance from <15km to 100km and Major EQs with distance from 25km to 100km	Same as above	Same as above	Negligible
Gangtok	DBE	(5.5-6Mw) + (25-50km)	Moderate EQs with distance upto 50km, Strong EQs with distance from 25km to 100km and Major EQs with distance from 25km to 100km	Za1, Za3, Za6, Za7, Za10, and Zb1	MCT, MBT and MFT	Significant
	MCE	(6.5-7Mw) + (25-50km)	Moderate EQ with distance less than 15km. Strong EQs with distance from 15km to 100km and Major EQs with	same as above	same as above	Significant

			distance from 25km to 100km			
Itanagar	DBE	(6-6.5Mw) + (25-50km)	Moderate EQ with distance upto 100km. Strong EQs with distance from 25km to 100km and Major EQs with distance from 50km to 100km	Za8, Za9, Za12, Za13, Zb3, Zb7, and Zb9	Kopili fault, Dighalpani-Kakijan fault, Naga Thrust, MCT, MBT, and MFT	Significant
	MCE	(7-7.5Mw) + (50-100km)	Strong EQs with distance from 25km to 100km and Major EQs with distance from 25km to 100km	same as above	same as above	Significant

5.14 Summary

The Present PSHA work suggests that the eastern region of Manipur and the southern part of Nagaland have the highest seismic hazard levels within NE India, both in the case of DBE and MCE. Similarly, Sikkim and parts of Arunachal Pradesh and Assam have the lowest seismic hazard values. Thus, based on the PGA distribution for NE, the highest PGA values have been found near the Indo-Burma ranges, followed by the Bhutan-Himalaya and then the Shillong Plateau. These areas had a high density of previous EQs and greater magnitude EQs. The lower CoV values (<0.4) in the case of PGA determination in the present work represent a low variability in the estimated PGA values throughout NE India.

The deaggregation results presented in Figure 5.17, Figure 5.18, Figure 5.19, and Figure 5.20 suggest that depending on the position of the above cities with respect to important sources, contributions from EQs in different magnitude and distance bins are different. These contributions mainly depend on the past seismic activity around the cities. Due to this reason, different cities in NE India show significantly varying seismic hazard values. In addition, current results also show that strong EQs play a dominant role in controlling PGA of DBE and MCE for the majority of the cities except for Itanagar. For Itanagar City, however, the contribution from major EQs, which had happened at larger distances, is the most. Further, for all the cities, the contributions from larger magnitudes is more in the case of MCE than DBE. This is primarily due to the fact that MCE has a larger return period in comparison to DBE.

It is also observed from the present results that the contribution of the Great EQs is significant (>1%) for Guwahati, Sikkim, Shillong, and Itanagar (Figure 5.17, and Figure 5.18) primarily because these cities are close to past great EQs. Other cities, however, are either located far away from the past great EQs or have a larger number of strong to major EQs in nearby regions, which overshadow the contribution of the great EQs. Such contribution from great EQs may increase with the increase in the return periods. The present study also finds that irrespective of the magnitudes, EQs in distance bins >200km have a negligible contribution to seismic hazard in DBE and MCE. Further, the contribution of EQs with magnitudes less than 5.5Mw is significantly lower.

Chapter 6. Seismic Hazard Analysis (SHA) of NE India by Utilising Kernel Density Method to Deal with the Uncertainty in Location and Magnitude of Past EQs

6.1 Background

SHA work described in Chapter 5 is based on the traditional PSHA methodology, which considers that the seismic activity is uniform throughout the entire seismic source (independent of whether it is an areal or linear source). However, in reality, it is seen that regardless of the approach employed to define seismic source zones, seismic activity is not evenly distributed across the whole source zone (Molchan et al. 1997; Ramanna and Dodagoudar 2012; Rout et al. 2018). Seismic activity can vary significantly throughout a fault or areal source, leading to localised regions of higher or lower seismicity. This poses a significant challenge to SHA, as traditional PSHA methods may not accurately capture it. Traditional PSHA methodology uses a recurrence relation (primarily the Guttenberg-Richter recurrence relation) to model the seismic hazard, which assumes that EQs occur periodically on a fault/ source zone. Development and application of the recurrence relation to a small area is doubtful due to the absence of an adequate number of EQ events. This limitation in the traditional PSHA can be overcome by using some new methods or techniques, such as zone-free or spatial smoothing techniques like the Kernel density method (Ramanna and Dodagoudar 2012; Rout et al. 2018). The Kernel density method is a method of estimating the probability density function from a data set (EQ data), which is point-based and not zone-based (Ramanna and Dodagoudar 2012).

Thus, the method can properly capture the non-uniform distribution of seismicity within a location by considering suitable grids. Besides, the Kernel density based method replaces the traditional Guttenberg-Richter recurrence relation with a spatially smoothened seismic activity rate probability density function. In the present work, the Kernel density method is utilised to perform PSHA for NE India. Besides using the EQ data, the present work utilises the uncertainty related to the magnitude and location of the EQs and the fault information of the region. This Chapter presents the entire procedure adopted for Kernel based PSHA for the NE India region in this thesis.

6.2 Utilisation of Kernel Density for SHA

The Kernel density method is a type of data smoothing method. Many researchers have implemented this method to forecast the seismicity rate of a particular magnitude at a site of interest (Woo 1996; Chan et al. 2010; Chan 2016; Rout et al. 2018). Based on the magnitude and distance of EQs, the mean annual cumulative rate of seismicity or seismicity rate density, $\lambda(\mathrm{M}, \mathrm{st})$ at the site of interest (st) can be found out by the following equation (Woo, 1996):

$$\lambda(M, st) = \sum_{i=1}^{N_M} \frac{K(M, r_i)}{T_M} \qquad 6.1$$

In the above equation, $\mathrm{K}(\mathrm{M}, \mathrm{r_I})$ represents a smoothing Kernel. $\mathrm{K}(\mathrm{M}, \mathrm{r_i})$ is a function of M (ie. the magnitude of the the i^{th} EQ) and the distance ($\mathrm{r_i}$) between the i^{th} EQ and the site of interest. Further, $\mathrm{T_M}$ is the period of completeness for magnitude threshold M and $\mathrm{N_M}$ is the total number of EQs with magnitudes larger than M. The smoothing Kernel function $\mathrm{K}(\mathrm{M}, \mathrm{r_I})$ can be obtained by using the following equation (Woo 1996; Ramanna and Dodagoudar 2012).

$$K(M, r_i) = \frac{PL-1}{\pi H^2(M)} \times \frac{1+\delta \cos^2 \emptyset}{1+\frac{\delta}{2}} \times \left[1 + \left(\frac{r_i}{H(M)}\right)^2\right]^{-PL} \qquad 6.2$$

Here, PL, also known as the “kernel fractal scaling index”, is the power law index (Molina et al. 2001). Molina et al. (2001) recommend a PL between 1.5 and 2.0 corresponding to a cubic or quadratic decay of the probability density function with epicentral distance. However, Chan et al. (2010) found that the differences between the results are insignificant when the PL is between these values. Thus, based on Chan (2016), an intermediate value of 1.75 is considered for the present analysis.

δ represents the degree of anisotropy. For an isotropic condition, the value of δ is considered as zero. Similarly, a higher value of δ refers to a significant anisotropic condition. The parameter $\emptyset$ (see Figure 6.1) is the angle subtended at the site of interest between the epicentre location and the intersection of the fault plane with the Earth's surface (Ramanna and Dodagoudar 2012). So, the parameter has a different value for each location of interest and EQ epicentre.

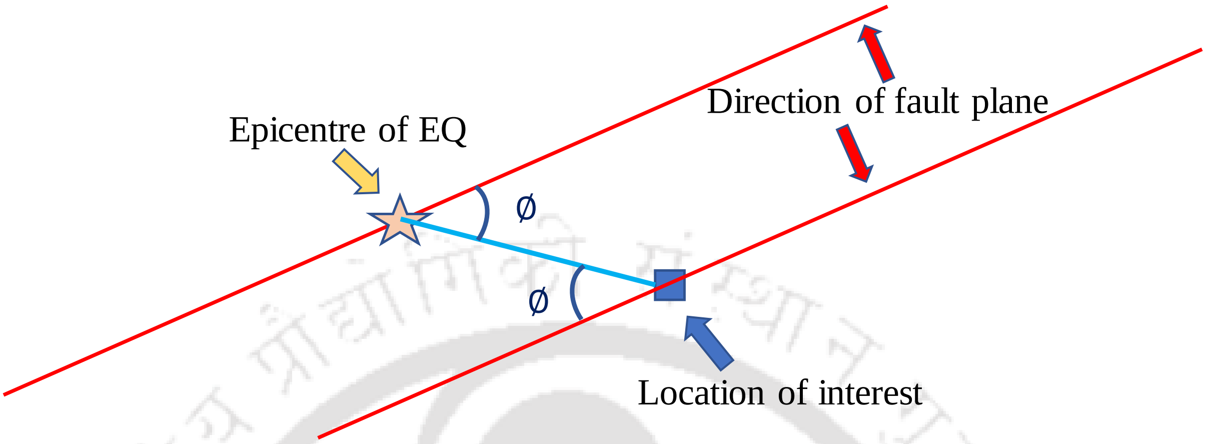

Figure 6.1 Figure showing the parameter Ø that is used in the Kernel function

Term $H(M)$ in Eq. 6.2 represents the bandwidth function, which is a function of M, as in Eq. 6.2 (Woo 1996). This is the most important term in Kernel density, which can significantly influence the output. . It is estimated based on the nearest distance of an event to other events for each magnitude bin, M_b. Based on the standard form of the logarithmic correlation between fault length and magnitude, (Woo 1996) suggested using the following functional form to determine $H(M)$.

$$H(M) = c \times e^{d \times M} \qquad 6.3$$

Where c and d are region-specific regression constants. To determine these parameters, various magnitude bins are created first, and then, for each EQ in the bin, the distance to the nearest EQ is determined. From these nearest distances, the mean nearest distance for each bin is determined. Then c and d values are obtained from regression analysis of the natural logarithm of the mean distance of the nearest event vs magnitude plot.

6.2.1 PSHA methodology

In the standard PSHA method, the following equation is used for finding the annual frequency of exceedance (v_z) of a ground motion level (z) (EM-1110-2-6050, 1999):

$$v_z = \sum_{n=1}^{N} \sum_{m_i=m^0}^{m_i=m^u} \lambda_n(m_i) \left[\sum_{r_j=r_{min}}^{r_j=r_{max}} P_n(R = r_j|m_i) P(Z > z|m_i r_j)\right] \qquad 6.4$$

In the above equation, v_z means the annual frequency of exceedance of a hazard level z (mainly in PGA and Sa), $\sum_N$ denotes the sum of results obtained from N seismic sources, $\lambda_n(m_i)$

stands for the annual frequency at which a magnitude m_i occurs on the n^{th} seismic source, $P_n(R = r_j|m_i)$ represents the probability of an EQ with magnitude m_i occurring at a distance r_j on the n^{th} source. $P(Z > z|m_i r_j)$ is the conditional probability of exceedance of hazard level z, considering that the magnitude of the EQ is m_i which is located at a distance r_j.

In the Kernel method (Ramanna and Dodagoudar 2012), to use the mean annual cumulative rate of seismicity, $\lambda(M, st)$ one need to change Eq. 6.4 significantly. Firstly, the entire region should be divided into grids with specific dimensions (following Rout et al. 2018). The finer grid can better capture the spatial distribution of $\lambda(M, st)$ values. As $\lambda(M, st)$ gives the density of exceedance (or the seismicity rate density) of an EQ of a certain magnitude in terms of numbers per year per sq. km, the rate of exceedance in that particular grid needs to be found out. This can be done by multiplying $\lambda(M, st)$ by the area of the grid (A_{st}). The present study considers 0.1°×0.1° of grid sizes. These grids will also be considered as the seismic sources for SHA. The annual rate of occurrence of a particular magnitude m_i in these grids can be found out by using the following equation (based on Rout et al. 2018 and EM-1110-2-6050, 1999):

$$\lambda_o(m_i, st) = \{\lambda(m_i - \Delta m/2, st) - \lambda(m_i + \Delta m/2, st)\} \times A_{st} \qquad 6.5$$

Where, Δm is the difference between the two subsequence magnitudes considered for finding the seismic hazard. In the present work, Δm to be considered as 0.5Mw. As the term $\lambda_o(m_i, st)$ gives the annual rate of occurrence of EQ of magnitude m_i at the grid source (st), this term can directly replace the term " $\lambda_n(m_i)$" and "$P_n(R = r_j|m_i)$" term in Eq. 6.5. Thus, the final equation for PSHA using Kernel density can be written as

$$v_z = \sum_{st=1}^{N} \sum_{m_i=m^0}^{m_i=m^u} \lambda_o(m_i, st) \times P(Z > z|m_i r_{st}) \qquad 6.6$$

where, r_{st}= the hypocentral distance between the site of interest and the grid source $P(Z > z|m_i r_{st})$= probability of exceedance of ground motion level z for an EQ of magnitude m_i and located at the grid source *st* (hypocentral distance is r_{st}). The term $P(Z > z|m_i r_{st})$ is calculated similarly to the term $P(Z > z|m_i r_j)$ in Eqs. 5.12 and 5.13, as described in Section 5.6.3 in Chapter 5. As mentioned in Chapter 5, three GMPEs namely: Toro (2002), NDMA (2010), and Anbazhagan et al. (2013) are used for the present work.

6.3 EQ Catalogue Used for the Analysis

The EQ catalogue used in the present study consists of EQs from 825 AD to 2020 AD. The information on the past EQs occurred between the latitude of 18°N and 32°N and the longitude of 85° E and 100°E are obtained from the USGS, NCS, IMD, ISC, and NDMA (2010). In addition, the EQ catalogue also contains some paleo EQs. Table 6.1 shows the past great paleo EQs that occurred within and around NE India with their time of occurrence, locations, and magnitudes. These events are also shown in Figure 6.2 on the map of NE India.

Table 6.1 Paleo EQ occurred in and around of NE India

Year of Occurrence	Magnitude	Location	Approximate coordinate considered	References
825AD-835 AD	>8Mw	Near Guwahati - southern Brahmaputra valley	(91.8°E,26.1°N)	Rajendran et al. (2004)
1100AD	8.4-8.8Mw	Nepal Region	(85°E,27.5°N)	Lavé et al. (2005)
1255AD	8.6-9Mw	Nepal region	(86°E, 27.07°N)	Mishra et al. (2016)
1663AD	8	Assam	(92°E, 22°N)	Alam and Dominey-Howes (2016)
1697AD	7.7-8.1Mw	Arunachal Pradesh	(94.6°E, 27.75°N)	Pandey et al. (2021)
1714AD	7.5-8.5Mw	Bhutan Region	(90.08°E,27°N)	Hetényi et al. (2016), Zhao et al. (2021)
1762 AD	8	Bangladesh: Chittagong	(92°E,22°N)	Alam and Dominey-Howes (2016)

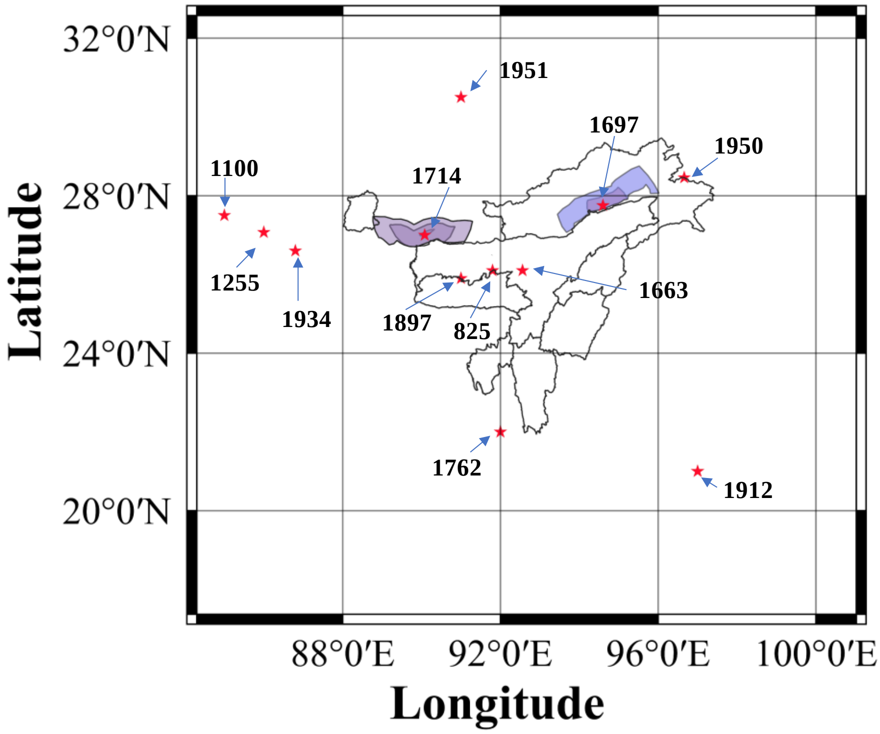

Figure 6.2 Location of Great EQs considered in the present study along with the year of occurrence (The coloured area is the possible location for the 1697AD and 11714AD EQs based on Pandey et al., 2021; Hetényi et al. 2016; and Zhao et al. 2021)

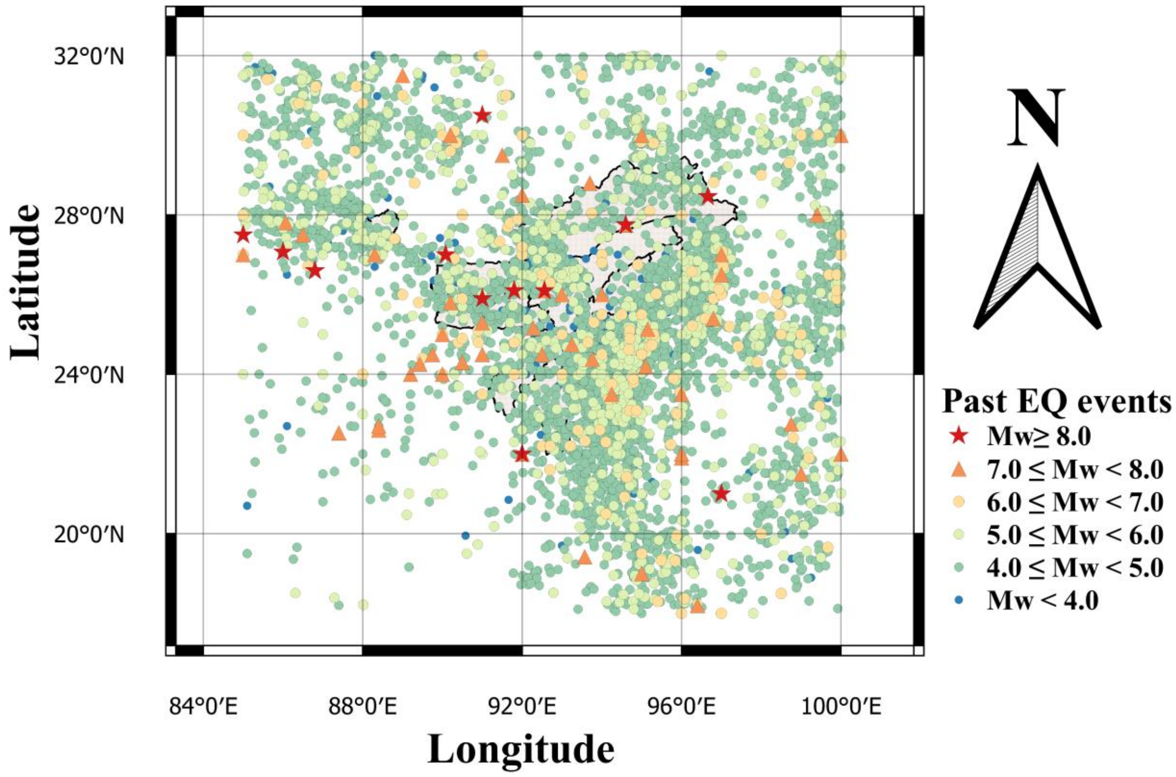

Figure 6.3 Past EQ events in and around NE India

After collecting the EQ information, homogenisation, removal of the duplicate events and declustering are carried out following the methods described in Chapter 3. For the declustering Uhrhammer (1976) method is used. After the declustering, 5876 numbers of the EQs remain in the catalogue, out of which 215 events have magnitude, Mw < 4.0, 4628 events have magnitude range 4.0 ≤ Mw < 5.0, 813 events have magnitude range 5.0 ≤ Mw < 6.0, 159 events have magnitude range 6.0 ≤ Mw < 7.0, 49 events have magnitude range 7.0 ≤ Mw < 8.0 and 12 events have magnitude, Mw≥ 8.0 (as shown in Figure 6.3).

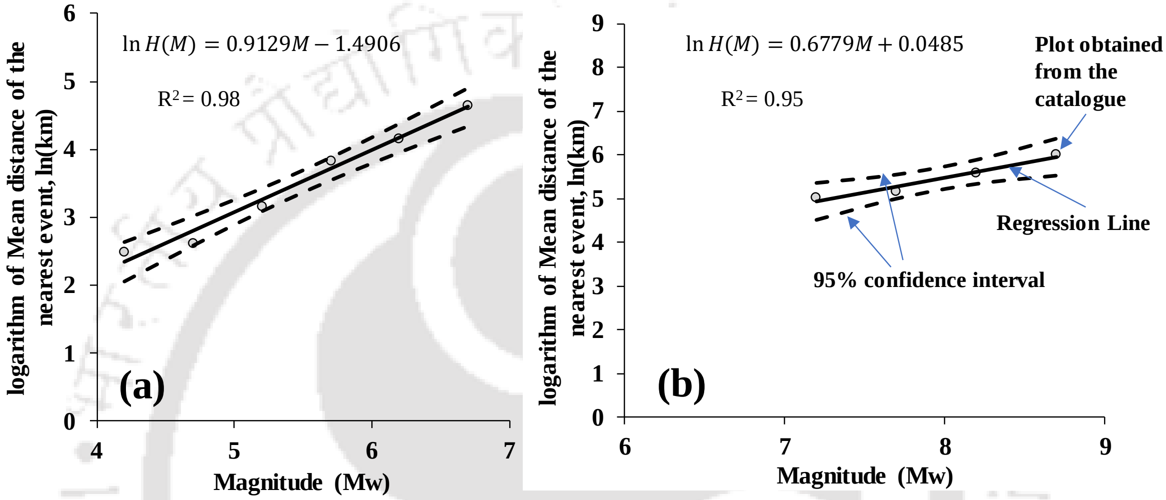

Figure 6.4 The derived bandwidth functions together; (a) for Z1 and (b) For EQs having magnitude greater than 7Mw for the entire NE India region

6.4 $H(M)$ Determination for Different Source Zones in NE India

As the bandwidth function parameters c and d are region-specific, their values differ for different parts of NE India. The present study considers the source zone model (see Figure 4.4) developed in Chapter 4. In this source zone model, there are 12 seismic source zones (Zone 1 or Z1 to Zone 12 or Z12). These source zones are based on the distribution of past EQ events. c and d values are calculated for each source zone (total 12). For each seismic source zone, the EQs are separated into magnitude bins considering a difference in magnitude of ΔM=0.5Mw in each bin. The used magnitude bins are 4≤Mw<4.5, 4.5≤Mw<5.0, 5.0≤Mw<5.5, 5.5≤Mw<6.0, 6.0≤Mw<6.5, 6.5≤Mw<7.0, 7.0≤Mw<7.5, 7.5≤Mw<8.0, 8.0≤Mw<8.5 and 8.5≤Mw. The shortest distance between the events is first evaluated for the EQ events present in each magnitude bin, and then the logarithm of the average/ mean shortest distance ($\ln H$) vs. magnitude (M) is plotted. From the plot, c and d values are determined based on the regression analysis. Figure 6.4 (a) shows the linear regression between the logarithm of the mean shortest distance (ln H) vs. magnitude (M) for zone Z1. As major and great EQs influence a larger area, the c and d are estimated separately for major and great EQs by considering the EQs of Mw≥7 occurred within the whole NE India (Figure 6.4 b). Table 6.2 presents the c and d values obtained for all the source zones.

Table 6.2 Region-wise values of c and d

Source zones	*c*	*d*
Z1	0.2252	0.9129
Z2	0.3013	0.9427
Z3	0.2254	0.9546
Z4	0.571	0.7942
Z7	0.2693	0.8525
Z8	0.0831	1.014
Z9	0.1519	0.9308
Z11	0.2503	0.9023
Z12	0.6835	0.8012
Z5+Z6+Z10	4.5208	0.5658
EQs > 7Mw	1.2021	0.6779

6.5 Completeness Analysis of the Catalogue

For the catalogue, T_M is found out by applying the CUVI method which is a graphical method (Mulargia and Tinti 1985). Figure 6.5 shows typical plots of the cumulative number of EQs vs year for Z1. Figure 6.5 (a) shows stable pattern in the occurrence of EQs after the year 1972, which was absent for the EQs that occurred before 1972. Thus, for the magnitude class 4≤Mw≤4.5, 1972 is the year after which the EQs that occurred in Z1 are complete. Accordingly, the T_M is 49 years (1972-2020). Similarly, for magnitude classes 4.5≤Mw<5.0, 5.0≤Mw<5.5, 5.5≤Mw<6.0, 6≤Mw<6.5, and 6.5≤Mw<7.0 magnitude classes, the years of completeness are 1964 (T_M=57 years), 1963 (T_M=58 years), 1923 (T_M=98 years), 1899 (T_M=122 years) and 1833 (T_M=188 years) respectively. Figure 6.6 depicts the T_M=years of completeness for EQs having a magnitude greater than 7Mw. For magnitude classes 7≤Mw<7.5, 7.5≤Mw<8.0 and 8.0≤Mw, the years of completeness are obtained as 1842 (T_M=179 years), 1762 (T_M=259 years), and 1663 (T_M=358 years) respectively. A summary of T_M for various magnitude classes are presented in Table 6.3.

Two types of catalogues are considered in the present study for SHA. For the first type of catalogue, only declustered EQs, which are complete in terms of magnitude and duration, are used (Catalogue 1). In the second type of catalogue (Catalogue 2), all EQs with a Mw≥7Mw are considered. Additionally, EQs with a magnitude less than 7Mw, which are complete with respect to magnitude and time, are kept in Catalogue 2.

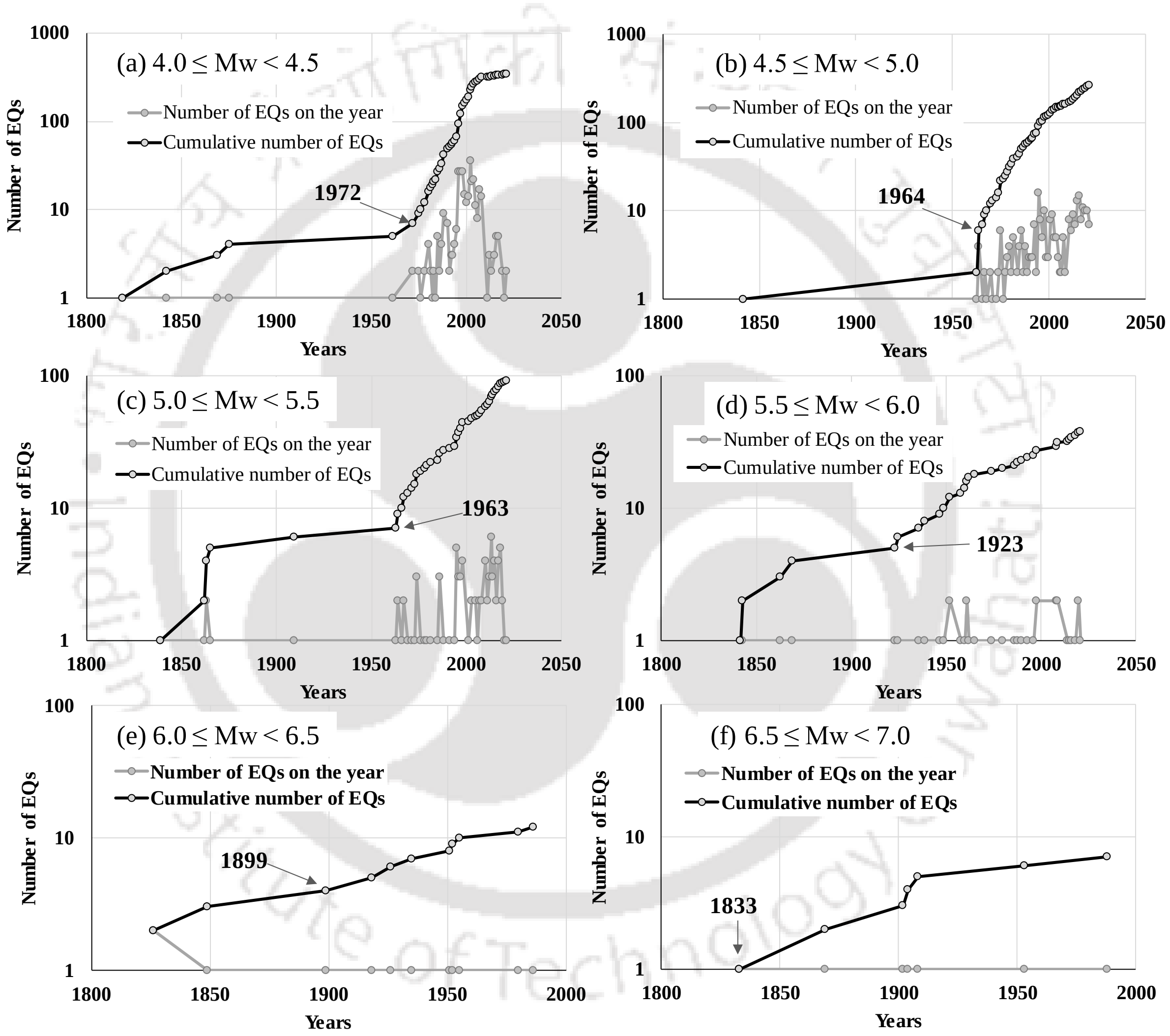

Figure 6.5 Years of completeness for Z1 determined based on CUVI method

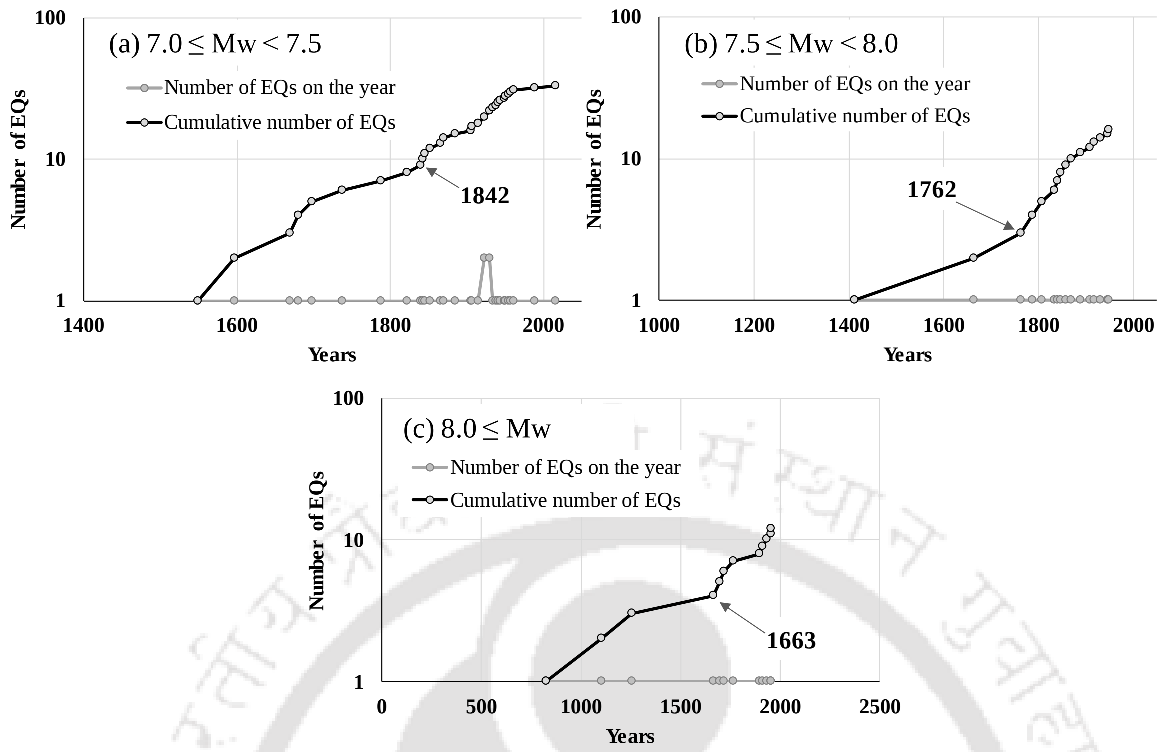

Figure 6.6 Years of completeness for EQs having magnitude greater than 7Mw obtained based on CUVI method

Table 6.3 Summary of T_M (in years) for various zones and various magnitude range

<table>
<tr><th rowspan="2">Source Zone</th><th colspan="9">T_M</th></tr>
<tr><th>4.0 ≤ Mw < 4.5</th><th>4.5 ≤ Mw < 5.0</th><th>5.0≤ Mw < 5.5</th><th>5.5 ≤ Mw < 6.0</th><th>6.0 ≤ Mw < 6.5</th><th>6.5 ≤ Mw < 7.0</th><th>7.0≤ Mw<7.5</th><th>7.5≤ Mw<8.0</th><th>8.0≤ Mw</th></tr>
<tr><td>Z1</td><td>49</td><td>57</td><td>58</td><td>98</td><td>122</td><td>188</td><td rowspan="10">179</td><td rowspan="10">259</td><td rowspan="10">358</td></tr>
<tr><td>Z2</td><td>32</td><td>49</td><td>53</td><td>97</td><td>86</td><td>119</td></tr>
<tr><td>Z3</td><td>53</td><td>59</td><td>57</td><td>96</td><td>83</td><td>--</td></tr>
<tr><td>Z4</td><td>59</td><td>55</td><td>57</td><td>91</td><td>98</td><td>--</td></tr>
<tr><td>Z7</td><td>42</td><td>56</td><td>63</td><td>89</td><td>78</td><td>106</td></tr>
<tr><td>Z8</td><td>49</td><td>56</td><td>62</td><td>97</td><td>101</td><td>94</td></tr>
<tr><td>Z9</td><td>59</td><td>60</td><td>58</td><td>97</td><td>96</td><td>107</td></tr>
<tr><td>Z5+Z6 +Z10</td><td>43</td><td>45</td><td>41</td><td>94</td><td>257</td><td>--</td></tr>
<tr><td>Z11</td><td>46</td><td>56</td><td>57</td><td>75</td><td>79</td><td>118</td></tr>
<tr><td>Z12</td><td>39</td><td>36</td><td>40</td><td>37</td><td>29</td><td>80</td></tr>
</table>

*-- Not determined as no EQ is available in the bin

6.6 Kernel Density

In this work, two types of Kernels are used for the determination of spatial distribution of $\lambda(\mathrm{M}, \mathrm{st})$ values in NE India. The two Kernels are the isotropic Kernel and the anisotropic Kernel. The EQ catalogue event information and $\mathrm{T_M}$ values are utilised for the determination of $\lambda(\mathrm{M}, \mathrm{st})$ values for different EQ magnitudes. The two types of Kernels are described in the following subsections.

6.6.1 Isotropic Kernel

For the isotropic Kernel, only information on past EQ events is utilised. δ value is considered as 0 in Eq. 6.2 to obtain the isotropic Kernel density using the following formula:

$$K(M, r_i) = \frac{PL-1}{\pi H^2(M)} \times \left[1 + \left(\frac{r_i}{H(M)}\right)^2\right]^{-PL} \quad 6.7$$

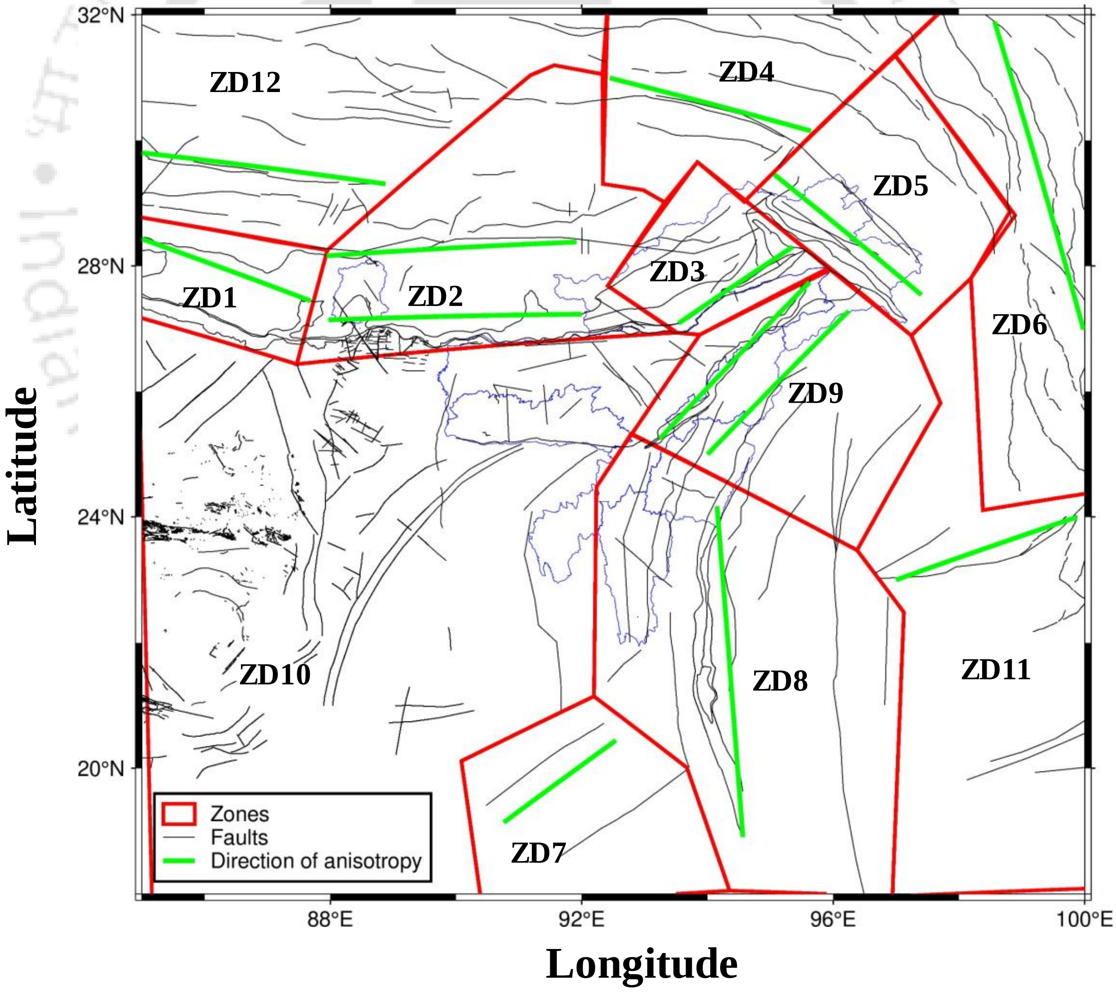

Figure 6.7 Faults and direction of anisotropy in NE India

6.6.2 Anisotropic Kernel

In the anisotropic Kernel, information of Faults/ lineaments is also utilised along with information of past EQ events. The anisotropy's direction is determined based on the fault alignment (Figure 6.7). Available fault information for NE India is obtained from Bhukosh, Kundu and Gahalaut (2013) and Chen (2020). Based on the orientation of the faults, the present study divided the entire NEI region into 12 zones. The same is illustrated in Figure 6.7, along with the direction of anisotropy. It should be mentioned here that the direction of anisotropy is not the $\emptyset$ value mentioned in Eq. 6.2. $\emptyset$ value is different and is calculated as mentioned in Section 6.2.1 . The direction of the anisotropy is given in Table 6.4 in terms of angle with respect to the north direction. The δ value is considered as 2 (highest anisotropy as per Ramanna and Dodagoudar, 2012) for the present work for all the zones except ZD10. For ZD10, no particular direction of anisotropy is considered, as no such dominant direction is in the zone (Figure 6.7). Thus, for ZD10 in Figure 6.7, the δ value is considered as 0 i.e. isotropic Kernel will be used for the zone.

Table 6.4 Direction of Anisotropy

Zone	Direction of anisotropy with respect to the North (in degree)
ZD1	112
ZD2	88
ZD3	53
ZD4	108
ZD5	113
ZD6	164
ZD7	53
ZD8	174
ZD9	39
ZD10	No particular direction is considered, as no such dominant direction is in the zone.
ZD11	68
ZD12	100

6.7 Incorporation of Uncertainties in the Magnitude and Location of EQs

The results of the Kernel density method depend entirely on the location and magnitude of the EQs in the catalogue. There is always uncertainty regarding the locations and magnitudes of the EQs (Kijko and Sellevoll 1992; Husen and Hardebeck 2010). For instrumental EQs, the uncertainty is low; however, for historical and paleoseismological EQs, the uncertainty is very high. If the magnitudes and locations are incorrect, the output results obtained from SHA or any other analysis are not reliable. Thus, the present work also considers the uncertainty related to the locations and magnitudes of EQs. For those EQs that occurred after 1964, uncertainty in magnitude is considered as ±0.3 (Mw) and the distance uncertainty is considered as ±10km. If the magnitude of an EQ available is M_{wa} according to available data, then it is considered that the magnitude of that particular EQ may vary between M_{wa}-0.3 and M_{wa}+0.3. Similarly, the location of the EQ may be anywhere within the 10km radius of the location available in the literature. For past EQs that occurred before 1964, the magnitude uncertainty is considered as ± 0.5Mw, and the location uncertainty is considered as ±30km. For paleo EQs, if information about the uncertainty is available in the literature, then the magnitude and location uncertainties are considered as per the literature (Table 6.1 and Figure 6.2). For those paleo EQs, where information regarding the uncertainty is not available in past literature, the magnitude uncertainty is considered as ± 0.5Mw, and the location uncertainty is considered as ±30km. As India's strong ground motion network was installed in 1963, the EQ information (size and location) of the EQs that occurred after 1964 is generally more accurate than the information available for EQs before 1964. Consequently, the uncertainty regarding the locations and magnitudes of the EQs that occurred after 1964 is lower than that of the EQs before 1964. Therefore, in this study, a higher level of uncertainty is assigned to EQs that occurred before 1964 than to those occurred afterwards. To incorporate the uncertainty related to EQ magnitudes and locations, the Latin hypercube sampling (LHS) technique with reduced correlation is employed. LHS is implemented using the 'lhsdesign' function with correlation control and the 'icdf' function in Matlab. Based on the uncertainty levels associated with the magnitude and location of each EQ, LHS is used to randomly generate the location coordinates within the required spatial boundaries and magnitudes within the specific range of magnitude (as mentioned earlier) for each EQ event. Subsequently, 100 synthetic EQ catalogues are generated separately from Catalogue 1 and Catalogue 2, resulting in a total of 200 synthetic EQ catalogues for this study. Considering all 200 synthetically generated EQ catalogues, spatial distributions of the seismicity rate density are determined using both isotropic and anisotropic Kernels by using Eqs. 6.1 and 6.2, respectively.

6.8 Results and Discussion

For the present work, the entire study area is divided into 0.1°×0.1° grid as mentioned in Section 6.2.1 . For each grid's centre, time-independent seismicity rate density, $\lambda(M, st)$ is estimated using Eqs. 6.1, 6.2, 6.3 based on the methodology presented in Section 6.2 . The $\lambda(M, st)$ values for each grid are determined considering both isotropic and anisotropic Kernel (Section 6.6) using 100 synthetic EQ catalogues obtained from Catalogue 1 and 100 synthetic EQ catalogues obtained from Catalogue 2 (Section 6.7). Subsequently, SHA is performed based on the λ(M,st) values estimated for the grids, utilising Eqs. 6.4 and 6.5. Further details are provided in the following subsections.

6.8.1 Distribution of seismicity rates in NE India

Based on the above discussion, a total of 400 sets (100 synthetic EQ catalogues × 2 types of EQ catalogues × 2 types of Kernel) of $\lambda(M, st)$ values are available for each of the grids. For each combination between the type of catalogue and the type of Kernel (Catalogue 1 + Isotropic Kernel, Catalogue 1 + Anisotropic Kernel, Catalogue 2 + Isotropic Kernel, and Catalogue 2 + Anisotropic Kernel), 100 sets of $\lambda(M, st)$ values are available by considering 100 synthetic EQ catalogues. The mean seismicity rate density for each grid is determined for each combination by taking the average of the 100 sets of $\lambda(M, st)$ values. Figure 6.8 and Figure 6.9 show the spatial distribution of the mean seismicity rate density in NE India based on synthetic EQ catalogues obtained for isotropic and anisotropy Kernel, respectively, by utilising EQ Catalogue 1. Similarly, Figure 6.10 and Figure 6.11 show the spatial distribution of the mean seismicity rate density in NE India based on synthetic EQ Catalogues obtained for isotropic and anisotropy Kernel, respectively, by utilising EQ Catalogue 2. It is observed from Figure 6.8 and Figure 6.9 that for M>4Mw EQs, the difference in the mean seismicity rate density is not that significant for isotropic and anisotropy Kernel in the case of Catalogue 1. However, for some locations near Faults, the seismic rate density is observed to be higher in the case of anisotropic Kernel than that of isotropic Kernel. Similar observations can be made for the spatial distribution of the mean seismicity rate density in NE India based on Synthetic EQ Catalogues obtained by utilising EQs in Catalogue 2 (Figure 6.10 and Figure 6.11). Figure 6.8, Figure 6.9, Figure 6.10, and Figure 6.11 manifest that the seismicity is high in the locations of the Indo-Burma ranges, the west side of Arunachal Pradesh, the Middle region of Assam, the Shillong plateau and the

Southern part of Sikkim. In Figure 6.8, Figure 6.9, Figure 6.10, and Figure 6.11, the EQs that occurred between 2021-2022 are also superimposed. Most of the EQs that occurred during this period fell within the regions of higher seismicity rate densities.

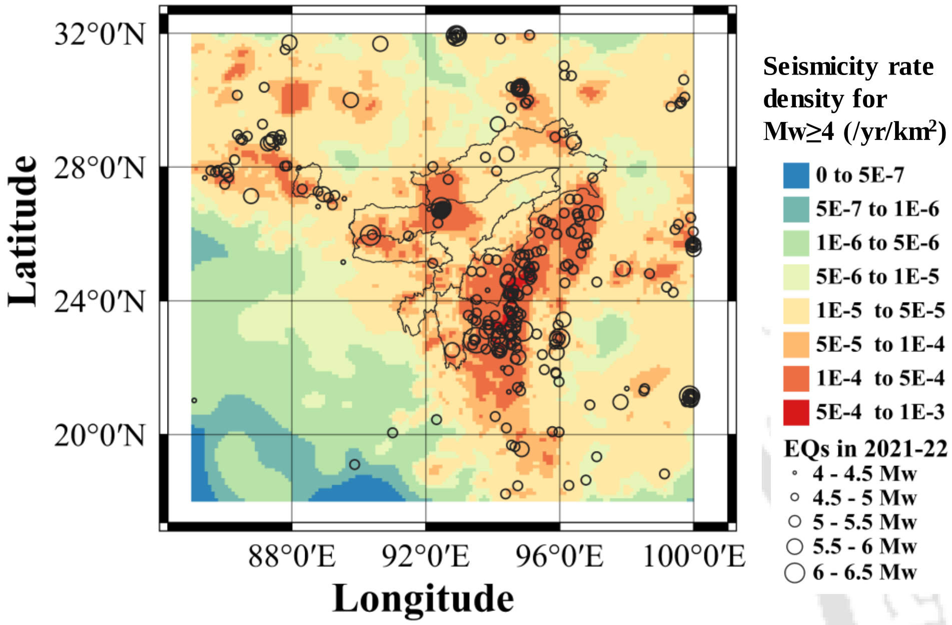

Figure 6.8 Mean seismicity rate density for EQs having magnitude equal and more than 4Mw, estimated by using synthetic EQs catalogues obtained from Catalogue 1 in the case of isotropic Kernel

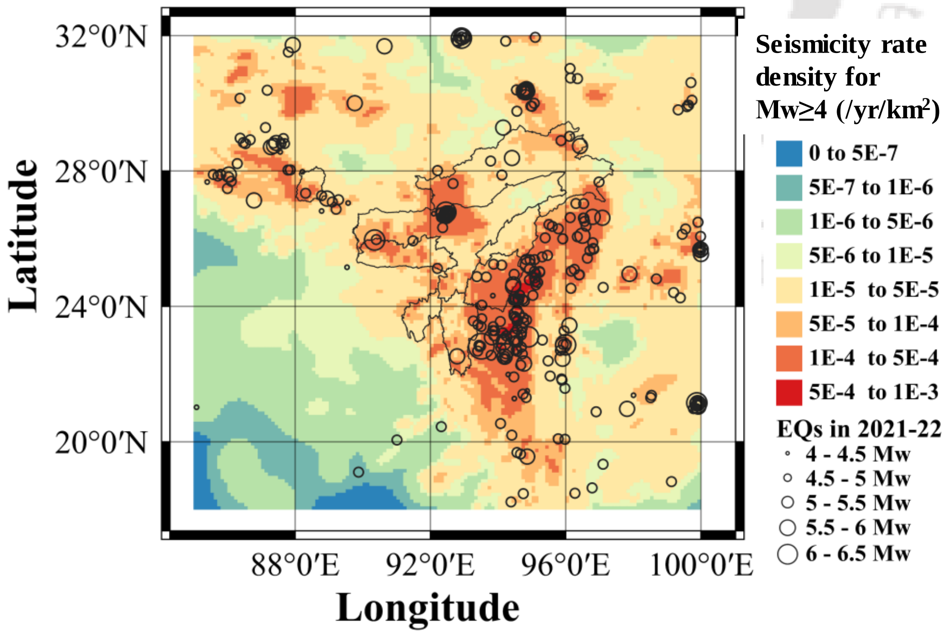

Figure 6.9 Mean seismicity rate density for EQs having magnitude equal and more than 4Mw, estimated by using synthetic EQs catalogues obtained from Catalogue 1 in the case anisotropic Kernel

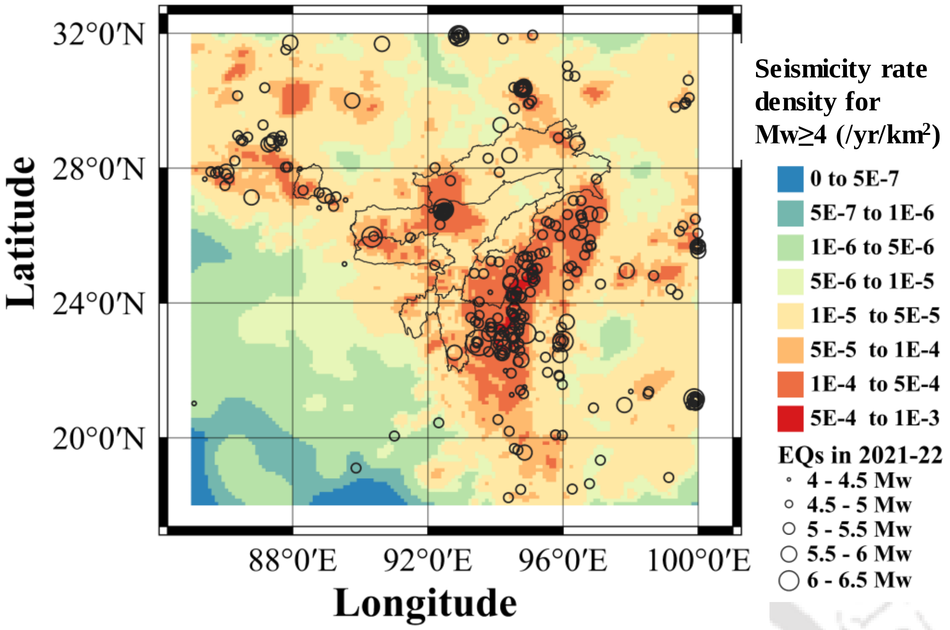

Figure 6.10 Mean seismicity rate density for EQs having magnitude equal and more than 4Mw, estimated by using synthetic EQ catalogues obtained from Catalogue 2 in the case of isotropic Kernel

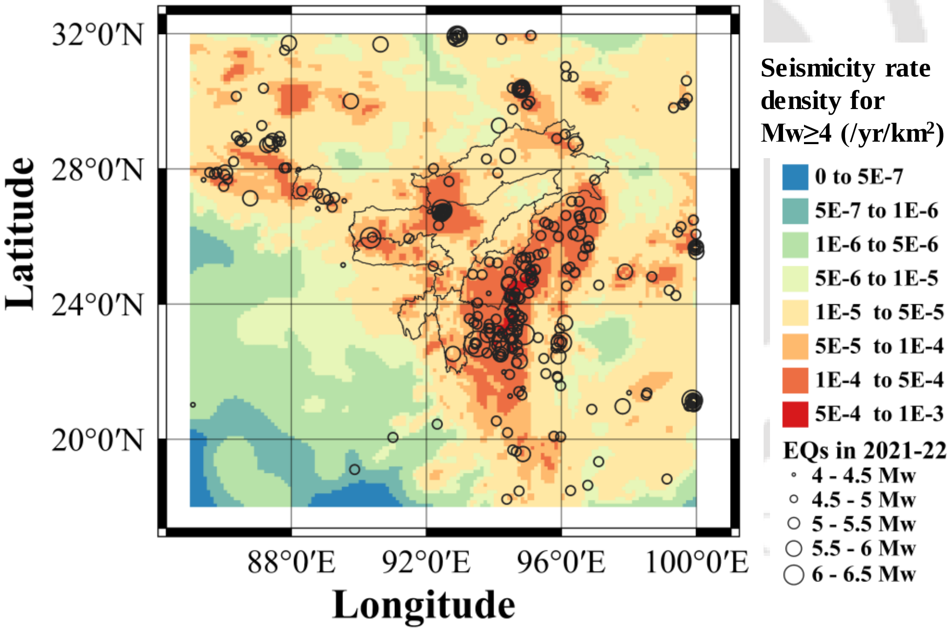

Figure 6.11 Mean seismicity rate density for EQs having magnitude equal and more than 4Mw, estimated by using synthetic EQ catalogues obtained from Catalogue 2 in the case of anisotropic Kernel

Similarly, the seismicity rate densities for EQ magnitude bins $4 \leq Mw < 4.5$, $4.5 \leq Mw < 5$, $5 \leq Mw < 5.5$, $5.5 \leq Mw < 6$, $6 \leq Mw < 6.5$, $6.5 \leq Mw < 7$, $7 \leq Mw < 7.5$, $7.5 \leq Mw < 8$, $8 \leq Mw < 8.5$, and $8.5 \leq Mw$ are determined for each grid. Figure 6.12, Figure 6.13, Figure 6.14, and Figure 6.15 show the seismicity rate density for higher magnitude bins such as $7 \leq Mw < 7.5$, $7.5 \leq Mw < 8$, $8 \leq Mw < 8.5$, and $8.5 \leq Mw$ respectively, in the case of Catalogue 1. Figure 6.12, Figure 6.13,

Figure 6.14, and Figure 6.15 clearly show the difference in seismicity rate densities obtained by using isotropic and anisotropic Kernel. For the case of anisotropic Kernel, the seismicity rate density is more near the faults than that for the case of isotropic Kernel.

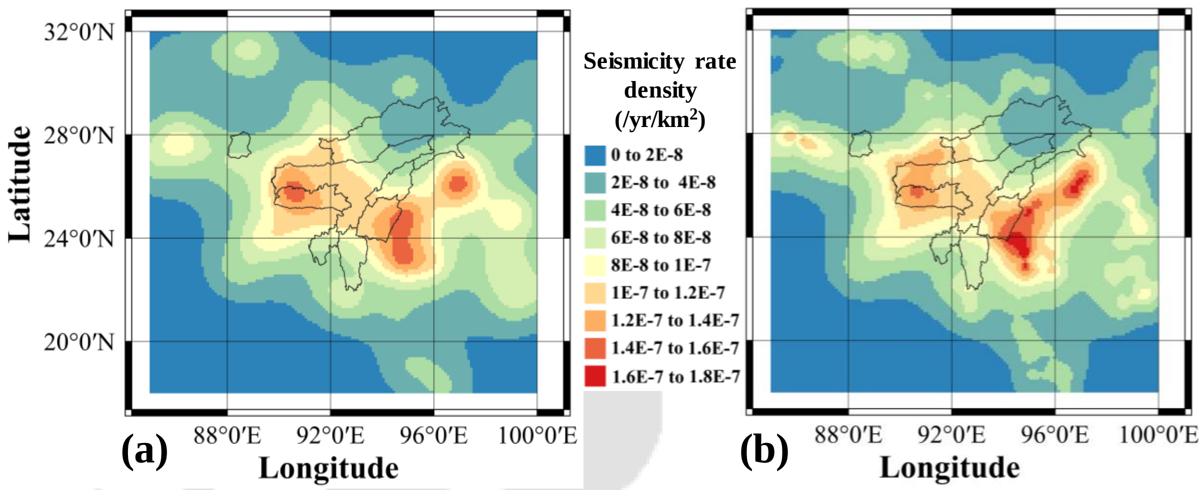

Figure 6.12 Spatial distribution of seismicity rated for the magnitude bin of 7≤Mw<7.5 obtain a. by using Isotropic Kernel in Catalogue 1, b. by using Anisotropic Kernel in Catalogue 1

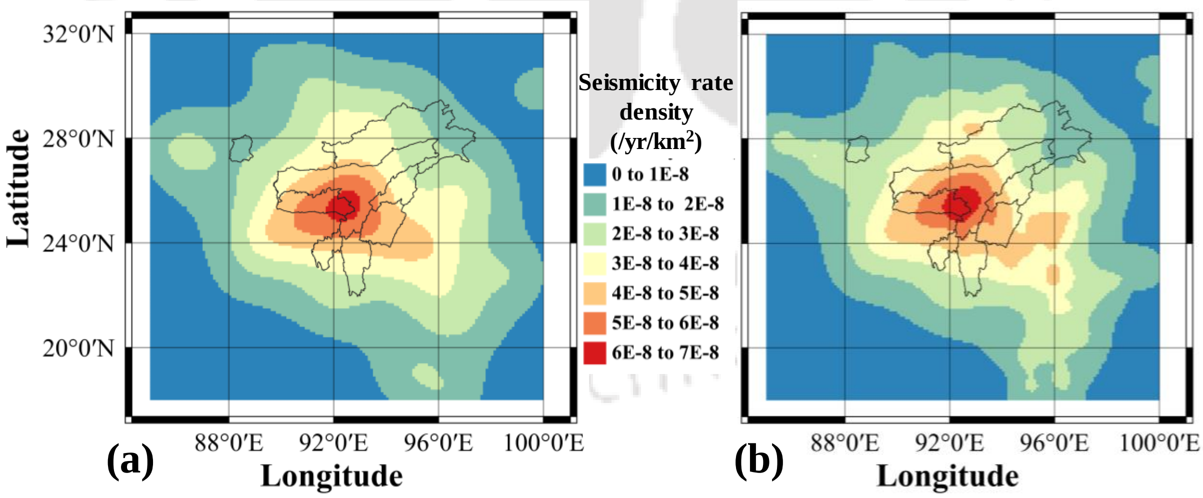

Figure 6.13 Spatial distribution of seismicity rated for the magnitude bin of 7.5≤Mw<8 obtain a. by using Isotropic Kernel in Catalogue 1, b. by using Anisotropic Kernel in Catalogue 1

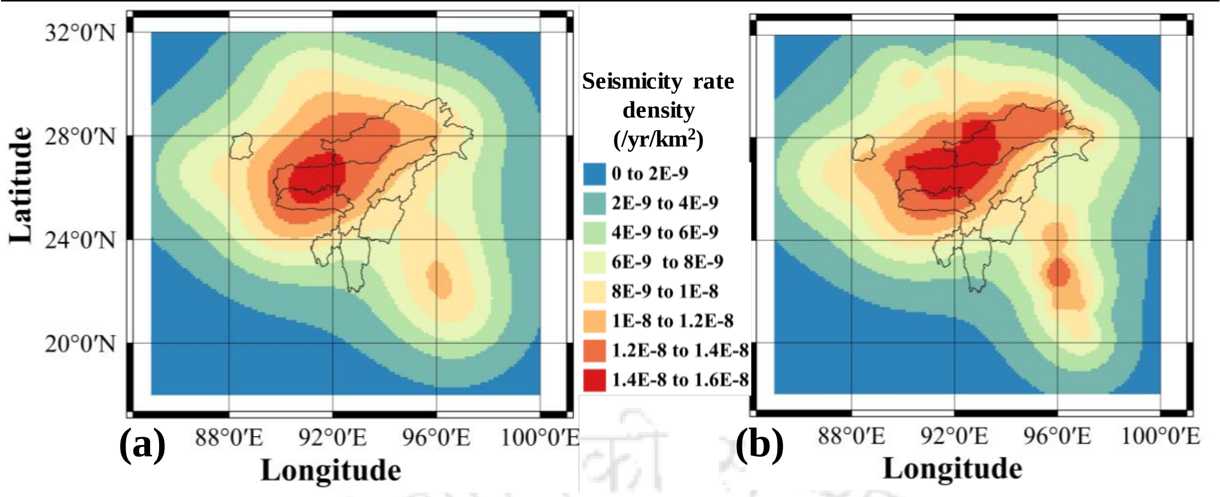

Figure 6.14 Spatial distribution of seismicity rated for the magnitude bin of 8≤Mw<8.5 obtain, a. by using Isotropic Kernel in Catalogue 1, b. by using Anisotropic Kernel in Catalogue 1

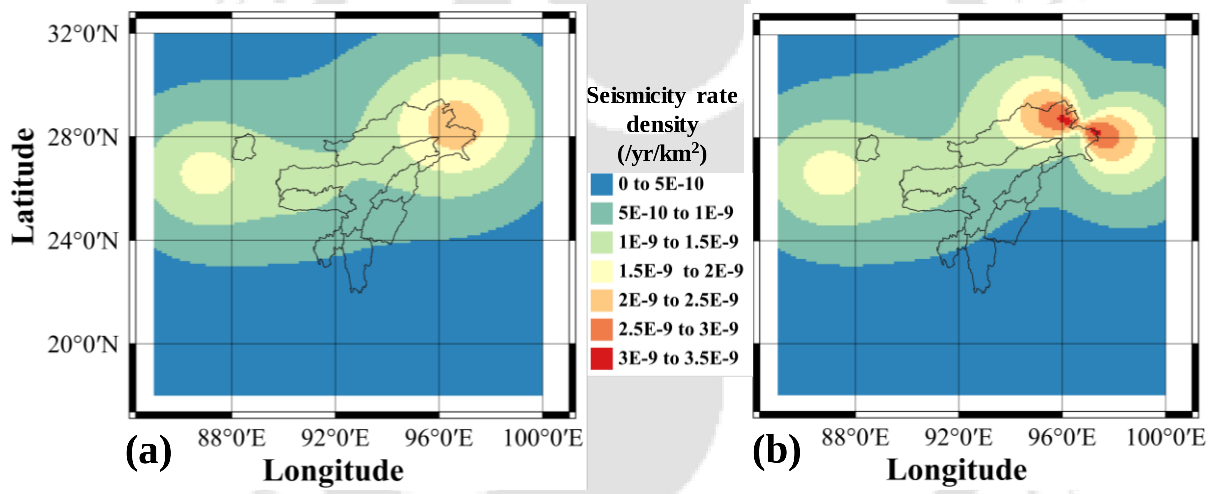

Figure 6.15 Spatial distribution of mean seismicity rated for the magnitude bin of 8.5≤Mw obtain a. by using Isotropic Kernel in Catalogue 1, b. by using Anisotropic Kernel in Catalogue 1

6.8.2 Comparison with the actual seismicity of the region

In the present work, the mean seismicity rate density obtained above is compared with the distribution of EQs that happened after 2021 A.D. by utilising a Molchan diagram (Molchan 1990; Zechar and Jordan 2008), as depicted in Figure 6.16. The Molchan diagram illustrates the “fraction alarm occupied space” vs. the “fraction of failure in forecast”. The term “fraction of alarm-occupied space” represents the proportion of events inside the specified region of interest having a seismicity rate greater than or equal to the “alarm” seismicity rate. The “fraction of failure in forecasting” is the proportion of subsequent EQs having a lesser seismicity rate than the alarm seismicity rate (Zechar and Jordan, 2008). To obtain the Molchan diagram, the seismicity rate of the location of each event is determined. Then, for a particular event, the total area with seismicity rates smaller than or equal to the location's seismicity rate is determined and expressed as a percentage of the whole area. All the events are subsequently sorted based on their respective percentages and plotted with respect to the event count, which is expressed as a percentage of the overall number of predicted events. The presence of a diagonal line in the Molchan diagram indicates the absence of any correlation between the observed and the forecast seismicity. Conversely, an upward arc signifies a negative correlation between the observed and the forecast seismicity, whereas a downward arc represents a positive correlation between the observed and forecast seismicity. A favourable outcome would be characterised by a minimal proportion of alarms and a low rate of failure in predicting seismicity.

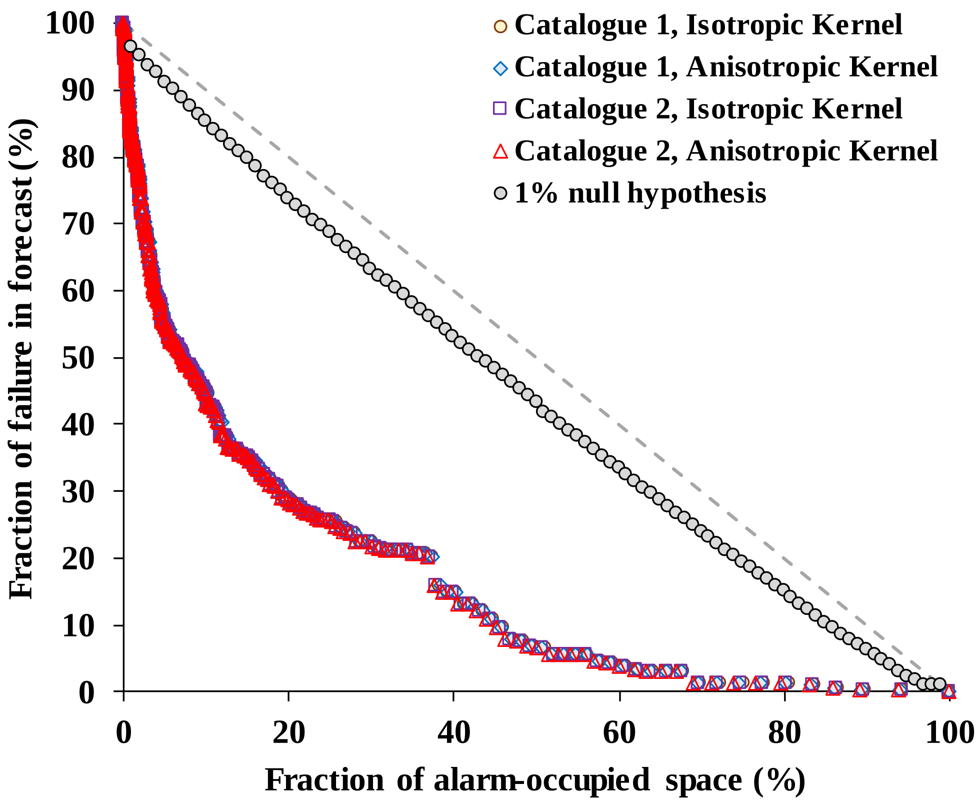

Figure 6.16 Molchan diagram to investigate the correlation between the forecasted seismicity rates obtained from the four cases and the forecast EQs (The grey circular dots represents the 99% significance level determined by 303 forecasting events that occurred in 2021-22)

In the present work, downward arc is observed for all the cases. In Figure 6.16, the 99% significant level (1% null hypothesis in the figure/ grey dots) is plotted considering the total number of forecasted EQs (303). The null hypothesis line is obtained by solving the following equation given by Zechar and Jordan (2008) for α=1%:

$$\sum_{n=h}^{N} B\left(h \middle| N, \widetilde{P_A}\right) = \alpha \qquad 6.8$$

Where,

$B\left(h\backslash N,\ \widetilde{P_A}\right) = \binom{N}{h}\widetilde{P_A}^{h}(1-\widetilde{P_A})^{N-h}$, $\widetilde{P_A}$= Probability weighted of the alarm region, N= Number of targeted EQs, here 303, h=Number of EQs located in the alarm region.

As in the Molchan diagram, the plots of the four cases are below the 1% null hypothesis plots; the results show robustness in the forecasting ability of the present work.

6.8.3 Spatial distribution of the seismic hazard level

Based on the seismicity rate distribution, the annual rate of occurrence $\lambda_o(m_i, st)$ of a particular magnitude m_i is determined (Eqs. 6.5). For the PSHA, in the present work, the mean magnitudes of the bins are considered for finding the seismic hazard values. These values are 4.2, 4.7, 5.2, 5.7, 6.2, 6.7, 7.2, 7.7, 8.2 and 8.7Mw. For these values, using Eqs. 6.5 and 6.6, PSHA is carried out by using the three GMPEs as mentioned in Section 6.2.1 to obtain the PGA for the entire NE India.

Figure 6.19 presents contours of mean PGA corresponding to DBE (10% probability of exceedance in 50 years) for NE India. Figure 6.19 illustrates that the variation in PGA for NEI is from 0.24g to 0.57g. The southern part of the state of Nagaland and the eastern part of the state of Manipur show the highest seismic hazard level within NE India (Figure 6.19). The most variations are seen in the region of Sikkim, parts of Arunachal and Indo-Burma ranges. For Manipur, PGA varies from 0.39g to 0.57g. Similarly, for Nagaland, PGA varies from 0.27g to 0.48g. These results can be attributed to past EQs' high density and the occurrence of higher magnitude events near Nagaland and Manipur, which are close to the Indo–Burma ranges. The present study observes the lowest seismic hazard value in northern parts of Sikkim, central parts of Arunachal Pradesh and eastern parts of Assam. The lowest seismic hazard value is observed in Sikkim and parts of Arunachal Pradesh and Assam. For Sikkim, the PGA varies from 0.26g to 0.38g. The south-central part of the state of Arunachal Pradesh and a minor section of the state of Assam (in the eastern part) show PGA ranging from 0.24g to 0.27g. These locations experience a notably lower occurrence of EQs in the past. However, the centre regions of Assam, which are in proximity to the Bhutan Himalaya and the Shillong Plateau, have a significantly high PGA value ranging from 0.39g to 0.42g. This region is also near to the location where the 1897 Assam EQ took place. Comparable findings can also be derived from the western and eastern areas of Arunachal Pradesh, where the PGA ranges from 0.42g to 0.48g. These regions saw a high concentration of previous EQs with greater magnitudes. Both Assam and Arunachal Pradesh encompass vast territories that include areas that have encountered rare to very frequent EQs. Consequently, there is a significant variation in the PGA values observed in the two states mentioned above, ranging from 0.24g to 0.48g. The PGA values in other states of NE India do not exhibit significantly wide variation.

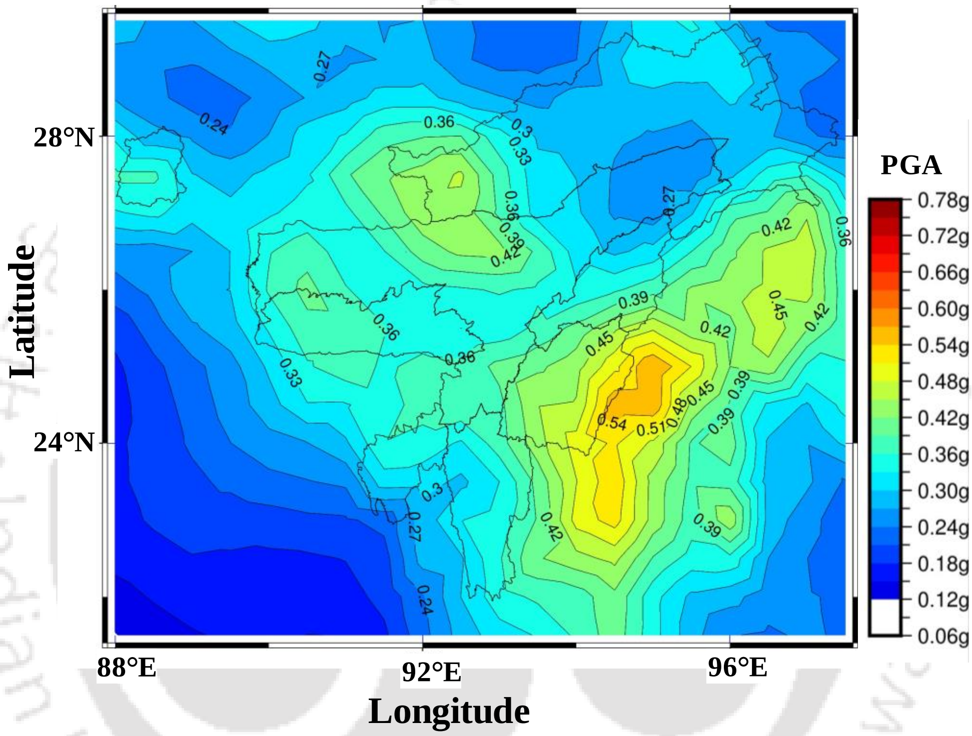

Figure 6.17 Spatial distribution of mean PGA for 10% Probability in 50 years

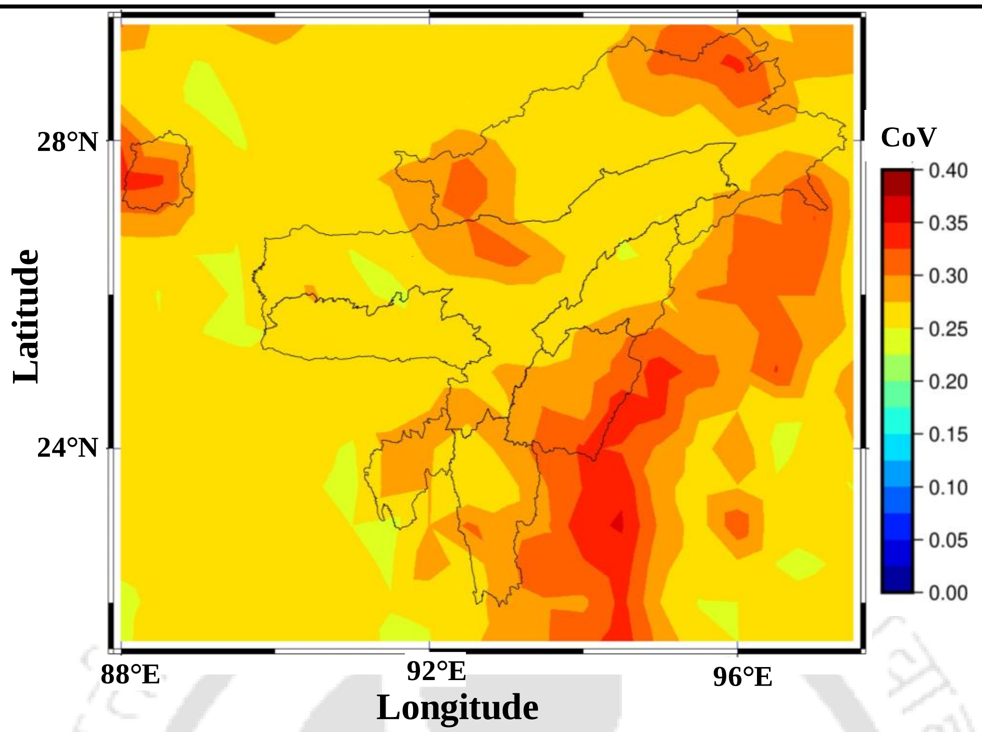

Figure 6.18 Spatial distribution of CoV of PGA for 10% Probability in 50 years

.

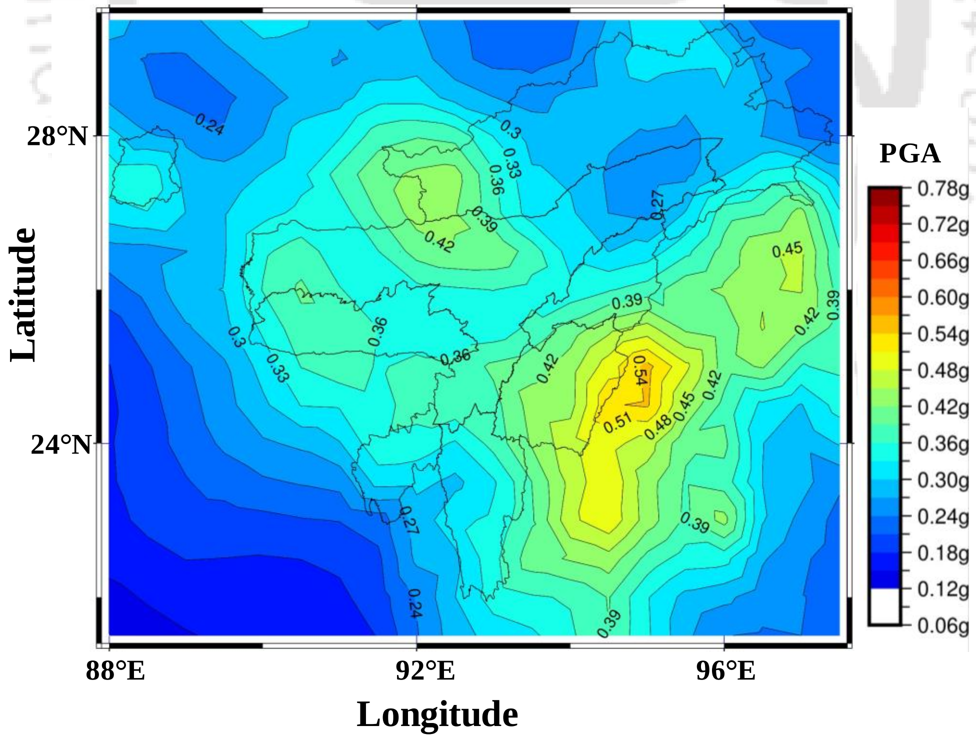

Figure 6.19 Spatial distribution of mean PGA for 10% Probability in 50 years considering Isotropic Kernel

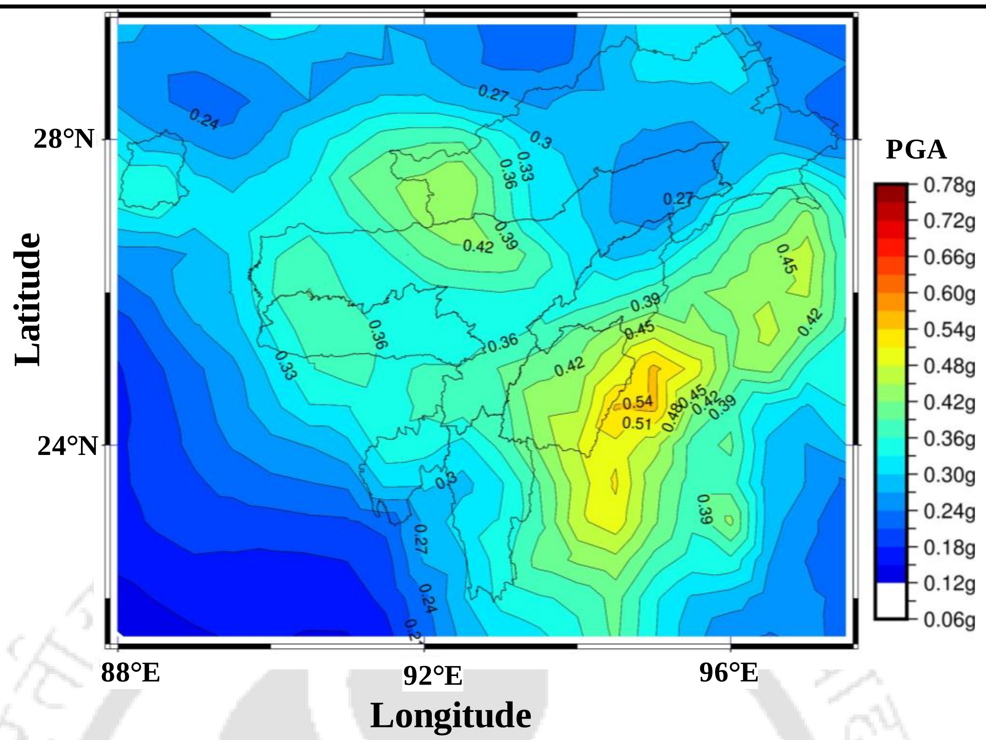

Figure 6.20 Spatial distribution of mean PGA for 10% Probability in 50 years considering Anisotropic Kernel

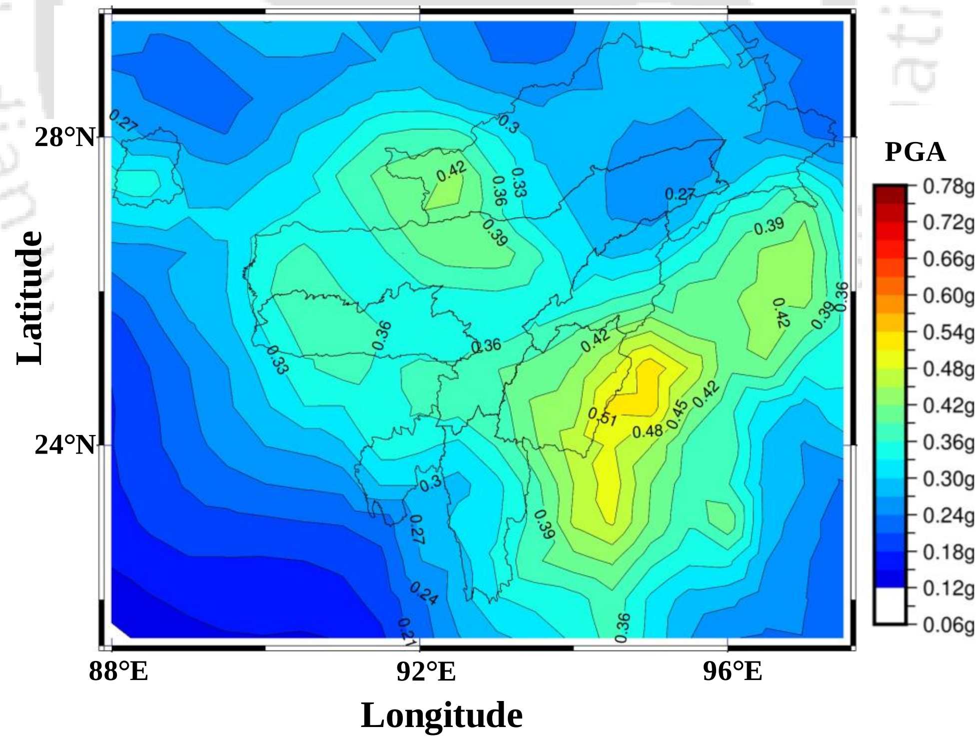

Figure 6.21 Spatial distribution of mean PGA for 10% Probability in 50 years considering Catalogue 1

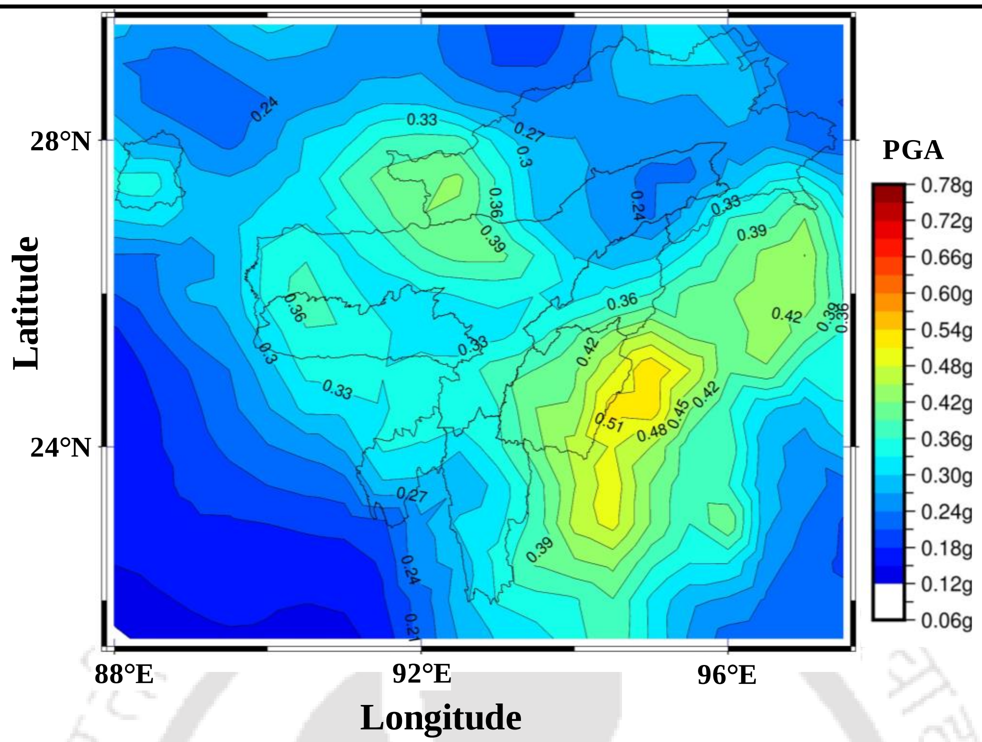

Figure 6.22 Spatial distribution of mean PGA for 10% Probability in 50 years considering Catalogue 2

The PGA values in the state of Meghalaya vary from 0.27g to 0.42g; PGA values in the state of Tripura vary from 0.24g to 0.39g, and for the state of Mizoram, the PGA vary from 0.3g to 0.42g (see Figure 6.19). Based on the distribution of PGA for NE, it can be stated that the highest PGA values are observed near the Indo-Burma ranges, followed by the Bhutan-Himalaya and then the Shillong Plateau. To evaluate the variability/uncertainty of the estimated seismic hazard levels, CoVs are calculated. The CoV map of hazard level for 10% probability in 50 years is depicted in Figure 6.18 as contours corresponding to PGA values. The CoV of PGA within NE India is within 0.4, which indicates that the variability in the estimated PGA value is low. The variability is more for Sikkim, west and east part of Arunachal Pradesh, Manipur and some parts of Nagaland and Mizoram region.

Figure 6.19 and Figure 6.20 show the spatial distribution of mean PGA values for isotropic Kernel and anisotropic Kernel, respectively. The difference in location-wise PGA values in both cases is not significant. In both cases, the range PGA values in NE India is from 0.24g to 0.54g. Similarly, Figure 6.21 and Figure 6.22 show the spatial difference of mean PGA values obtained from EQ catalogues generated from Catalogue 1 and Catalogue 2, respectively. Though most of the places, the PGA values in both cases have no difference, many areas have significantly higher PGA values when Catalogue 1 is considered than that when Catalogue 2 is considered. The reason for this is that the EQs which are complete with respect to time, are used in the case of Catalogue 1. Thus, the number of EQs per year is more for higher magnitude EQs (>7Mw) than that of Catalogue 2, where the entire catalogue time is considered.

Similarly, Figure 6.23 presents PSHA outcomes for 2% probability of exceedance in 50 years (MCE). It is observed that for MCE, PGA values in the NE region vary from 0.45g to 0.85g (Figure 6.23). Similar to the DBE, the highest hazard level is seen in the eastern region of Manipur, southern region of Nagaland. After this, some parts of Arunachal Pradesh and Assam and Shillong plateau have higher values of the hazard level. In this case, also CoV values are within 0.4 Figure 6.24

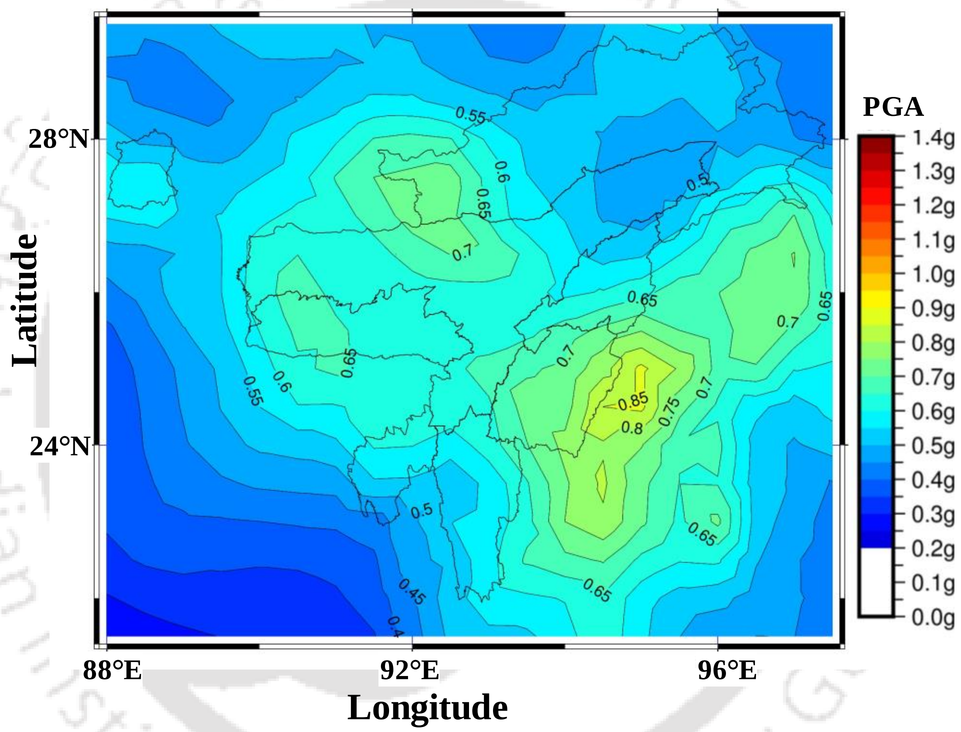

Figure 6.23 Spatial distribution of mean PGA for 2% Probability in 50 years

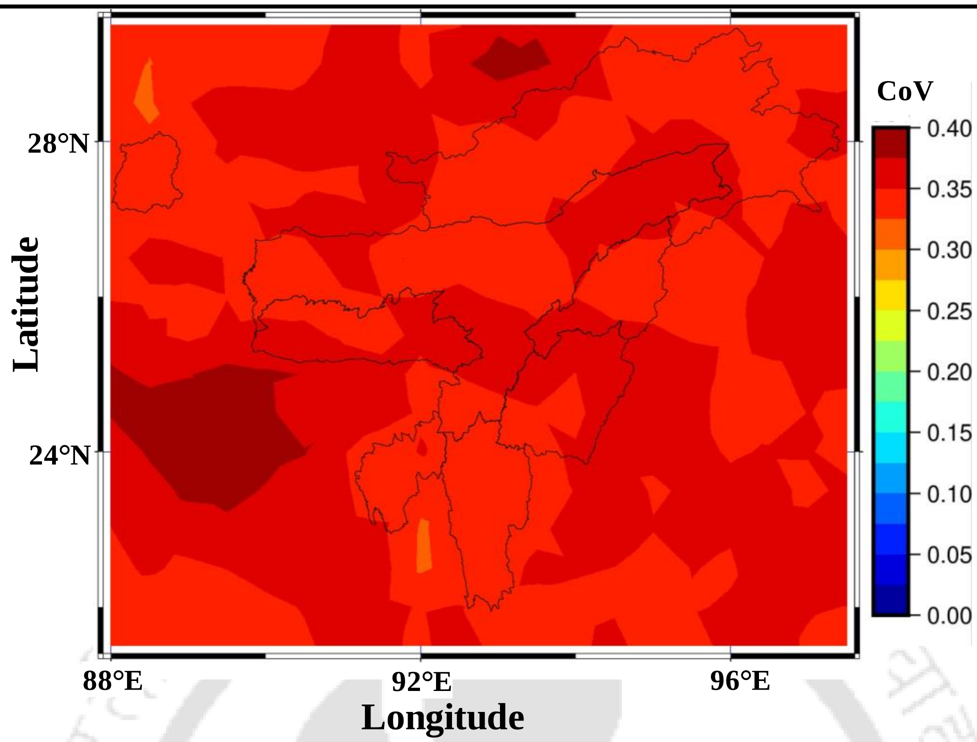

Figure 6.24 Spatial distribution of CoV of PGA for 2% Probability in 50 years

Figure 6.25 and Figure 6.26 show the spatial distribution of mean PGA values obtained considering isotropic and anisotropic Kernels. In both cases, the PGA values are in the range of 0.45 to 0.9g. However, location wise, a significant difference is observed in the PGA values for both cases. In the case of anisotropic Kernel, PGA values are observed to be more near the Fault region (Figure 6.26) than that of isotropic Kernel. Higher PGA values are seen in the case of anisotropic Kernel in locations south of Sikkim, the west side of Arunachal Pradesh, and a small part of Manipur. Similarly, for some locations, such as eastern parts of Assam and in Shillong, lower PGA values are observed in the case of anisotropic Kernel than those in the case of isotropic Kernel. Figure 6.27 shows the spatial distribution of mean PGA values obtained by using EQ catalogues generated from Catalogue 1 and Figure 6.28 shows the spatial distribution of mean PGA values obtained by using EQ catalogues generated from Catalogue 2. Similar to the DBE, in both cases, the PGA values are higher in the case of Catalogue 1, where a complete EQ catalogue is considered. The difference in PGA values in both cases is much more significant in case of MCE than DBE.

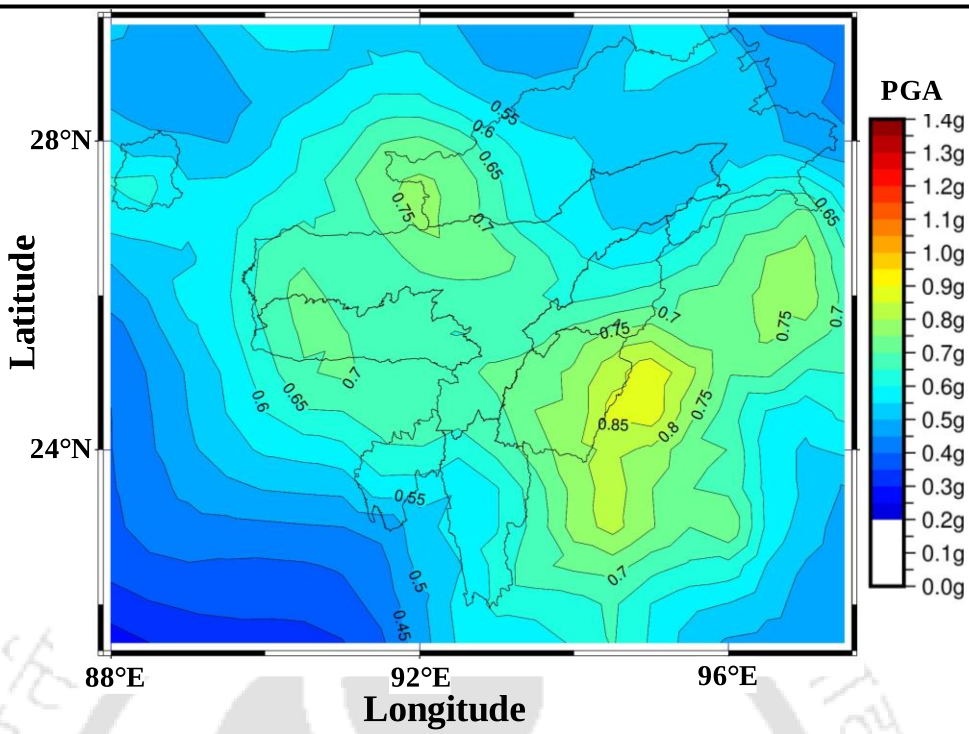

Figure 6.25 Spatial distribution of mean PGA for 2% probability in 50 years considering isotropic Kernel

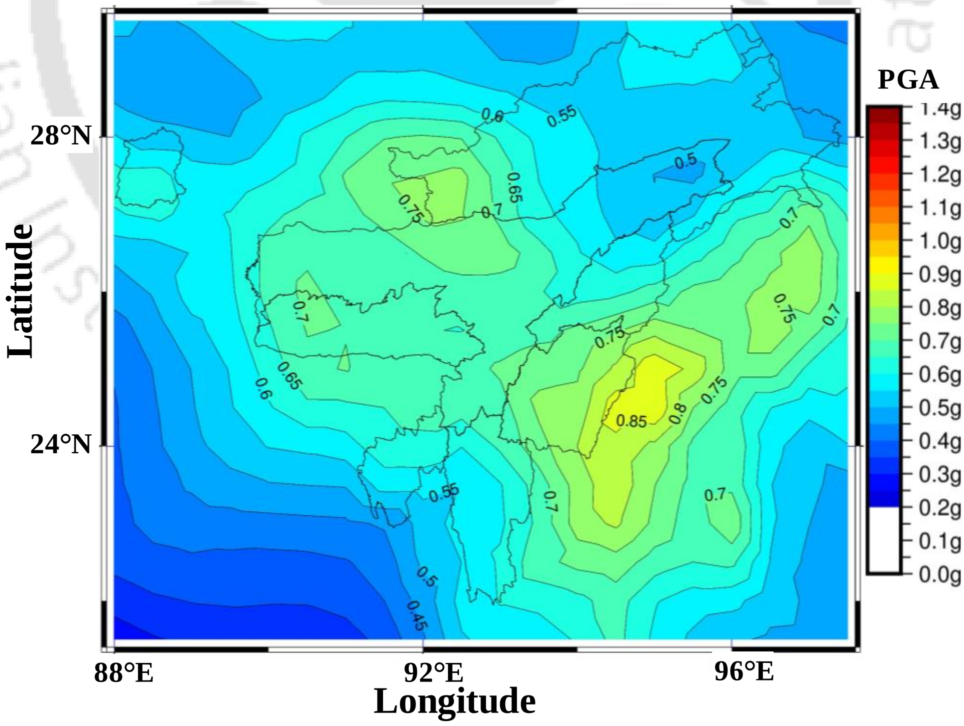

Figure 6.26 Spatial distribution of mean PGA for 2% probability in 50 years considering anisotropic Kernel

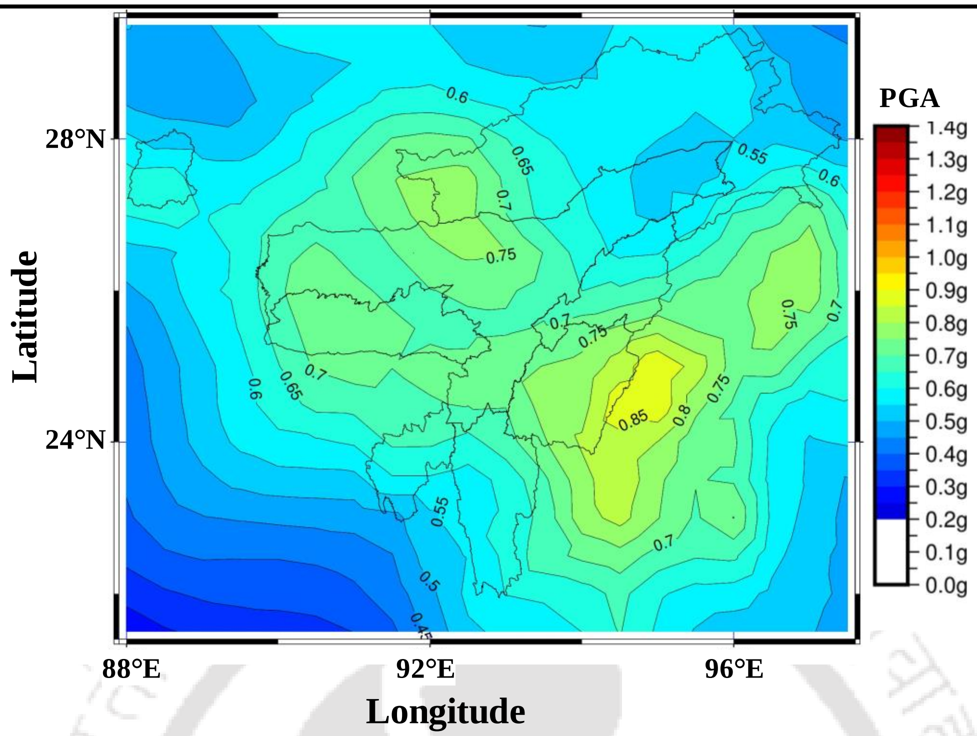

Figure 6.27 Spatial distribution of mean PGA for 2% probability in 50 years considering Catalogue 1

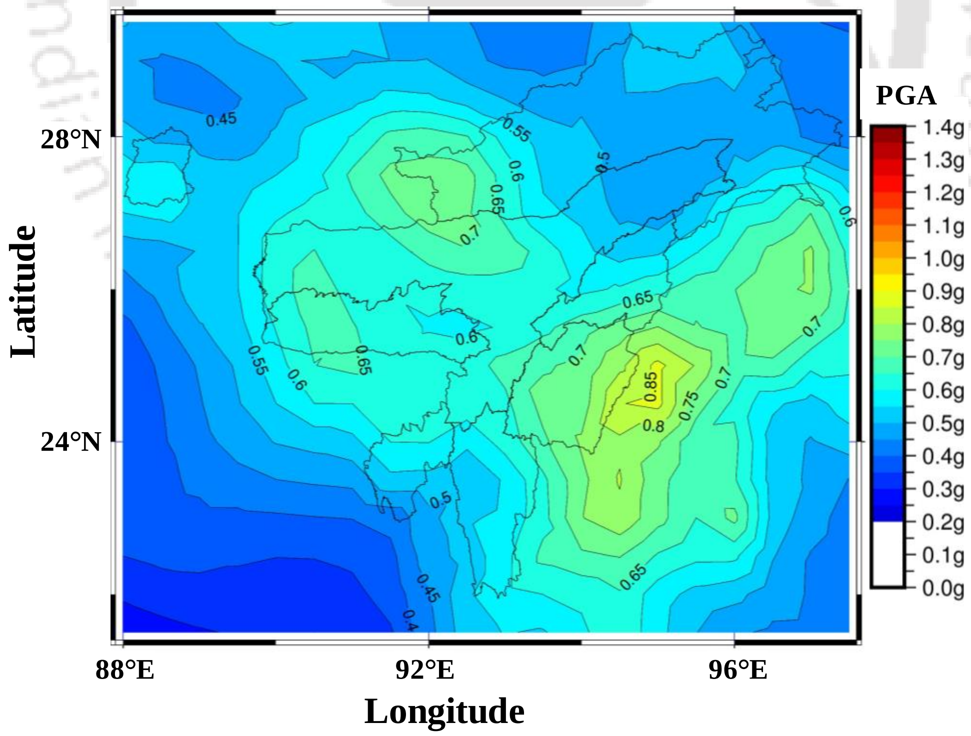

Figure 6.28 Spatial distribution of mean PGA for 2% probability in 50 years considering Catalogue 2

6.9 UHS for different locations

For the 8 important Cities (see Section 5.10, Chapter 5), Guwahati, Agartala, Imphal, Shillong, Aizawl, Kohima, Gangtok, and Itanagar, the hazard level in terms of Sa is determined for different time periods for DBE and MCE. Based on Sa values obtained by utilising the 200 catalogues, mean, median, 5th, 16th, 84th, and 95th percentile UHS are developed for all the Cities. Figure 6.29 a, b, c, and d show the DBE based UHS obtained for Guwahati, Agartala, Imphal, and Shillong, respectively. Figure 6.30 a, b, c, and d show the DBE based UHS obtained for Aizawl, Kohima, Gangtok, and Itanagar city centres, respectively It is observed from the results that the mean hazard level in terms of PGA is highest (PGA=0.43g) for Imphal (Figure 6.29 c) followed by Kohima (PGA=0.38g), Guwahati (PGA=0.36g), Gangtok (PGA=0.36g), Shillong (PGA=0.34g), Agartala (PGA=0.32g), Aizwal (PGA=0.31g), and Itanagar (PGA=0.30g). Similarly, Figure 6.31 a, b, c, and d, show the MCE based UHS obtained for Guwahati, Agartala, Imphal, and Shillong, respectively. Figure 6.32 a, b, c, and d show the MCE based UHS obtained for Aizawl, Kohima, Gangtok, and Itanagar city centres, respectively. . For MCE , the mean hazard level in terms of PGA is highest (PGA=0.75g) for Imphal (Figure 6.31 c) which is followed by Kohima (PGA=0.69g), Guwahati (PGA=0.67g), Shillong (PGA=0.65g), Gangtok (PGA=0.63g), Agartala (PGA=0.60g), Aizwal (PGA=0.59g), and Itanagar (PGA=0.59g). Table 6.5 shows the variability in estimation Sa values in terms of CoV for all the cities for different time periods. It is observed that for all the cases, the variability is within 0.4. The highest CoV in determining PGA is observed for Gangtok (0.35). For all the other cities, the CoV values for MCE are higher than those of DBE.

Table 6.5 CoV in Sa estimation for different cities

Time periods (s)	Scenerio	Guwahati	Agartala	Imphal	Shillong	Aizawl	Kohima	Gangtok	Itanagar
0	DBE	0.28	0.29	0.30	0.27	0.28	0.29	0.35	0.27
	MCE	0.30	0.30	0.33	0.30	0.30	0.32	0.34	0.30
0.1	DBE	0.24	0.25	0.27	0.24	0.23	0.24	0.32	0.22
	MCE	0.30	0.29	0.32	0.30	0.28	0.30	0.32	0.29
0.2	DBE	0.25	0.26	0.28	0.25	0.24	0.26	0.32	0.23
	MCE	0.32	0.31	0.34	0.32	0.31	0.33	0.34	0.31
0.4	DBE	0.20	0.20	0.22	0.20	0.20	0.21	0.25	0.20

	MCE	0.26	0.24	0.27	0.26	0.24	0.26	0.27	0.26
1	DBE	0.25	0.25	0.28	0.25	0.27	0.27	0.27	0.27
	MCE	0.29	0.28	0.30	0.28	0.28	0.30	0.29	0.30
2	DBE	0.31	0.31	0.32	0.31	0.32	0.32	0.31	0.33
	MCE	0.34	0.33	0.35	0.34	0.33	0.35	0.35	0.36

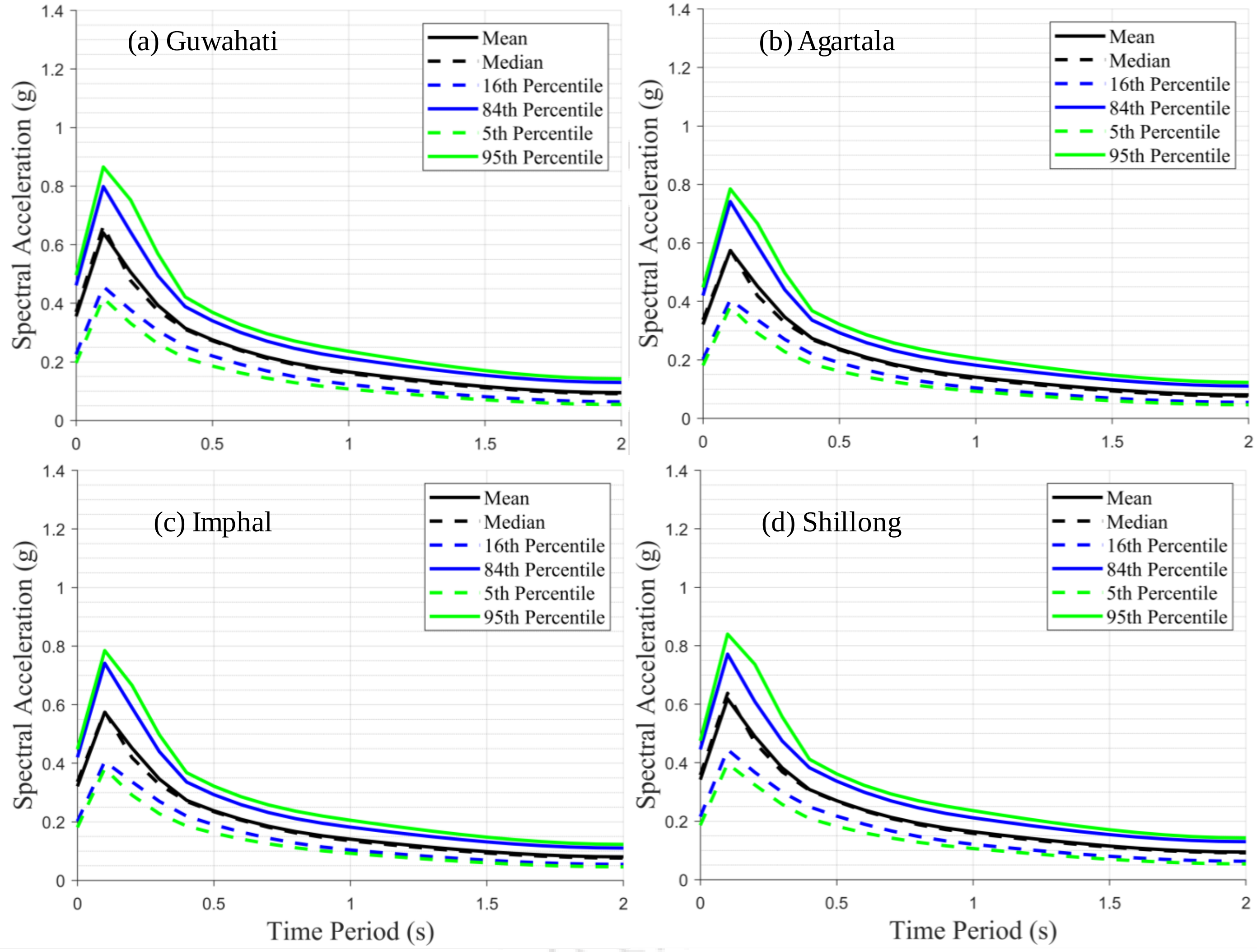

Figure 6.29 UHS for DBE for a. Guwahati, b. Agartala, c. Imphal and d. Shillong

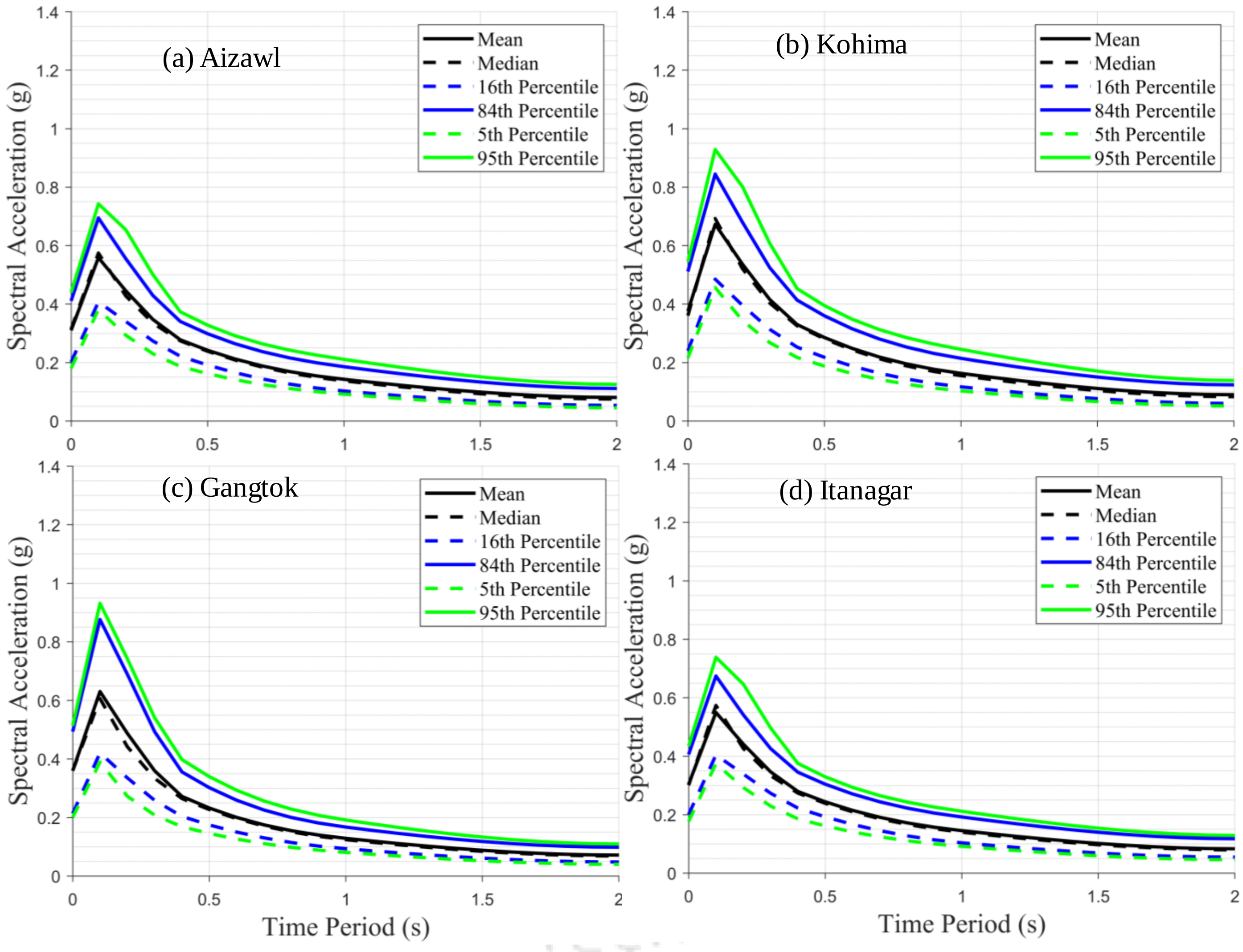

Figure 6.30 UHS for DBE for a. Aizawl, b. Kohima, c. Gangtok and d. Itanagar

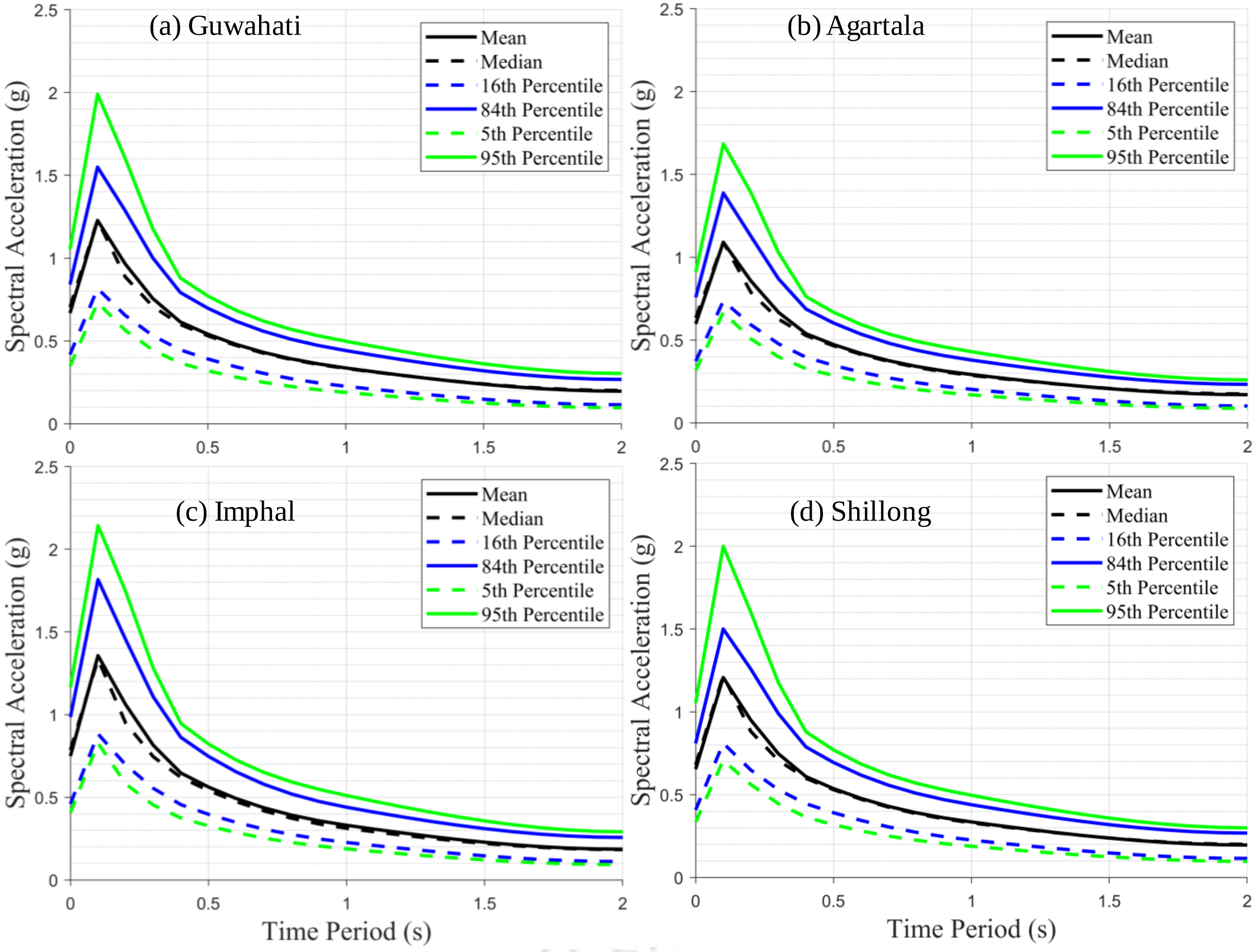

Figure 6.31 UHS for MCE for a. Guwahati, b. Agartala, c. Imphal and d. Shillong

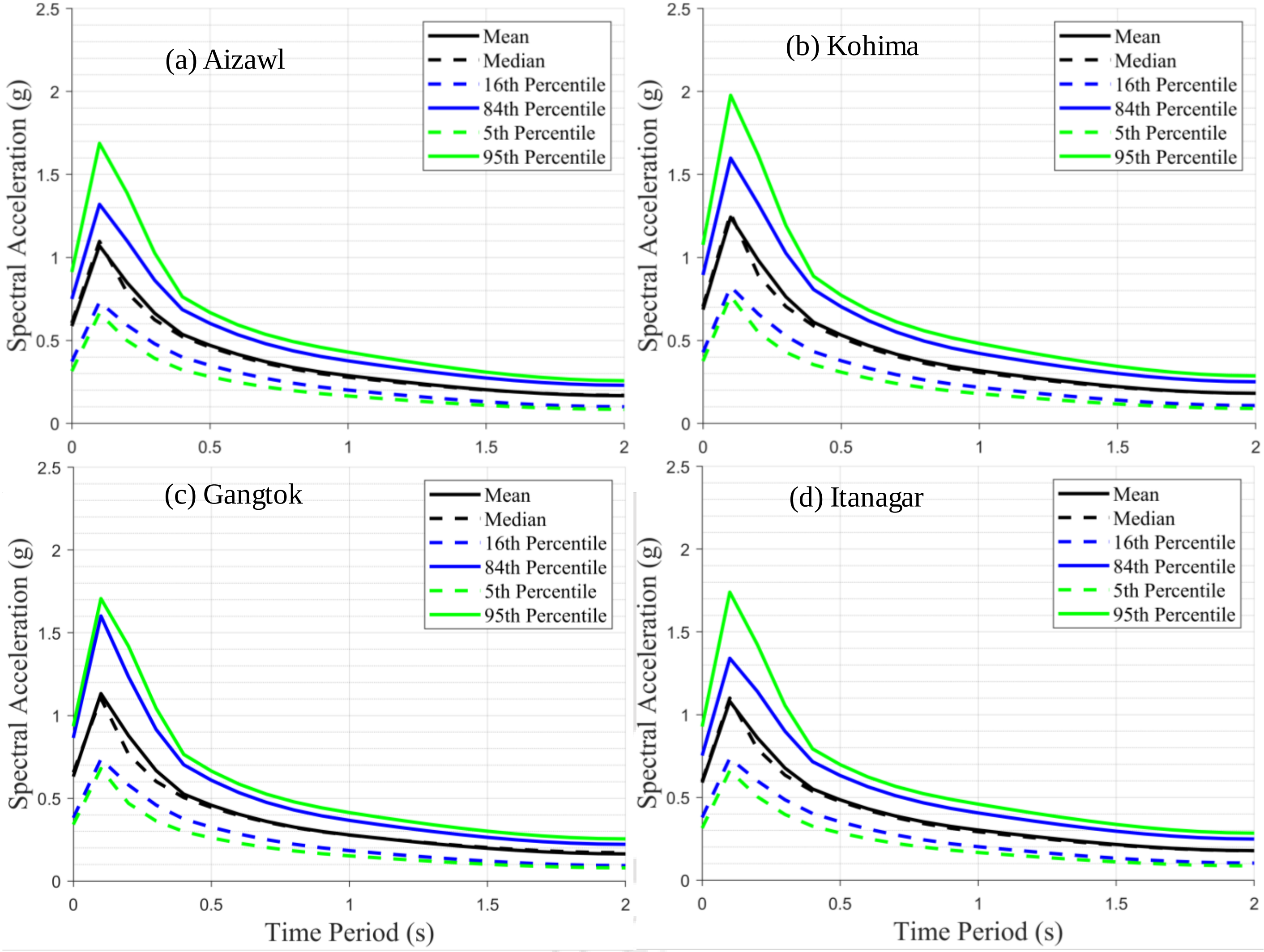

Figure 6.32 UHS for MCE for a. Aizawl, b. Kohima, c. Gangtok and d. Itanagar

6.10 Summary

This chapter utilises Kernel density method for PSHA. The uncertainties with respect to magnitude and locations of EQs are utilised for the determination of seismic hazard levels for NE India. The spatial distribution of the seismic activity rate density suggests that the overall seismic activity is more for the Indo-Burmese Ranges, the Sikkim Himalaya, the Arunachal Himalaya, the Shillong Plateau and the Assam Valley Region. For Great EQs (M>8Mw), seismic activity is more in the Bhutan-Arunachal Himalaya and the Shillong Plateau. The mean seismic activity rate density obtained from the synthetic EQ catalogues is then compared with the location and magnitude of the recently occurred EQs (between 2021-2022 A.D.). Most of the EQs that occurred during this period fell within the region of higher seismicity rate. The seismic activity rate density is further used for PSHA. It is observed from the PSHA result that the mean PGA for the MCE scenario of NE India varies from 0.45 to 0.85g, and CoV is within 0.4. The mean PGA for the DBE varies from 0.24 to 0.57g. Though the range in the isotropic and anisotropic Kernel case is the same, location-wise difference in PGA value is observed. In the case of the anisotropic Kernel, as fault direction along with EQ data is utilised, for some locations near faults, PGA values are more than that obtained in the case of the isotropic Kernel. Similarly, it is also found that in the case of Catalogue 1, where EQs complete with respect to magnitude and time are present, the PGAs obtained are more than those obtained by using Catalogue 2 for all the locations. This suggests that in consideration of the catalogue where EQ are complete with respect to time and magnitude, hazard values are not underestimated with respect to the location. Further, UHS are developed for the 8 important cities in NE India. These UHS can be used for designing and evaluating the seismic risk of structures in the region.

Chapter 7. Development of GMPE for NE India, Applicable to Different Site Classes

7.1 General

GMPE or attenuation relation is an empirical relation where the ground motion characteristics at the site of interest are correlated with EQ size and distance of the site from EQ source. This relation is crucial, as seen in Chapter 5, where different GMPEs are used for seismic hazard estimation. Referring to the discussion in Chapter 2, it should be mentioned here again that though regional GMPEs for NE India are available, these are limited for specific magnitude and distance range. In addition, most of the existing GMPEs are only applicable to SC A/ B conditions and not to other SCs. Keeping in mind the damages and induced effects witnessed during various EQs in NE India, GMPE should be able to predict ground motion at various SCs. For this reason, in this chapter, an attempt is made to develop a new GMPE for the NE India region. The lack of adequate ground motion records (discussed in Chapter 2) is dealt with by using synthetic ground motions generated based on regional parameters and available model. Further, site coefficients are estimated based on ground response analyses to utilise the present GMPE for different SCs. This Chapter presents the entire procedure followed for the development of GMPE for NE India.

7.2 Generation of Synthetic Ground Motions

Synthetic ground motions have been used successfully around the world for seismic hazard studies where either recorded ground motions are not available or are available in scarce (Nath et al. 2005; Rietbrock et al. 2013; Edwards and Fäh 2013; Drouet and Cotton 2015; D'Amico et al. 2018). Brune (1970) gave a point source model to generate synthetic ground motion. Later, Boore (1983) modified Brune's (1970) point source model and introduced a band-limited white noise stochastic model to generate synthetic ground motions based on the point source model. This modified model of Boore (1983) was used in many studies (Boore and Atkinson 1987; Irikura 1983; Sitharam and Anbazhagan 2007) to simulate ground motions and estimate seismic hazard values. It must be highlighted here that the ground motion characteristics (amplitude, frequency content and duration) at a site are controlled by the directivity and rupture propagation path effect (Hartzell 1978; Irikura 1983). However, according to Anbazhagan et al. (2013) and Harinarayan and Kumar (2020), the point source model is unable to incorporate those effects accurately. Such limitation can be overcome by considering a finite fault model for generating synthetic ground motions. The finite fault model was first introduced by Hartzell (1978), where the entire rupture area was divided into several sub-faults. Further, each sub-fault was modelled as a single point source, and the main event was obtained by adding the contribution from each sub-fault considering suitable time delay (Hartzell 1978; Hanks and McGuire 1981; Heaton and Hartzell 1989). Finite fault models have been effectively used worldwide for seismic hazard analysis (for e.g., Kanamori, 1977; Heaton and Hartzell, 1989; Irikura, 1983). In most of the above-mentioned studies, however, the source and path model were considered by means of empirical correlations.

Later, Beresnev and Atkinson (1997) introduced the FINSIM code developed on the finite fault simulation model based on the shear dislocation theory of Richards and Aki (1980). The seismic source and propagation path attenuation components were included using the time function, which characterizes the real physical process. Afterwards, Motazedian and Atkinson (2005) improved the FINSIM model by incorporating the concept of dynamic corner frequency (f_0) and developed EXSIM (Extended Finite-Fault Simulation). The EXSIM code characterizes f_0 as a time-dependent parameter, and the rupture history determines the frequency content of each simulated sub-fault's time series. The rupture begins with a substantial value of f_0 and progresses towards a lesser value of f_0 as the ruptured area expands. The rupture process is designed such that the slip takes place on a specific section of the fault (sub-fault) at any given time while other sub-faults at that time remain idle. A ruptured sub-fault governs its f_0, whereas the amplitude of ground motion is limited by a set of sub-faults in the f_0 calculations (Motazedian and Atkinson 2005). In the current study, keeping in mind that limited regional ground motion records are available for the study region and further such motions are not available from across NE India, a larger dataset of synthetic ground motions is generated using EXSIM.

EXSIM models the acceleration spectrum of the shear-wave portion of the ground motion of j^{th} sub-fault at i^{th} recording station ($U_{ij}(f)$) as:

$$U_{ij}(f) = \left\{ \frac{\left[\frac{R^{\theta}FV}{4\pi\rho\beta_s^{3}}\right] H_{ij} M_{0j} (2\pi f)^2)}{\left[1+\left[\frac{f}{(f_{0j})}\right]^2\right]} \right\} \left\{ e^{\left(-\pi f R_{ij}/_{Q_s\beta_s}\right)} Gr \right\} \left\{ e^{(-\pi\kappa f)} \right\} \{D_i\} \qquad 7.1$$

in Eq. 7.1, $e^{(-\pi\kappa f)}$ represents the near-surface high-cut filter, where κ represents the linear decay of spectral amplitude of the S-wave part of the accelerogram at higher frequencies. Furthermore, $\left\{ e^{\left(-\pi f R_{ij}/_{Q_s\beta_s}\right)} Gr \right\}$ defines the path attenuation factor, where Q_s is the quality factor representing anelastic attenuation of the study area. Further, Gr in Eq. 7.1 is the geometric spreading. The term $\left\{ \left[\frac{R^{\theta}FV}{4\pi\rho\beta_s^{3}}\right] H_{ij} M_{0j} (2\pi f)^2)/[1+\left({}^{f}/_{f_{0j}}\right)^2] \right\}$ constitutes source factor, where M_{0j} is the seismic moment and f_0 is the corner frequency. Term D_i in Eq. 7.1 stands for site component. In Eq. 7.1, R^{θ} represents radiation pattern, F is the free surface amplification, V deals with the partitioning of energy in orthogonal directions, ρ is the crustal density, β_s is the crustal shear wave velocity and H_{ij} is the scaling factor responsible for converging the energy at the high-frequency spectral level of the sub-fault. A comprehensive description of the EXSIM model is not given in this Chapter and can be read from Motazedian and Atkinson (2005).

7.3 Regional Parameters for Ground Motion Simulation

To develop a GMPE, a regional dataset of recorded ground motions over a broader range of magnitude and distance is required. For NE India, however, the regional ground motion records database has gaps in terms of higher M_w and larger distance. In addition, an improper understanding of SCs adds uncertainty if recorded ground motions from various recording stations are used to develop GMPE for any SCs. As a result, a larger set of ground motion records is generated synthetically in this work by using EXSIM rather than taking recorded ground motions into account.

Table 7.1 Details of EQs considered for synthetic ground motion generation along with corresponding model parameters

Sl. No.	Date (dd-mm-yyyy)	Lat. (°) (N)	Lon. (°) (E)	Focal Depth (Km)	Mw	Stress drop (MPa)	Source of Stress drop	Strike (°)	Dip (°)	Source of Fault plane solution
1	1950-Assam 03-05-1950	28.4	96.7	35	8.7	6.6	Nath et al. (2009)	333.5	57.5	Nath et al. (2009)
2	1987 Shilong (12-06-1987)	26.0	91.0	35	8.1	15.9	Nath et al. (2009)	292	40	Nath et al. (2009)
3	1934-Bihar	27.0	87.0	34.2	8.2	27.5	Bajaj and Anbazhagan, (2019)	275	18	Pettenati et al. (2017)
4	2015 Nepal (12-05-2015)	27.8	86.06	15	7.3	3.6	Prakash et al. 2016	312	11	Prakash et al. 2016
5	1988-Manipur (06-08-1988)	25.1	95.1	91	7.2	28.2	Nath et al. (2009)	284	45	Nath et al. (2009)
6	2011-Sikkim (18-09-2011)	27.7	88.1	20	6.9	11.5	Chopra et al. (2014)	220	78	Chopra et al. (2014)
7	2016-Myanmar (24-	21.1	94.52	90	6.8	2.0	Raghukanth and	323	8	Shiddiqi et al (2018)

	08-2016)						Somala ,(2009)			
8	6-5-1995	25.01	95.34	122	6.4	12.8	Raghukanth and Somala , (2009)	281	48	Kumar et al. (2015)
9	8-5-1997	24.89	92.25	34	6.0	15.6	Raghukanth and Somala , (2009)	239	79	Kumar et al. (2015)
10	06-11-2013	29.5	93.5	10	5.2	11.9*	Kumar et al. (2019)	310	75	Kumar et al. (2015)
11	11-05-2012	26.6	93.0	20	5.4	3.5	Kumar et al. (2019)	332	84	Kumar et al. (2015)
12	29-07-2012	22.83	94.32	55	5.8	6.64	Kumar et al. (2019)	176	55	Kumar et al. (2015)
13	19-08-2012	22.9	94.3	41	4.0	7.3	Kumar et al. (2019)	120	77	Kumar et al. (2015)
14	04-05-2012	27.50	95.10	15	4.4	1	Kumar et al. (2019)	295	11	Kumar et al. (2015)
15	19-08-2009	26.51	92.45	46	4.8	5.78	Kumar et al. (2019)	200	40	Kayal et al. (2012)
16	01-05-2012	26.2	92.3	15	3.7	2.1	Kumar et al. (2019)	90.4	70	Kumar et al. (2015)

Keeping in mind the seismicity of NE India, the magnitude of past EQ that occurred in the region and the possibility of induced effects even during distant EQ, the development of synthetic ground motions is done in this work up to a maximum value of epicentral distance of 400km and for magnitude range between 4 and 9 (Mw). Ground motion simulation models requires various parameters to generate synthetic ground motions. These parameters include; Crustal density (ρ), Crustal shear-wave velocity (β_S), stress drop ($\Delta\sigma$), strike and dip of the faults, Kappa, Quality factor (Q_s), Kappa factor (κ_H), Geometric spreading, Rupture propagation speed, Duration model, Radiation coefficient etc. Following Hazarika and Kumar (2012), β_S and ρ are considered as 3.8km/s and ρ=2.7g/cc respectively in all the simulations. $\Delta\sigma$ is a critical ground motion model parameter which controls the amplitude of the acceleration spectrum at higher frequencies. To understand the $\Delta\sigma$ variations in NE India, $\Delta\sigma$ values of 16 past EQs are obtained from various literature as summarized in Table 7.1 (Nath et al. 2009; Raghukanth and Somala 2009; Kayal et al. 2012; Chopra et al. 2014; Kumar et al. 2015, 2019; Prakash et al. 2016; Pettenati et al. 2017; Shiddiqi et al. 2018; Bajaj and Anbazhagan 2019). Corresponding values of Mw, strike, dip of the past EQs are also summarized in Table 7.1. Based on $\Delta\sigma$ values of these past EQs, synthetic ground motions will be generated considering $\Delta\sigma$ values ranging from 20bars to 300bars. In case of EXSIM based ground motion simulation, two other important parameters required are; rupture length (L) and rupture area (A). In the present work, these two parameters for each EQ are calculated as per Wells and Coppersmith (1994) based on the magnitude of the EQs.

For all the simulations, the hypocenter is positioned at the centre of the main fault, following Anbazhagan et al. (2013). The rupture process begins at the hypocenter and extends radially to the neighbouring subfaults with a rupture propagation velocity yβ. The value of "y" is considered as 0.8 (following Atkinson and Boore, 2006). Q_s is a region-specific factor responsible for Fourier spectrum attenuation due to the propagation path. For the present study, the two Q_s correlations given by Raghukanth and Somala (2009) for NE India are considered. The geometry factor ($\mathrm{Gr}[\mathrm{f}, \mathrm{R}]$) relation as per Singh et al. (2004) is used for the study region in this work. κ_H, which governs the attenuation of the Fourier spectrum at high frequency, is also a regional factor. For the present study, κ_H is considered to be equal to 0.006 for hard rock corresponding to shear wave velocity (Vs) of 2800m/s following Singh et al. (2016) and Campbell (2003). Further, the ground motion duration term (T) is considered as per Bajaj and Anbazhagan (2019) for all the simulations. Various parameters and correlations which are used to estimate the above-mentioned parameters are summarized in Table 7.2. In the present work, the above-mentioned model parameters are utilized to simulate synthetic ground for bedrock conditions motions by using EXSIM without considering any site amplification. The procedure is described in the following section.

Table 7.2 List of modelling parameters used for the simulation in present work

Parameters	Value	Distribution	Reference
Crustal density (ρ)	2.7g/cc	Constant	Hazarika and Kumar, (2012)
Crustal shear-wave velocity (β_s)	3.8km/s	Constant	Hazarika and Kumar, (2012)
Kappa (κ_H)	Rock Site: 0.006s	Constant	Singh et al. (2016) Campbell (2003)
Quality factor	$Q = (431 \pm 9)f^{(0.73 \pm 0.01)}$ for Indo-Burma $Q = (224 \pm 10)f^{(0.93 \pm 0.01)}$ for Shillong Plateau	Normal	Raghukanth and Somala (2009)
Geometric spreading	$\mathrm{Gr}[f,R] = \begin{bmatrix} 1/R & R \le 100\text{ km} \\ 1/\sqrt{100R} & R > 100\ km \end{bmatrix}$	Constant	Singh et al. (1999), Singh et al. (2004)
Rupture propagation speed	0.8 β_s		Atkinson and Boore, (2006)
Duration model	$T = R \times 16.8/60$ R<60 km $T = 16.8 + 0.05(R - 60)$ R ≤ 60 km	Constant	Bajaj and Anbazhagan, (2019)
Site component	Bedrock level		

M	4-9Mw	Uniform	Based on past EQ data
R	Up to 400 km	Uniform	
Depth	15-45km	Uniform	Based on past EQ data
Stress drop	20-300bars	Uniform	Based on past EQ data
Fault Dip	Reverse fault: FD≤40° Normal fault: 40°<FD≤75° Strike-slip fault:75°<FD	Uniform	Bajaj and Anbazhagan, (2019)
Azimuth of the apparent stations	0°-360°	Uniform	
Radiational coefficient (R^{θ})	0.48-0.64	Uniform	Singh et al. (2016)
Cut off Frequency	20–50 Hz	Uniform	Singh et al. (2016)

7.4 Development of Synthetic Ground Motions for Bedrock Condition

It should be noted that the variations in the model parameters mentioned above can alter the simulated ground motion and, subsequently the proposed GMPE. Hence, the effect of possible variations in each parameter should be considered to address such uncertainties and minimize the error in the proposed GMPE. Different model parameters associated with synthetic ground motion generation include: κ_H, Q_s, $\Delta\sigma$, path attenuation parameters, fault plane solution (strike and dip values) and hypocentral distance. For the generation of the synthetic ground motions, the values of some model parameters such as ρ, β_s, and κ_H are kept the same throughout all the simulations. Similarly, some model parameters, such as $Gr[f, R]$, and T are considered as functions of R. Other parameters, such as $\Delta\sigma$, Q_s, radiational coefficient (R^{θ}), cut-off frequency, fault dip etc., possible variations are considered as per the past literature. Table 7.2 shows the variation in the values of these parameters considered in this study. To include the variation in $\Delta\sigma$, the value of $\Delta\sigma$ is considered to be uniformly distributed between the values 20bars and 300bars. R^{θ} is considered uniformly varying between 0.48 and 0.64 (Singh et al. 2016). The cut-off frequency is considered to vary uniformly between 20 to 50 Hz (Singh et al. 2016). In the present study, ground motions are separately generated for reverse, normal and strike-slip faults. For reverse mechanism, the fault dip is considered to be uniformly distributed between 0° and 40°. For the strike-slip mechanism, the dip is considered to be uniformly distributed between 75° and 90°. For the normal mechanism, the dips are considered to be uniformly distributed between 40° and 75° (following Bajaj and Anbazhagan 2019). In order to incorporate the concept of source finiteness into the simulations, the present work generates the positions of apparent stations as defined by Anbazhagan et al. (2013), by assuming a uniform distribution of epicentral distances ranging from 0km to 400km around the epicentres with azimuth to vary between 0° and 360° uniformly. As a result of the various fault plane solutions and the randomly chosen hypocentral distance, the position of the apparent station around the epicentre becomes a random variable for the simulation. As highlighted earlier, the two Q_s correlations, given by Raghukanth and Somala (2009) for Indo-Burma region and Shillong plateau region each are used in the present study (referring to NDMA, 2010). In both correlations, Q_s is related to frequency (f) by the relation $Q_s = Q_0 f^n$. Q_0 and n are considered to be normally distributed in the present work. For Indo-Burma region, the mean and standard deviation of Q_0 is considered to be 431 and 9 respectively, and the mean and the standard deviation in n is taken as 0.73 and 0.01 respectively. Similarly, for Shillong Plateau region, the mean and standard deviation in Q_0 is considered to be 224 and 10 respectively. Further, the mean and standard deviation of n is taken as 0.93 and 0.01 respectively. Ground motions are generated considering above two correlations separately. For separate fault mechanisms and Q_s correlation, a total 1000 synthetic ground motions are generated for bedrock condition. To generate 1000 synthetic ground motions, 1000 sets of input model parameters are selected randomly based on Latin Hypercube sampling technique with reduced correlation following above mentioned variations in model parameters (McKay et al., 2000). The distribution type, range, mean and standard deviation of these parameters are summarized in Table 7.2. In total 6000 (=1000×3×2) ground motions are simulated for bedrock condition. For all the simulations, a damping value of 5% is considered.

7.5 Development of Ground Motion Prediction Equation for Bedrock Level

In Section 7.4 , 2000 ground motions are simulated for bedrock conditions for each faulting type. Hence, a total of 6000 simulated ground motions at bedrock are now available in order to develop new GMPE following a suitable functional form. It should be highlighted here that numerous functional forms of regression models for GMPEs are available in the literature

(Abrahamson and Litehiser 1989; Atkinson and Boore 2006; Campbell 1981; Fukushima and Tanaka 1990). In general, the functional form of any GMPE is decided based on the availability of information related to EQ parameters for the region. In the present case, two components, namely M and R are considered for the regression analysis and accordingly the following functional form (similar to NDMA 2010) is considered for the proposed GMPE;

$$ln\left(\frac{S_a}{g}\right) = C_1 + C_2 M + C_3 M^2 + C_4 R + C_5\, ln(R + C_6 e^{C_7 M}) + C_8\, log(R)\, f_R + ln(\varepsilon) \qquad 7.2$$

where, $f_R = max(ln\, R/100)\, , 0)$, S_a represents the spectral acceleration, M is the EQ magnitude in Mw scale, R is the hypocentral distance in kilometres, C_1 to C_8 are frequency-dependent regression coefficients, which are estimated based on nonlinear regression analyses. These values of the above regression coefficients (C_1 *to* C_8) are determined separately for the three different types of faults, as well as combining all the types of faults. Table 7.3, Table 7.4, Table 7.5, and Table 7.6 summarise the values of C_1 *to* C_8 for different frequencies, up to 100Hz for Normal, Reverse, Strike-slip and all faulting types respectively.

Using the above values of regression coefficients, once the value of M, R and fault types are known (or unknown), values of spectral acceleration at various frequencies of interest can be determined for bedrock level using the above proposed GMPE. In the correct form, proposed GMPE can only be used for bedrock condition. In order to determine the Sa value for different SCs, site coefficients can be used in addition to the components of bedrock motion. Next section discusses the procedure for determining the site coefficient for different SCs.

Table 7.3 Regression coefficient for Normal fault

f (Hz)	C1	C2	C3	C4	C5	C6	C7	C8	σ (ln(Sa/g))
0.25	-27.532	6.181	-0.311	-0.003	-1.128	0.012	0.911	0.227	0.43
0.40	-25.483	6.213	-0.329	-0.003	-1.272	0.801	0.434	0.249	0.44
0.50	-24.347	6.074	-0.325	-0.002	-1.305	0.880	0.401	0.203	0.44
0.81	-20.933	5.299	-0.284	-0.002	-1.218	0.574	0.413	0.192	0.47
1.00	-18.634	4.929	-0.260	-0.003	-1.326	1.798	0.346	0.228	0.47
1.25	-17.302	4.479	-0.233	-0.004	-1.197	0.879	0.391	0.262	0.48
1.60	-15.046	3.968	-0.200	-0.003	-1.232	0.661	0.420	0.232	0.48
2.00	-12.413	3.479	-0.166	-0.003	-1.379	1.226	0.397	0.255	0.48
2.50	-11.097	3.165	-0.146	-0.003	-1.383	1.047	0.408	0.288	0.48

3.33	-10.132	2.861	-0.126	-0.004	-1.301	0.619	0.458	0.325	0.49
5.00	-8.217	2.298	-0.088	-0.003	-1.252	0.101	0.656	0.213	0.49
10.00	-6.236	1.727	-0.049	-0.005	-1.150	0.003	1.066	0.318	0.49
100.00	-6.730	1.778	-0.051	-0.005	-1.259	0.018	0.890	0.336	0.50

Table 7.4 Regression coefficient for Reverse fault

f (Hz)	C1	C2	C3	C4	C5	C6	C7	C8	σ (ln(Sa/g))
0.25	-28.008	6.222	-0.315	-0.004	-1.023	0.001	1.211	0.263	0.43
0.40	-25.964	6.266	-0.333	-0.004	-1.184	0.349	0.546	0.290	0.44
0.50	-24.707	6.170	-0.333	-0.003	-1.280	0.642	0.462	0.237	0.45
0.81	-21.506	5.425	-0.295	-0.003	-1.166	0.294	0.491	0.230	0.47
1.00	-19.118	5.028	-0.269	-0.003	-1.283	0.898	0.433	0.239	0.48
1.25	-17.757	4.548	-0.240	-0.004	-1.129	0.345	0.505	0.303	0.49
1.60	-15.559	4.051	-0.208	-0.004	-1.157	0.245	0.547	0.290	0.49
2.00	-12.982	3.518	-0.171	-0.004	-1.272	0.385	0.533	0.263	0.49
2.50	-12.004	3.234	-0.153	-0.004	-1.220	0.203	0.591	0.297	0.49
3.33	-10.924	2.913	-0.131	-0.005	-1.153	0.070	0.721	0.336	0.49
5.00	-8.814	2.322	-0.091	-0.004	-1.129	0.002	1.123	0.195	0.50
10.00	-6.721	1.777	-0.053	-0.007	-1.045	0.002	1.170	0.382	0.50
100.00	-7.190	1.789	-0.053	-0.005	-1.154	0.001	1.308	0.303	0.50

Table 7.5 Regression coefficient for Strike-slip fault

f (Hz)	C1	C2	C3	C4	C5	C6	C7	C8	σ (ln(Sa/g))
0.25	-28.138	6.291	-0.320	-0.005	-1.017	0.070	0.760	0.352	0.44
0.40	-25.737	6.266	-0.333	-0.005	-1.199	1.344	0.426	0.395	0.45
0.50	-24.551	6.153	-0.332	-0.004	-1.271	1.664	0.381	0.319	0.46
0.81	-21.457	5.391	-0.292	-0.004	-1.118	0.897	0.406	0.330	0.48
1.00	-18.571	5.025	-0.267	-0.005	-1.344	3.069	0.355	0.393	0.49
1.25	-17.338	4.564	-0.240	-0.006	-1.189	1.704	0.391	0.415	0.49
1.60	-15.132	4.057	-0.208	-0.006	-1.211	1.520	0.397	0.405	0.50
2.00	-12.021	3.564	-0.172	-0.005	-1.452	2.952	0.366	0.401	0.50
2.50	-10.996	3.243	-0.152	-0.006	-1.392	2.183	0.387	0.443	0.50

3.33	-10.369	2.937	-0.132	-0.006	-1.250	1.021	0.467	0.452	0.50
5.00	-8.468	2.397	-0.096	-0.005	-1.216	0.398	0.566	0.316	0.50
10.00	-6.690	1.851	-0.059	-0.008	-1.075	0.113	0.721	0.451	0.51
100.00	-7.206	1.942	-0.063	-0.007	-1.216	0.192	0.672	0.428	0.51

Table 7.6 Regression coefficient for All fault

f (Hz)	C1	C2	C3	C4	C5	C6	C7	C8	σ (ln(Sa/g))
0.25	-27.928	6.228	-0.315	-0.004	-1.046	0.009	0.973	0.280	0.43
0.40	-25.743	6.253	-0.332	-0.004	-1.218	0.790	0.458	0.314	0.44
0.50	-24.550	6.135	-0.330	-0.003	-1.284	1.030	0.409	0.255	0.45
0.81	-21.323	5.375	-0.290	-0.003	-1.164	0.553	0.431	0.253	0.47
1.00	-18.819	4.997	-0.266	-0.004	-1.311	1.780	0.373	0.290	0.48
1.25	-17.521	4.533	-0.238	-0.005	-1.162	0.835	0.424	0.328	0.49
1.60	-15.302	4.029	-0.206	-0.004	-1.191	0.665	0.447	0.311	0.49
2.00	-12.570	3.524	-0.170	-0.004	-1.351	1.204	0.423	0.307	0.49
2.50	-11.452	3.220	-0.151	-0.004	-1.318	0.896	0.444	0.344	0.49
3.33	-10.550	2.910	-0.130	-0.005	-1.223	0.411	0.529	0.372	0.49
5.00	-8.608	2.343	-0.092	-0.004	-1.179	0.057	0.748	0.240	0.50
10.00	-6.632	1.763	-0.052	-0.007	-1.058	0.003	1.114	0.378	0.50
100.00	-7.142	1.827	-0.055	-0.006	-1.183	0.011	0.969	0.352	0.50

7.6 Determination of Site Coefficients

GMPE developed in the last section is corresponding to bedrock condition. Knowledge about regional site amplification is required to obtain the ground motion level for different SCs. In this chapter, in order to obtain the surface level ground motion level ($SA_{surface}$) or ground motion spectra, an amplification factor (Amp) for bedrock-level acceleration (SA_{rock}) is utilised as;

$$SA_{surface} = Amp \times SA_{rock} \qquad 7.3$$

The dependency of the amplification factor (Amp) with SA_{rock} values for all SCs can be expressed by the following amplification model (referring to the work by Singh et al. 2016):

$$ln(Amp) = A' \times SA_{rock} + B' + ln\, \sigma_s \qquad 7.4$$

Where, A' and B' are the regression coefficients or the site coefficients.

In order to determine the values of A' and B' for various SCs, extensive Ground Response Analysis (GRA) is done for SC A, B, C, D and E, as discussed in further subsections.

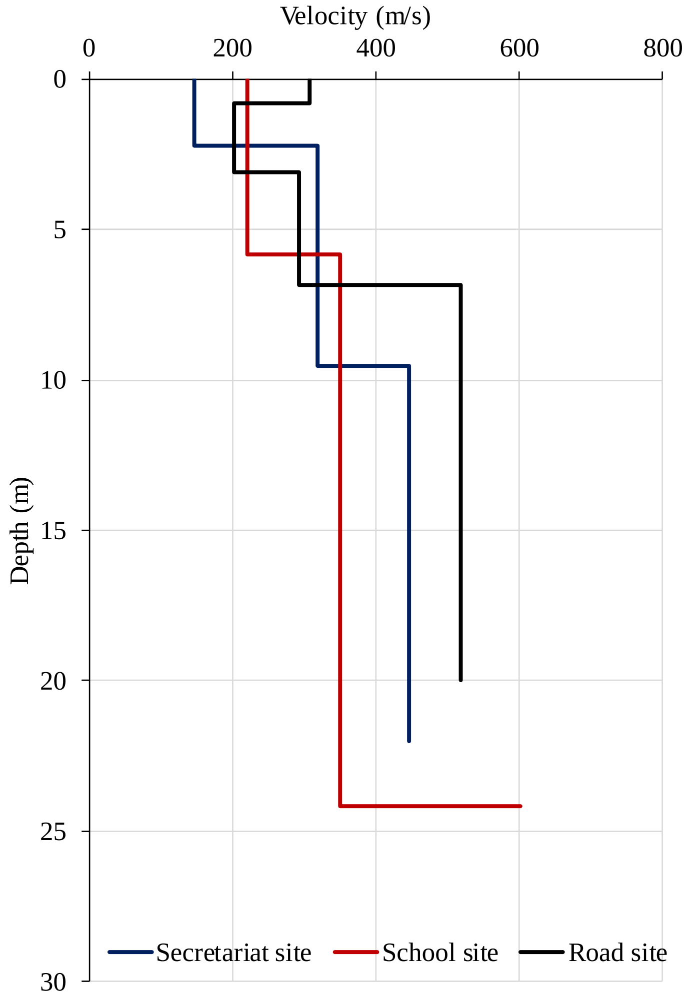

Figure 7.1 Vs profiles at selected sites in Sikkim (modified after Anbazhagan and Murali Krishna 2015)

7.6.1 Lithology

To account for change in ground motion due to local soil for NE India, available soil profiles from multiple locations across NE India are considered in the present analysis. These include 100 boreholes and 4 V_s profiles from different locations across NE India. Three V_s profiles [corresponding to Secretariat site (SS), School site (SCS) and Road site (RS)] from Gangtok, Sikkim are collected from Anbazhagan and Krishna (2015). Similarly, one V_s profile from Meghalaya [Mawprem site (MS) in Shillong] is collected from Baruah et al. (2020). Further, fifty borelogs (up to 30m depth) for Guwahati city each from SC D and SC E are obtained from Guwahati Metropolitan Development Authority (GMDA, 2016). V_s profiles obtained from Sikkim are shown in Figure 7.1. From Figure 7.1, it can be observed that across various locations within Sikkim, the V_s profiles vary. While the Secretariat site (SS) and School site (SCS) belong to SC D with V_{s30} value of 262m/s and 338m/s, respectively, the Road site (RS) belongs to SC C with V_{s30} value of 384m/s. According to Anbazhagan and Krishna (2015), lithology at these three sites consists of sandy soil with loose to medium density followed by bedrock. V_s profile Mawprem site (MS) in Shillong, obtained from Baruah et al. (2020) is shown in Figure 7.2. This V_s profile has an average V_{s30} value of 457m/s (SC C). According to Baruah et al. (2020), the lithology of Shillong city is dominated by soft rocks and stiff soil with V_s value in the top 30m ranging from 456.8m/s to 1122.9m/s.

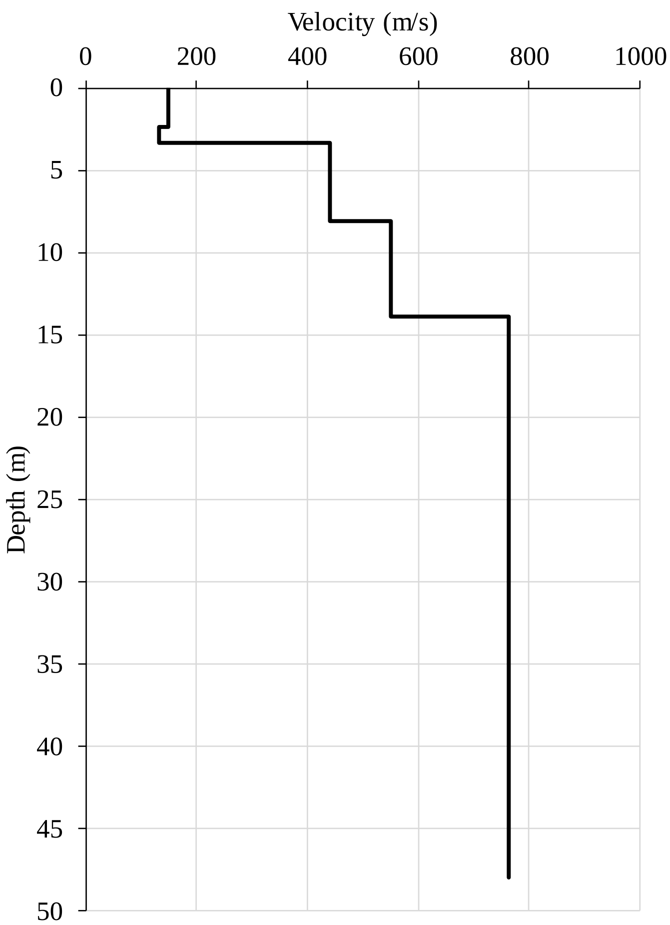

Figure 7.2 Vs profile at Mawprem area in Shillong (modified after Baruah et al. 2020)

One typical borelog (belonging to SC E) from Guwahati is also shown in Table 7.7. It can be observed from Table 7.7 that the subsoil lithology for Guwahati mainly consists of clay having low to high compressibility, followed by silty sand layers. Similar types of soil were also reported by Raghukanth et al. (2011) for Guwahati city. According to Halder et al. (2019), soft soil (SPT-N<15 according to (NEHRP & BSSC 2003) is present up to a depth of 10m. Stiff soil (having 15<SPT-N<50) is found between 10m and 15m depth, followed by hard soil (SPT-N>50). The depth of the ground water table (GWT) in all these boreholes was reported to vary from 0m to 6.7m after 24 hours of observation. Similar information is obtained from all the 100 boreholes for Guwahati. Borelogs for Guwahati have SPT-N value and not Vs values directly. These SPT-N values are converted to V_s values by using correlation developed by Kumar et al. (2018) for Guwahati. Thus, collective information on subsoil lithology is obtained in the form of Vs and (or) SPT-N variation with depth from different locations of NE India. Using the above information collectively, random Vs profiles are generated to account for possible subsoil layer stratification configurations and incorporate them while performing GRA, as discussed in the next subsection.

Table 7.7 Typical borlog from Guwahati (SC E)

Borhole 21, Location Garigaon			
GWT at 1.3m below GL			
Depth (m)	soil type	Bulk Density (g/cc)	N-value
1.5		1.66	4
3		1.62	3
4.5		1.66	2
6		1.66	2
7.5		1.66	3
9		1.66	11
10.5		1.66	10
12		1.66	9
13.5		1.67	11
15	CL	1.67	6
16.5		1.67	9
18	CH	1.68	14
19.5		1.69	21
21	SP	1.76	31

22.5		2.21	50
24		1.68	13
25.5		1.67	10
27	CI	1.68	14
28.5		1.69	21
30	SP	1.69	22

7.6.2 Random V_s profiles

NE India covers a larger geographical area. Hence, to determine site amplification spectra for different SCs, ideally, an adequate number of V_s profiles in the top 30m, corresponding to each of the SCs should be known from across NE India. However, for many locations in NE India, such subsoil information is not available in the existing literature. To overcome this limitation, a larger set of random Vs profiles are generated using known though limited information for the NE India. As discussed above, Vs profiles for SC C, D and E are available for parts of NE India (for Sikkim, Shillong and Guwahati regions). In the present work, the layer thicknesses in each known Vs profile are used as stratification details of random profiles. The stratification details such as, soil type, their density, and the maximum and minimum Vs values are utilised for the generation of the random Vs profiles. 100 random Vs profiles are generated using Matlab command “rand” based on the above stratification details for each of the three SCs (SC C, D and E). These random Vs profiles are generated, ensuring that V_{s30} values remain within the range of V_{s30} of each SC. This way, referring to the actual Vs profiles, 100 random Vs profiles for each of the three SCs are generated. In total, 300 (3x100) Vs profiles are generated randomly. These include 100 random Vs profiles for SC C, D and E each. Figure 7.3 a, b, and c depict 100 random V_s profiles corresponding to SC C, D, and E each respectively. It can be observed from Figure 7.3 (a) that the V_s values in these profiles vary from 150m/s to 1500m/s for SC C up to 30m depth. Similarly, Figure 7.3 (b) depicts the details of 100 V_s profiles that are generated based on the stratification details obtained from SC D. It can be observed from Figure 7.3 (b) that V_s values in the range of 80m/s to 649m/s are present. Further, 100 random profiles belonging to SC E (that are generated based on stratification detail from MS sites in Shillong and Guwahati) are shown in Figure 7.3 (c). The V_s values, in this case, vary from 70m/s to 299m/s (see Figure 7.3 c). It should be noted here that all the 300 randomly generated V_s profiles are for 30m depth, ensuring V_{s30} value of each of the 300 profiles falls within the range of V_{s30} for that SC (SC C:360-760 m/s, SC D:180-360 m/s, SC E:<180m/s).

For SCs A and B, information related to average layer thickness, average V_s values and average density are taken from Singh et al. (2016). Based on these average values, random Vs profiles are generated following the method mentioned above, ensuring Vs values remain within the range of Vs of each SC (SC A: >1500m/s, SC B: 760-1500m/s). For each SC (A and B), a total of 100 random boreholes are generated.

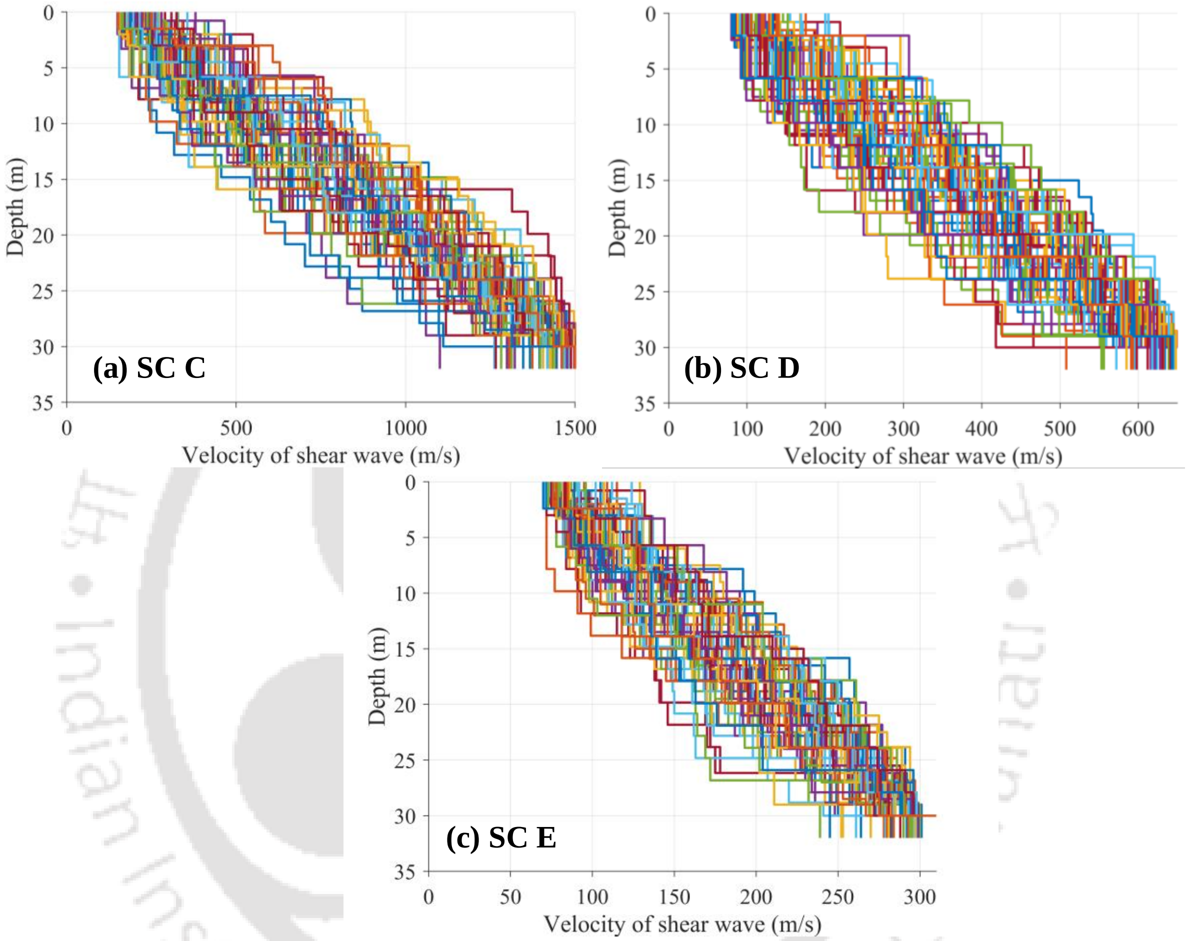

Figure 7.3 Randomly generated Vs profiles 100 for SC C, D, and E each

Based on the discussion in the last two paragraphs, a total of 500 synthetic random Vs profiles (100 corresponding to SC A, B, C, D and E each) are now available. Ultimately, all these 500 Vs profiles are used for Equivalent Linear Ground Response Analyses (ELGRAs). Further details on the ELGRAs are given in the next section.

7.6.3 ELGRA

Using 500 randomly identified Vs profiles as discussed in the last subsection, ELGRA is done in this section to determine site amplification spectra for SC A, B, C, D and E. Details on the selection of dynamic soil properties, input motions considered, and the analyses results can be found in the subsections.

7.6.3.1 Dynamic soil properties

Surface geology and subsoil geotechnical and geophysical properties are the most important input parameters for GRA of any location. As the different levels of ground motion can generate a wide range of shear strains (γ) in the soil, dynamic soil properties define the soil resistance in terms of shear modulus (G) and damping ratio (β), which the soil can offer against seismic excitation at any γ. These are expressed as variations in G (or G/G_{max}) and β versus γ for each soil type or dynamic soil property curves (DSPCs). G/G_{max} and β curves of a soil type can be obtained from different laboratory tests like cyclic tri-axial and resonant column tests (Doroudian and Vucetic 1995; Stewart et al. 2000, 2001). In the absence of regionally available DSPCs, depending upon site-specific soil characteristics such as soil type, over consolidation ratio, Plasticity Index (PI) etc., suitable DSPCs out of widely adopted curves for each material type, can be selected. In this analysis, five types of soils are considered depending upon the information obtained from earlier discussed V_s profiles and borelogs. These include silty sand, low plasticity clays (CL), medium plasticity clays (CI), high plasticity clays (CH) and rock. Thus, while defining DSPCs, G/G_{max} and β curves for average sand by Seed and Idriss (1970) are used for silty sand. For clay, appropriate G/G_{max} curves (based on PI value) developed by Sun et al. (1988) are selected. Further, as per Seed and Idriss (1970), β curve for clays are independent of PI, and thus average β curve for clays as per Seed and Idriss (1970) is used in the present analysis. Further, G/G_{max} and β ratio curves as per Schnabel (1973) for dense soils/ rocks are used for the modelling of rock layers. Figure 7.4 summarises all the selected G/G_{max} and β curves for the present work.

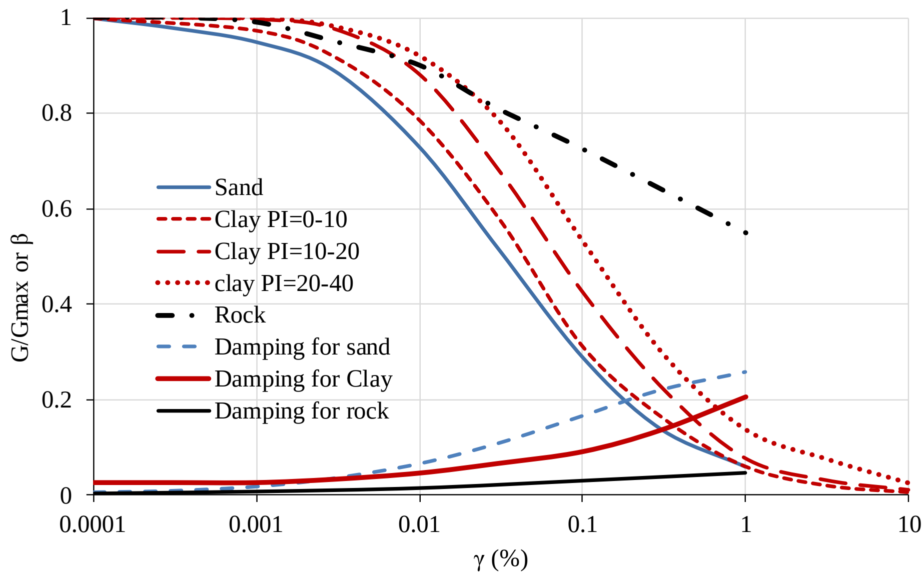

Figure 7.4 Details of DSPC used for ELGRA

7.6.3.2 Input motions

It is to be noted here that the same soil can show no to large amplification to base motion depending upon input motion characteristics. Hence, in addition to the DSPCs of soil, the selection of input motion for ELGRA is another important input. For scenario-specific study, corresponding ground motion can be utilized. However, for the present work, site amplification coefficients, which are valid for wider range of ground motions, are required to be estimated for each of the SCs separately. Hence, a larger set consisting of 540 synthetically generated ground motions at bedrock level are used as input motions. These ground motions are generated considering different combinations of magnitude and distances, and using model parameters discussed in section 6.5 earlier. The combinations of magnitudes and distances are taken in such a way that the generated ground motions contain a wide range of PGA values (here from 0.001g to 0.7g). To achieved this, the magnitude are taken in the range between 4 to 9Mw and epicentral distance are taken in a range from 10 km to 150 km. These 540 synthetic bedrock-level ground motions will be used as input motions for ELGRA in the present study.

7.6.3.3 Analyses

ELGRAs are performed using the MATLAB code developed by Kumar and Mondal (2017). Each of the 500 V_s profiles is modelled in layers of not more than 4m thickness. In case a soil layer has a thickness of more than 4m, it is subdivided into sublayers of 4m or lesser thicknesses for the modelling. Value of densities (ρ) for different values of Vs, as observed from borelogs for Guwahati, Sikkim and Shillong, are referred to assign ρ for each of the soil layers in the profiles according to the SCs. These ρ values are further used along with the V_s values to calculate G_{max} values. In addition, to initiate an analysis, an initial value of β of 5% (corresponding to low γ) is assigned to all the soil layers. Once all the input motions, dynamic soil properties, random Vs profiles and G_{max} values are ready, ELGRA are performed. It should be highlighted here that all the 540 ground motions are applied at the base of each of the 100 V_s profiles (refer to Section 7.6.2 earlier), and the responses at the ground surface are obtained. This way, a total of 540x100=54,000 ELGRAs are performed in this work for each SC, and the response spectra at the ground surface and at the bedrock level are obtained from each of the analyses.

Table 7.8 Coefficients for Site amplification for different frequencies

Frequency (Hz)	Site Coefficient												
	SC A (A'=0)		SC B (A'=0)		SC C			SC D			SC E		
	B'	$\ln(\sigma_S)$	B'	$\ln(\sigma_S)$	A'	B'	$\ln(\sigma_S)$	A'	B'	$\ln(\sigma_S)$	A'	B'	$\ln(\sigma_S)$
0.25	0.0005	0.0005	0.0065	0.0071	-0.0785	0.0187	0.0153	-1.8758	0.2608	0.2365	-2.5785	0.4802	0.2983
0.4	0.0010	0.0008	0.0122	0.0116	-0.0861	0.0333	0.0231	-0.6392	0.2202	0.1596	-0.0115	0.4715	0.2499
0.5	0.0013	0.0009	0.0157	0.0140	-0.1037	0.0458	0.0289	-0.3260	0.2349	0.1549	0.6115	0.5374	0.2482
0.81	0.0027	0.0013	0.0326	0.0260	-0.1281	0.0947	0.0543	0.0979	0.4194	0.2402	0.0382	0.9371	0.3006
1	0.0037	0.0018	0.0453	0.0380	-0.1269	0.1226	0.0697	0.1462	0.5615	0.2825	-1.3936	1.2777	0.3184
1.25	0.0051	0.0018	0.0610	0.0476	-0.2026	0.1956	0.1113	0.0173	0.7903	0.3268	-2.7712	1.6339	0.3477
1.6	0.0081	0.0025	0.0918	0.0652	-0.1790	0.2858	0.1442	-0.2805	1.1102	0.3792	-2.4134	1.4999	0.4268
2	0.0113	0.0034	0.1067	0.0602	-0.0339	0.3017	0.1446	-0.4812	1.2987	0.4312	-1.7228	1.2106	0.3382
2.5	0.0194	0.0064	0.1563	0.0786	-0.0111	0.4555	0.2059	-1.0162	1.6117	0.4538	-1.8126	1.1859	0.3340
3.33	0.0315	0.0060	0.2721	0.1392	0.0855	0.7423	0.3066	-1.4801	1.8413	0.5319	-2.3699	1.3331	0.4324
5	0.0719	0.0127	0.5794	0.1855	-0.1818	1.5350	0.4282	-1.1807	1.6347	0.4206	-2.1516	1.0527	0.4319
10	0.2865	0.0309	2.2525	0.3505	-0.2825	1.6509	0.3086	-1.3322	1.4545	0.5145	-1.6619	0.5388	0.4994
100	0.2666	0.0743	2.0649	0.2318	-1.4173	1.8549	0.2673	-2.7986	1.5504	0.3899	-3.1704	0.8885	0.3592

7.6.4 Determination of site coefficients for different SCs

As highlighted in the previous section, the response spectra at the bedrock level and corresponding response spectra at the ground surface are available for each SC from 54000 ELGRA. The amplification (Amp in Eq. 7.3) defined as the ratio of spectral ordinates between the ground surface and the bedrock at different frequencies, is calculated for each analysis based on surface response spectra and response spectra at bedrock. Once the value of SA_{rock} corresponding to each input motion and Amp is known, a regression analysis is carried out to determine the regression coefficients A' and B' (referring to Eq. 7.4). This exercise is repeated for different frequencies, and each of the five SCs (SC A, B, C, D and E).

Table 7.8 summarises the values of the above regression coefficients and standard deviation ($\ln \sigma_s$) for different frequencies and each SC. Once these values are known, the response spectra can be developed for different SCs based on the response spectra obtained from Eq. 7.3, and the site coefficients listed in Table 7.8.

At this stage, using the GMPE developed for bedrock level and the site coefficients obtained in last section, for different SCs are known. To obtain the ground motion level for different SCs, firstly, Eq. 7.2 can be utilised to obtain the bedrock ground motion level (SA_{rock}) for the required magnitude and hypocentral distance. Then, site coefficients of the required SC (from Table 7.8) are utilised to obtain Amp value using Eq. 7.4. Once Amp value is obtained, the same can be used to find the surface ground motion level using Eq. 7.3. By following the above mentioned steps, proposed GMPE for bedrock level and the site coefficients for specific SC can be utilised to determine ground motion level that SC for specific magnitude and distance values.

7.7 Validation of the GMPE

The values of regression coefficients determined in the previous section are based on the user-defined dataset. Thus, the effectiveness of the proposed GMPE should be verified by comparing the predicted ground motion response spectra with the regional records. For the purpose, ground motions recorded during various EQs, with recorded stations located at different SCs, are used as part of the validation. There are several ground motion recording stations spreading across the region of NE India. The records from such stations are available for different SCs in COSMOS (https://www.strongmotioncenter.org/vdc/scripts/default.plx) and PESMOS (https://pesmos.org/). It should be noted here that in the case of PESMOS, site classification is based on site geology and not based on subsurface exploration or V_{s30} based in-situ measurement. As per Mittal et al. (2012), the PESMOS station SCs are divided into three classes SC A, SC B, and SC C. V_{s30} for PESMOS SC A (firm/hard rocks) ranges from 700 m/s to 1620 m/s, which falls mainly under NEHRP SC A (> 1500 m/s) and B (760 m/s–1500 m/s). Similarly, PESMOS SC B (soft to firm rock having Vs^{30} in the range of 375 m/s–700 m/s) is equivalent to NEHRP SC C (V_{s30}: 360 m/s–760 m/s), and PESMOS SC C (soils having V_{s30} in the range of 200 m/s-375ms/) is equivalent to NEHRP SC D (V_{s30}: 180 m/s–360 m/s). Thus, in this Chapter, the results for SC A and B will be compared to the corresponding data from PESMOS SC A, SC C will be compared to the data from PESMOS SC B, and SC D will be compared to the data from PESMOS SC C. It should be noted here that as none of the available recording stations of PESMOS is located on SC E, validation of the proposed GMPE and Site coefficients for SC E is not performed here.

Some of the recording stations of PESMOS SC A conditions are located at Nongstoin, Gangtok and Shillong. Figure 7.5 shows the comparison of the response spectra based on recorded ground motions (two horizontal components) with the three sets of response spectra (Mean, Mean+ Standard deviation/ Mean+σ, and Mean-σ) obtained based on the proposed GMPE in this work and site coefficients for NEHRP SC B condition. The details of the EQs (magnitude, focal depth and epicentre location) considered in the study are collected mainly from the GCMT database. For those EQs which are absent in GCMT, the details are collected from the USGS EQ catalogue. The EQs considered for the validation include 11-08-2009 Indo-Myanmar EQ (5.5Mw), 03/06/2011 Nepal EQ (5.1Mw), 29/10/2009 Bhutan EQ (5.1Mw), 21/09/2009 Myanmar EQ (6.2Mw), 06/08/1988 Indo-Burma EQ (7.2Mw), and 26/02/2010 Tibet EQ (5.1Mw). It can be observed from Figure 7.5 that the response spectra based on the proposed GMPE are very close to the response spectra obtained as per recorded ground motions. However, for recordings in Gangtok station, it is seen that for higher natural frequency, the Sa value obtained based on the proposed GMPE is significantly higher than that of those based on the actual recording (Figure 7.5 b, d, g). Hence, the response spectra for this station are also compared with the response spectra obtained from the proposed GMPE and using the site coefficient for SC A condition (see Figure 7.6). From Figure 7.6, it can be observed that the response spectra based on the proposed GMPE considering SC A condition are very close to the response spectra as per recorded ground motions.

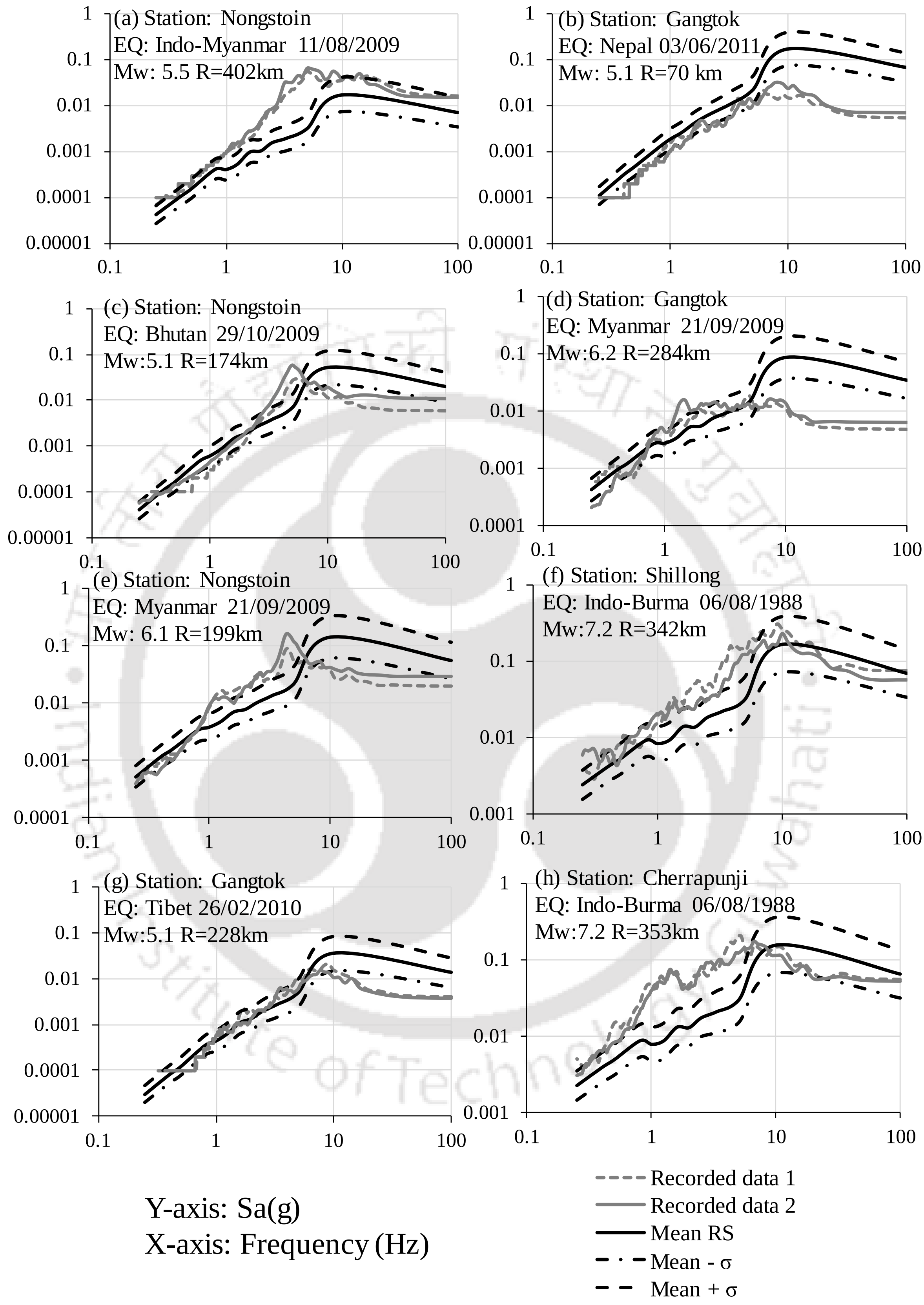

Figure 7.5 Comparison of response spectra from proposed GMPE and site coefficients for SC B with the recorded ground motions on Nongstoin, Shillong, Cherapunji, and Gangtok recording stations

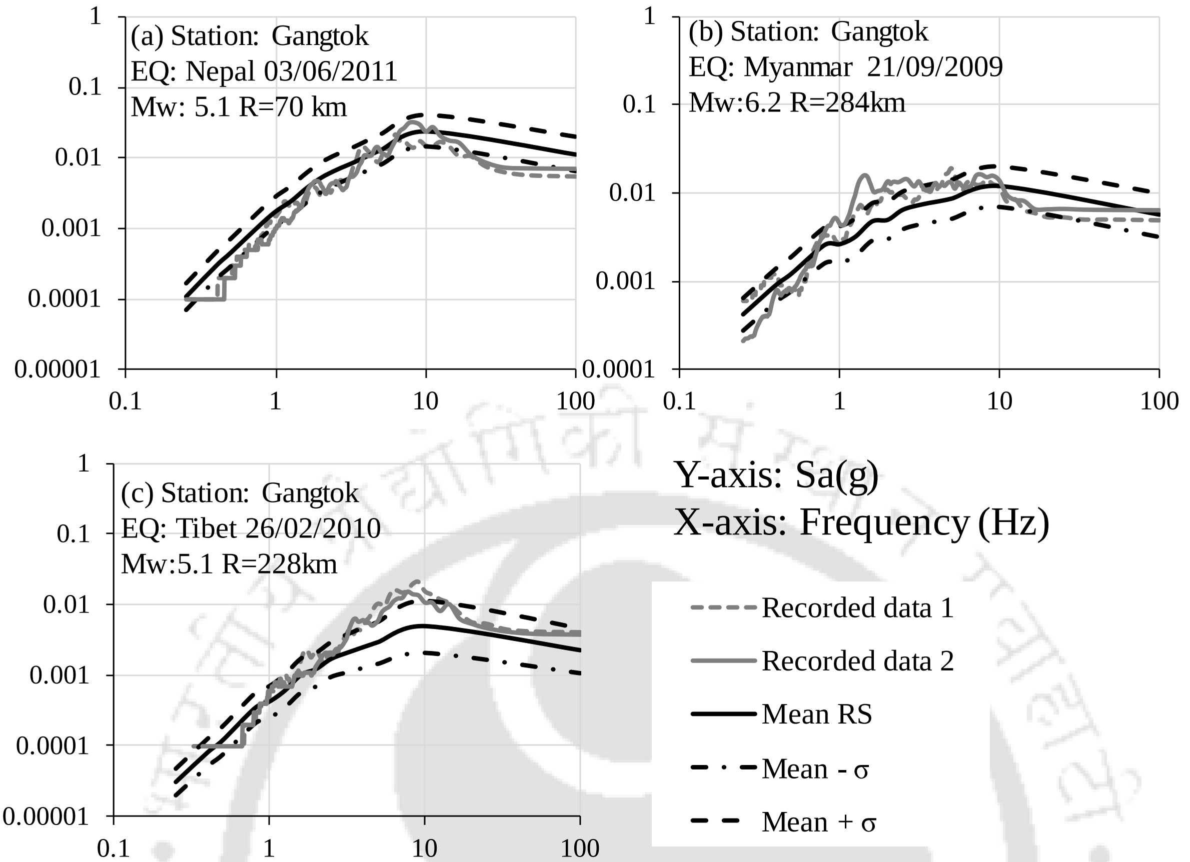

Figure 7.6 Comparison of response spectra from proposed GMPE and site coefficients for SC A with the recorded ground motions on Gangtok recording stations

Ground motions at the PESMOS stations Tejpur, Diphu, and Golaghat are utilised to validate the proposed GMPE and site coefficients for SC C conditions. For this, response spectra from the ground motions at Tejpur station based on ground motions recorded during 11 August 2009 Myanmar EQ (5.5Mw), 19 August 2012 Tejpur EQ (5.2Mw), and 02 October 2012 NE India EQ (4.9Mw) are compared with the response spectra obtained from the GMPE and site coefficients for SC C condition (Figure 7.7 a, b, c). Response spectra from the ground motions at Diphu station recorded during 03 September 2009 Myanmar EQ (5.9Mw) and 24 February 2009 NE India EQ (5.2Mw) are compared with the response spectra obtained from the GMPE and site coefficients for SC C (Figure 7.7 d, e). Response spectra from the ground motions at Golaghat station recorded during 01 July 2012 Nagaland EQ (5.6Mw), 11 May 2012 NE India (5.4Mw) and 09 January 2013 Myanmar EQ (5.8Mw) are compared with the response spectra obtained from the GMPE and site coefficients for SC C (Figure 7.7 f, g). It is observed in Figure 7.7 that the response spectra based on the proposed GMPE are very close to the response spectra as per recorded ground motions.

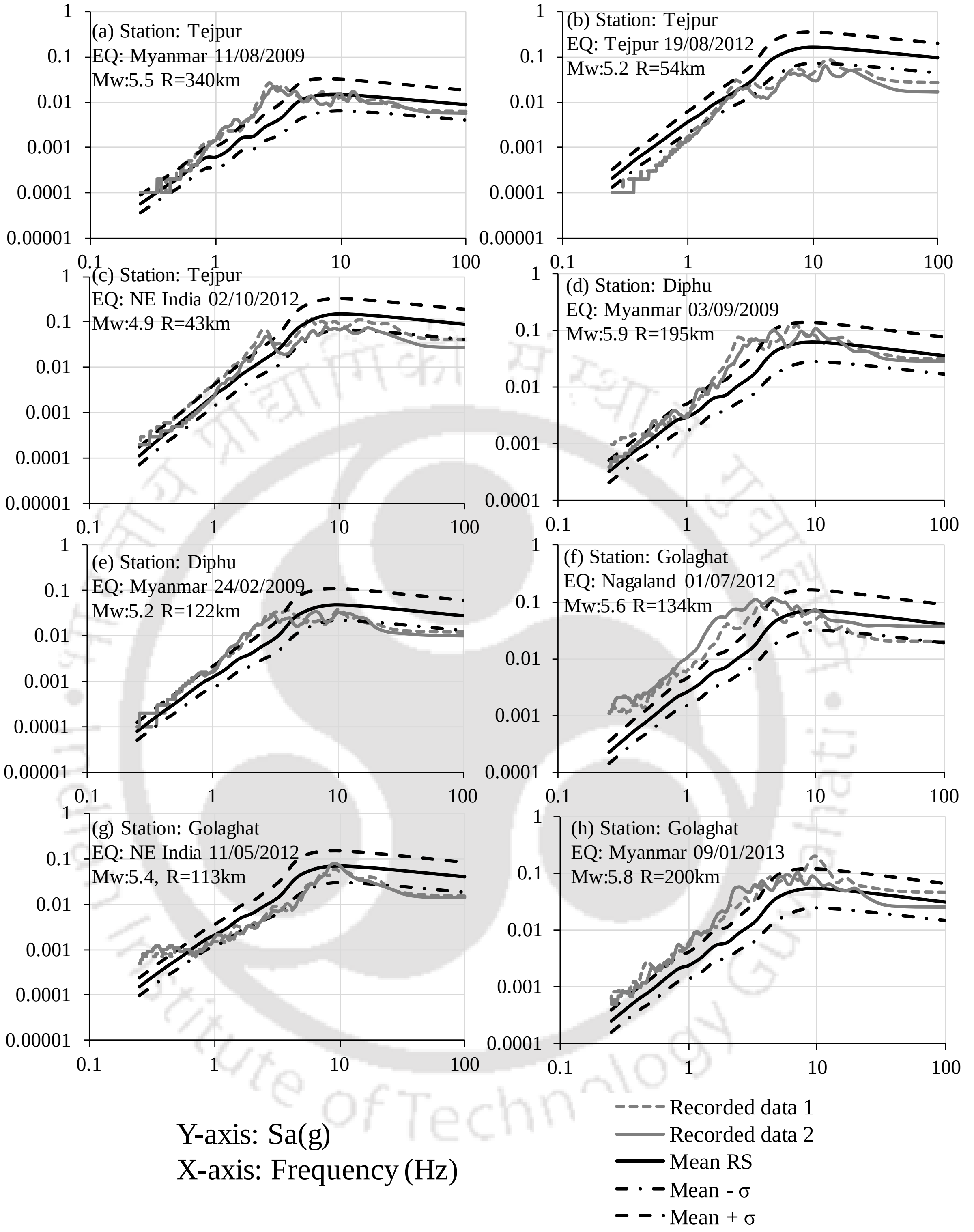

Figure 7.7 Comparison of response spectra from proposed GMPE and site coefficients for SC C with the recorded ground motions on Tejpur, Diphu and Golaghat recording stations

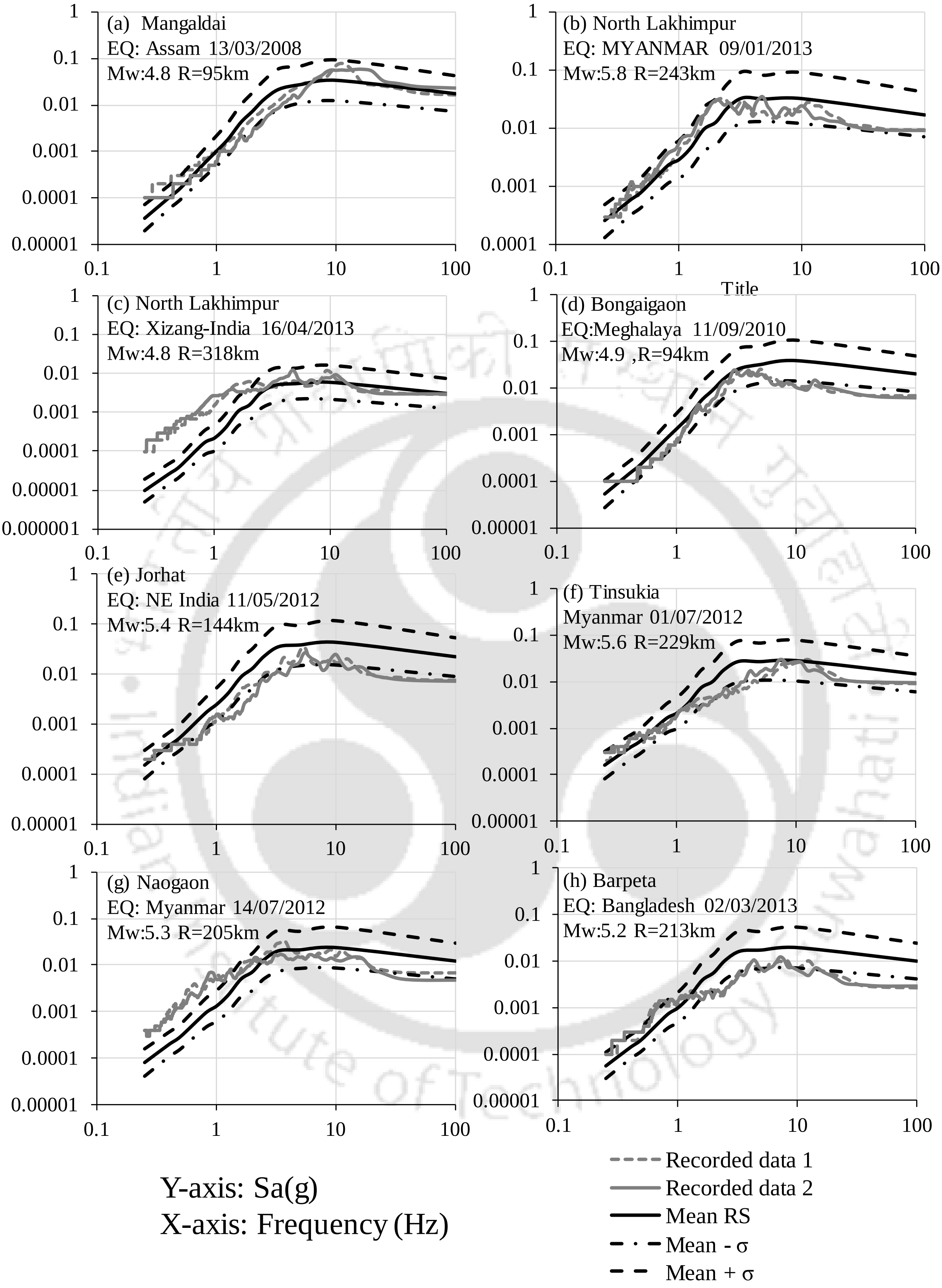

Figure 7.8 Comparison of response spectra from proposed GMPE and site coefficients for SC D with the recorded ground motions on Mangaldai, North Lakhimpur, Bongaigaon, Jorhat, Tinsukia, Naogaon, and Barpeta recording stations

Similarly, Ground motion recorded at Mangaldai, N Lakhimpur, Bongaigaon, Jorhat, Tinsukia, Naogaon and Barpeta stations are used for the validation of proposed GMPE for SC D. For this, response spectra from recorded ground motions at Mangaldai station during 13 March 2008 Assam EQ (4.8Mw) is compared with the response spectra obtained from the proposed GMPE and considering site coefficients for SC D condition (Figure 7.8 a). Further recorded response spectra from ground motions at N Lakhimpur station during 9 January 2013 Myanmar EQ (5.8Mw), and 16 April 2013 Xizang-India EQ (4.8Mw) are compared with the response spectra obtained from the proposed GMPE and considering site coefficients for SC D condition (Figure 7.8 b, c). Similarly, response spectra from ground motions at Bongaigaon station during 11 September 2010 Meghalaya EQ (4.9Mw), Jorhat station during 11 May 2012 NE India EQ (5.4Mw), Tinsukia station during 1 July 2012 Mynmar EQ (5.6Mw), Naogaon station during 14 July 2012 Myanmar EQ (5.3) and Barpeta station during 2 March 2013 Bangladesh EQ (5.2Mw) are compared with the response spectra obtained from the proposed GMPE and considering site coefficients for SC D condition (Figure 7.8 d, e, f, g, h). It can be observed from Figure 7.8 that the two sets of response spectra are matching very wells, clearly indicating that the proposed GMPE is able to predict response spectra closely for SC D up to 100Hz.

Based on the response spectra comparison shown in Figure 7.5, Figure 7.6, Figure 7.7, and Figure 7.8, it can be concluded that the response spectra as per the proposed GMPE and site coefficients match very well with recorded ground motion-based response spectra for SCs A, B, C, and D respectively. It should be noted here that the recording stations used for the validation in this section are located at different parts of NE India. Close matching of response spectra from these stations clearly states that the proposed GMPE can predict response spectra for SC A, B, C and D across NE India close to the actual record.

7.8 Summary

In the present Chapter, based on 6000 synthetic ground motion datasets, a GMPE is developed for NE India. These ground motions are developed for the bedrock level, considering regional parameters for NE India. While simulating ground motions, possible variations in ground motion model parameters have been considered using the Latin Hypercube sampling technique. To utilise the proposed GMPE for different SCs, site coefficients are obtained by performing 54, 000 site response analyses on random soil profiles for each of the five SCs. The random profiles are developed based on soil profiling information from Sikkim, Meghalaya and Guwahati city. As a part of the validation, comparisons of response spectra based on recorded ground motion with response spectra as per proposed GMPE are made for SC A, B, C and D each. Based on the comparison, it is observed that the spectral accelerations as per the proposed GMPE and proposed site coefficients match closely with the recorded ground motion-based response spectra. Thus, the present GMPE and the site coefficients can be utilised to estimate the ground motion level in NE India for different site conditions.

Chapter 8. Conclusions

8.1 General

The NE India region is well-known for its high seismic activity, with numerous catastrophic EQs in the past. The presence of numerous active faults in the region is primarily responsible for the high level of seismic activity. Thus, the region's seismic hazard levels must be thoroughly assessed to prepare for future EQs appropriately. In this study, an attempt is made to estimate the seismic hazard level for NE India. For this purpose, the present work begins with a seismic source zonation and then performs PSHA for the NE India region. The outcome of the study reveals differing hazard levels across different areas and sheds light on how different seismic sources influence these hazard levels. In addition, the present study intends to develop a region-specific GMPE that can be used for different SCs. The important conclusions of this study are discussed in detail in the following subsections.

8.1.1 Seismic source zonation for NE India

1. The hierarchical clustering technique is used to cluster past EQs based on their latitudes and longitudes, resulting in 12 seismic source zones. This seismic source ensures the uniform distribution of seismicity throughout each seismic source zone.
2. The spatial distribution of seismicity parameters is determined based on LSM and MLM. By utilising this spatial distribution of seismicity parameters, seismicity-based zonation is carried out.
3. The seismicity-based zonation is combined with the EQ location-based zonation to obtain the final seismic source zonation of NE India. This process ensures a uniform distribution of past EQs and seismicity parameters, dividing the entire NE India into 11 source zones.

8.1.2 Probabilistic seismic hazard analysis for NE India

1. The present study conducts PSHAs for NE India using both traditional PSHA methodology and Kernel density-based methodology.
2. Results from traditional PSHA methodology show that mean peak ground acceleration (PGA) values for DBE range from 0.23g to 0.51g, whereas for the MCE, the range falls between 0.40g and 0.85g.
3. The eastern region of Manipur and the southern region of Nagaland exhibit the most significant seismic hazard level due to the substantial occurrence of prior EQs and the size of seismic events in these areas.
4. Sikkim and certain Arunachal Pradesh and Assam regions have the least seismic hazard level.
5. The deaggregation of seismic hazards for eight important cities in NE India reveals the contribution of different magnitudes and distances to the seismic hazard level. The deaggregation results help understand the contribution of seismic sources around cities.
6. Strong EQs mainly control seismic hazards in NE India, both for MCE and DBE.
7. Itanagar receives most contributions from major EQs occurring at a larger distance, while other cities have significant contributions from larger EQs happening at close distances.
8. Great EQ contributions are significant in Guwahati, Gangtok, Shillong, and Itanagar.
9. Itanagar receives most contributions from major EQs occurring at a larger distance.
10. For all the cities, the contributions from larger magnitudes are more in the case of MCE than DBE. This is primarily due to the fact that MCE has a larger return period in comparison to DBE.

8.1.3 Seismic Hazard Analysis (SHA) of NE India by utilising Kernel Density method to deal with the uncertainty in location and magnitude of past EQs

1. Kernel-based PSHA is used to incorporate the non-uniform distribution of seismicity along with the uncertainty of magnitude and location of paleo EQs.
2. The results obtained from Kernel-based PSHA suggest that the overall seismic activity is more for Indo-Burma Ranges, Sikkim Himalaya, Arunachal Himalaya, Shillong Plateau and Assam Valley Region.
3. It is observed from the PSHA result that the mean PGA for the MCE and BDE scenario of NE India varies from 0.45 to 0.85g and 0.24 to 0.57g, respectively.

4 The location-wise difference in PGA value is observed in the case of the isotropic and anisotropic Kernel. For some locations near faults, PGA values obtained in the case of the Anisotropic Kernel are more than those obtained in the case of the Isotropic Kernel.

5 It is found that in the case of the catalogue, where EQs complete with respect to magnitude and time are present, the PGAs obtained are more than that obtained by using the catalogue where all EQs with magnitude more than 7Mw were present.

8.1.4 Development of GMPE for NE India, applicable to different site classes

1 The present study also developed a GMPE for NE India for different SCs based on regional parameters.

2 The developed GMPE is based on synthetic ground motions generated for the bedrock condition using the EXSIM program.

3 The present work also estimated the site coefficients, which can be used to obtain the ground motion level at the surface of different SCs.

4 The comparison between proposed GMPE and actual regional ground motion response spectra shows that the proposed GMPE is able to predict response spectra for SCs A, B, C and D across NE India close to the actual record, giving confidence about the outcome of the GMPE.

5 Developed GMPE can be used for different SCs for an extensive range of magnitude (4.0-9.0Mw) and distances (up to 400km).

It is to be mentioned here that the developed GMPE has not been used for the PSHA in the present thesis. This exceeds the scope of the present thesis task. However, in the future, this GMPE can be used to determine seismic hazard levels for NE India for different site classes, and the results can also be compared with the output obtained from the other GMPEs.

8.2 Major contributions of the present study

1 The present study developed different source zones for the NE India region based on past seismic activity, which can be used for Seismic Hazard studies.

2 The Seismic hazard map and the deaggregation results obtained for different locations can be utilised for infrastructure development and disaster preparedness for the region.

3 Paleo EQs and their size and location uncertainties are utilised in the present study. The results can add valuable input in understanding the region's seismic hazard level.

4 A new GMPE is developed, which can be utilised for different SCs in the region.

8.3 Limitations of the present study

1. The present source zonation work did not consider the uncertainty in the EQs' location and magnitude, which may influence the outcome.
2. Present PSHA did not consider the topography effect that may influence the final hazard values.
3. The Kernel density base PSHA utilised only 2-D Kernel for the analysis.
4. The present study does not incorporate the seismic gap present in NE India.
5. The GMPE developed in the study does not incorporate the V_{s30} in the equation. Thus, for particular Vs30 values, the ground motion level cannot be determined from the present GMPE.

8.4 Recommendation for Future Scope

1. The effect of the topography can be incorporated into the SHA.
2. 3-D Kernel density can also be utilised in SHA.
3. With the help of the present SHA results, studies related to seismic vulnerability can be performed.
4. Vs30-based Next Generation GMPE can be developed.

References

Abrahamson NA, Litehiser JJ (1989) Attenuation of vertical peak acceleration. Bull - Seismol Soc Am 79:549–580

Adams J (1990) Paleoseismicity of the Cascadia Subduction Zone: Evidence from Turbidites off the Oregon-Washington Margin. Tectonics 9:569–583

Aki K, Richards P (1980) Quantitative Seismology, Theory and Methods

Alam E, Dominey-Howes D (2016) Earthquakes in the north-eastern coastline of Indian Ocean with a particular focus on the Bay of Bengal–a synthesis and review. Nat Hazards 81:2031–2102

Algermissen ST, Perkins DM, Thenhaus PC, et al (1982) Probabilistic estimates of maximum acceleration and velocity in rock in the contiguous United States. US Geological Survey,

Álvarez-Gómez JA (2019) FMC—Earthquake focal mechanisms data management, cluster and classification. SoftwareX 9:299–307. https://doi.org/10.1016/j.softx.2019.03.008

Ambraseys N, Bilham R (2000) A note on the Kangra M s = 7 . 8 earthquake of 4 April 1905. Curr Sci 79:45–50

Ambraseys N, Bilham R (2003) Reevaluated intensities for the great Assam Earthquake of 12 June 1897, Shillong, India. Bull Seismol Soc Am 93:655–673. https://doi.org/10.1785/0120020093

Ambraseys NN, Douglas J (2004) Magnitude calibration of north Indian earthquakes. Geophys J Int 159:165–206. https://doi.org/10.1111/j.1365-246X.2004.02323.x

Anbazhagan P, Krishna AM (2015) Forensic investigation of earthquake induced failures during Sikkim 2011 earthquake, India. 15th Asian Reg Conf Soil Mech Geotech Eng ARC 2015 New Innov Sustain 201–206. https://doi.org/10.3208/jgssp.IND-19

Anbazhagan P, Kumar A, Sitharam TG (2013) Ground motion prediction equation considering combined dataset of recorded and simulated ground motions. Soil Dyn Earthq Eng 53:92–108. https://doi.org/10.1016/j.soildyn.2013.06.003

Angelier J, Baruah S (2009) Seismotectonics in Northeast India : a stress analysis of focal mechanism solutions of earthquakes and its kinematic implications. Geophys J Int 178:303–326. https://doi.org/10.1111/j.1365-246X.2009.04107.x

Ansari A, Firuzi E, Etemadsaeed L (2015) Delineation of seismic sources in probabilistic seismic-hazard analysis using fuzzy cluster analysis and Monte Carlo simulation. Bull Seismol Soc Am 105:2174–2191. https://doi.org/10.1785/0120140256

Atkinson GM, Boore DM (2006a) Earthquake ground-motion prediction equations for eastern North America. Bull Seismol Soc Am 96:2181–2205

Atkinson GM, Boore DM (2006b) Earthquake ground-motion prediction equations for eastern North America. Bull Seismol Soc Am 96:2181–2205. https://doi.org/10.1785/0120050245

Atkinson GM, Boore DM (2003) Empirical Ground-Motion Relations for Subduction-Zone Earthquakes and Their Application to Cascadia and Other Regions. Bull Seismol Soc Am 93:1703–1729

Aung HH (2020) The 2020 M 5.9 Falam Earthquake the Subduction-Induced Strike-Slip Earthquake, West Myanmar. Curr J Appl Sci Technol 39:59–67. https://doi.org/10.9734/cjast/2020/v39i1530717

Bahuguna A, Sil A (2018) Comprehensive Seismicity, Seismic Sources and Seismic Hazard Assessment of Assam, North East Comprehensive Seismicity, Seismic Sources and Seismic. J Earthq Eng 00:1–44. https://doi.org/10.1080/13632469.2018.1453405

Bajaj K, Anbazhagan P (2019) Regional stochastic GMPE with available recorded data for active region – Application to the Himalayan region. Soil Dyn Earthq Eng 126:105825. https://doi.org/10.1016/j.soildyn.2019.105825

Bapat A, Kulkarni RC, Guha SK (1983) Catalogue of Earthquakes in India and Neighborhood from Historical Period up to 1979, Indian Society of Earthquake Technology, Roorkee, India. BARUAH

Baro O, Kumar A (2015) A review on the tectonic setting and seismic activity of the Shillong plateau in the light of past studies. Disaster Adv 8:34–45

Baro O, Kumar A (2017a) Seismic source characterization for the Shillong Plateau in Northeast India. J Seismol 21:1229–1249. https://doi.org/10.1007/s10950-017-9664-2

Baro O, Kumar A (2017b) Seismic source characterization for the Shillong Plateau in Northeast India. J Seismol. https://doi.org/10.1007/s10950-017-9664-2

Baro O, Kumar A, Ismail-zadeh A (2018a) Seismic hazard assessment of the Shillong Plateau, India. Geomatics, Nat Hazards Risk 9:841–861. https://doi.org/10.1080/19475705.2018.1494043

Baro O, Kumar A, Ismail-Zadeh A (2020) Seismic hazard assessment of the Shillong Plateau using a probabilistic approach. Geomatics, Nat Hazards Risk 11:2210–2238. https://doi.org/10.1080/19475705.2020.1833989

Baro O, Kumar A, Ismail ZA (2018b) Seismic hazard of the Shillong Plateau, Geomatics. Nat Hazard Risk

Baruah S, Baruah S, Gogoi NK, et al (2009) Ground motion parameters of Shillong plateau: One of the most seismically active zones of northeastern India. Earthq Sci 22:283–291. https://doi.org/10.1007/s11589-009-0285-2

Baruah S, Boruah GK, Sharma S, et al (2020) Seismic vulnerability assessment of earthquake-prone mega-city Shillong, India using geophysical mapping and remote sensing. Georisk 14:112–127. https://doi.org/10.1080/17499518.2019.1598560

Beauval C, Hainzl S, Scherbaum F (2006) Probabilistic seismic hazard estimation in low-seismicity regions considering non-Poissonian seismic occurrence. Geophys J Int 164:543–550. https://doi.org/10.1111/j.1365-246X.2006.02863.x

Bender B, Perkins DM (1987) SEISRISK III: a computer program for seismic hazard estimation. US Government Printing Office

Benítez HD, Flórez JF, Duque DP, et al (2013) Spatial pattern recognition of seismic events in South West Colombia. Comput Geosci 59:60–77. https://doi.org/10.1016/j.cageo.2013.04.028

Beresnev IA, Atkinson GM (1997) Modeling finite-fault radiation from the ωn spectrum. Bull Seismol Soc Am 87:67–84

Berryman KR, Ota Y, Hull AG (1989) Holocene paleoseismicity in the fold and thrust belt of the Hikurangi subduction zone , eastern North Island , New Zealand. Tectonophysics 163:185–195

Besheli PR, Zare M, Umali RR, Nakhaeezadeh G (2015) Zoning Iran based on earthquake precursor importance and introducing a main zone using a data-mining process. Nat Hazards 78:821–835. https://doi.org/10.1007/s11069-015-1745-3

Bhatia SC, Kumar MR, Gupta HK (1999) A probabilistic seismic hazard map of India and adjoining regions. Ann Di Geofis 42:

Bhukosh Bhukosh. https://bhukosh.gsi.gov.in/Bhukosh/MapViewer.aspx

Bhuvan Bhuvan:Indian Geo-Platform of ISRO. https://bhuvan.nrsc.gov.in/home/index.php

Bilham R (2004) Earthquakes in India and the Himalaya : tectonics , geodesy and history. Ann Geophys 47:839–858. https://doi.org/10.4401/ag-3338

Bilham R (2001) Slow tilt reversal of the Lesser Himalaya between 1862 and 1992 at 78 deg. E, and bounds to the southeast rupture of the 1905 Kangra earthquake. Geophys J Int 144:713–728

Bilham R, England P (2001) Plateau Pop-up during the 1897 Assam earthquake. Nature 410:806–809

BIS (2016) IS-1893 2016 part 1: Criteria for Earthquake Resistant Design of Structures sixth revision

Biswas R, Baruah S (2017) Assessing site response through H/V ratio in a seismic prone area of northeast India. arXiv Prepr arXiv170502952

Boore D, Atkinson G (1987) Stochastic prediction of ground motion and spectral response parameters at hard-rock sites in eastern North America. Bull Seismol Soc Am 77:440–467

Boore DM (1983) Stochastic simulation of high-frequency ground motions based on seismological models of the radiated spectra. Bull Seismol Soc Am 73:1865–1894

Boore DM, Atkinson GM (2008) Ground-Motion Prediction Equations for the Average Horizontal Component of PGA , PGV , and 5 % -Damped PSA at Spectral Periods between

0 . 01 s and 10 . 0 s. Earthq Spectra 24:99–138. https://doi.org/10.1193/1.2830434

Borah N, Kumar A, Dhanotiya R (2021) Seismic Source zonation for NE India on the basis of past EQs and spatial distribution of seismicity parameters. J Seismol

Brune J (1970) Tectonic stress and the spectra of seismic shear waves from earthquakes. J Geophys Res 75:4997–5009

BSSC (2003) NEHRP recommended provisions for seismic regulations for new buildings and other structures (FEMA 450). Washington DC.

Campbell K. W. (1981) Near-Source Attenuation of Peak Horizontal Acceleration

Campbell KW (2003) Prediction of Strong Ground Motion Using the Hybrid Empirical Method and Its Use in the Development of Ground-Motion (Attenuation) Relations in Eastern North America. Bull Seismol Soc Am 93:1012–1033

Chan CH (2016) Importance of three-dimensional grids and time-dependent factors for applications of earthquake forecasting models to subduction environments. Nat Hazards Earth Syst Sci 16:2177–2187. https://doi.org/10.5194/nhess-16-2177-2016

Chan CH, Sørensen MB, Stromeyer D, et al (2010) Forecasting italian seismicity through a spatio-temporal physical model: Importance of considering time-dependency and reliability of the forecast. Ann Geophys 53:129–140. https://doi.org/10.4401/ag-4761

Chandra U (1992) Seismotectonics of Himalaya. Curr Sci 62:40–71

Chen B (2020) Stress-induced trend: the clustering feature of coal mine disasters and earthquakes in China. Int J Coal Sci Technol 7:676–692. https://doi.org/10.1007/s40789-020-00334-z

Chopra S, Sharma J, Sutar A, Bansal BK (2014) Estimation of source parameters of M w 6.9 Sikkim Earthquake and modeling of ground motions to determine causative fault. Pure Appl Geophys 171:1311–1328

Cornell CA (1968) Engineering seismic risk analysis. Bull Seismol Soc Am 58:1583–1606

Cramer CH (2001) A seismic hazard uncertainty analysis for the New Madrid seismic zone. Eng Geol 62:251–266. https://doi.org/10.1016/S0013-7952(01)00064-3

Cramer CH, Wheeler RL, Mueller CS (2002) Uncertainty analysis for seismic hazard in the southern Illinois basin. Seismol Res Lett 73:792–805. https://doi.org/10.1785/gssrl.73.5.792

D'Amico M, Lanzano G, Santulin M, et al (2018) Hybrid GMPEs for Region-Specific PSHA in Southern Italy. Geosciences 8:. https://doi.org/10.3390/geosciences8060217

Das R (2013) Probabilistic seismic hazard assessment of Northeast India region. IIT Roorkee

Das R, Sharma ML, Wason HR (2016) Probabilistic Seismic Hazard Assessment for Northeast India Region. Pure Appl Geophys 173:2653–2670. https://doi.org/10.1007/s00024-016-1333-9

Dasgupta S, Nandy D (1982) Seismicity and Tectonics of Meghalaya plateau, northeastern India. In: Proc. VII Syrup. on Earthquake Engineering, Roorkee. Roorkee, pp 19–24

Davis SD, Frohlich C (1991) Single-Link Cluster Analysis of Earthquake Aftershocks' Decay Laws and Regional Variation. J Geophys Res 96:6335–6350

Dhanya J, Raghukanth STG (2020) Neural network-based hybrid ground motion prediction equations for Western Himalayas and North-Eastern India. Acta Geophys 68:303–324. https://doi.org/10.1007/s11600-019-00395-y

Doroudian M, Vucetic M (1995) A direct simple shear device for measuring small-strain behavior. Geotech Test J 18:69–85

Drouet S, Cotton F (2015) Regional Stochastic GMPEs in Low-Seismicity Areas: Scaling and Aleatory Variability Analysis—Application to the French Alps. Bull Seismol Soc Am 105:. https://doi.org/10.1785/0120140240

Dziewonski AM, Chou T, Woodhouse JH (1981) Determination of earthquake source parameters from waveform data for studies of global and regional seismicity. J Geophys Res Solid Earth 86:2825–2852

Edwards B, Fäh D (2013) A Stochastic Ground-Motion Model for Switzerland. Bull Seismol Soc Am 103:78–98. https://doi.org/10.1785/0120110331

Ekström G, Nettles M, Dziewoński AM (2012) The global CMT project 2004–2010: Centroid-moment tensors for 13,017 earthquakes. Phys Earth Planet Inter 200:1–9

EM (1999) Response Spectra and Seismic Analysis for Concrete Hydraulic Structures. DEPARTMENT OF THE ARMY U.S. Army Corps of Engineers, Washington, DC

England P, Bilham R (2015) The Shillong Plateau and the great 1897 Assam earthquake. Tectonics 34:1792–1812. https://doi.org/10.1002/2015TC003902

Esteva L (1969) Seismic Prediction-A baysen approach.pdf. In: Fourth World Conf. on Earthquake Engineering. Chile, pp 13–18

Esteva L (1970) Seismic risk and seismic design decisions. Massachusetts Inst. of Tech., Cambridge. Univ. of Mexico, Mexico City

Fahmi KJ, Ayar BS, Hassan IA (1988) The Takab seismic gap-forecasting a major earthquake occureence off northeastern Iraq. J Geodyn 10:43–72. https://doi.org/10.1016/0264-3707(88)90005-1

Fraley C, Raftery A (1998) How many clusters? Which clustering method? Answers via model-based cluster analysis. Tech Rep no 329, Deparment Stastics Univ og Washingt

Frohlich C, Davis SD (1990) Single-link cluster analysis as a method to evaluate spatial and temporal properties of earthquake catalogues. Geophys J Int 100:19–32

Fukushima Y, Tanaka T (1990) A new attenuation relation for peak horizontal acceleration of strong earthquake ground motion in Japan. Bull - Seismol Soc Am 80:757–783

Gahalaut VK, Kundu B (2016) The 4 January 2016 Manipur earthquake in the Indo-Burmese wedge, an intra-slab event. Geomatics, Nat Hazards Risk 7:1506–1512. https://doi.org/10.1080/19475705.2016.1179686

Gardner J k., Knopoff L (1974) Is The Sequence of Earthquakes in Southern California, with Aftershocks Removed, Poissonian? Bull Seismol Soc Am 64:1363–1367. https://doi.org/10.1785/0120160029

Georgoulas G, Konstantaras A, Katsifarakis E, et al (2013) "Seismic-mass" density-based algorithm for spatio-temporal clustering. Expert Syst Appl 40:4183–4189. https://doi.org/10.1016/j.eswa.2013.01.028

Ghione F, Poggi V, Lindholm C (2021) A hybrid probabilistic seismic hazard model for Northeast India and Bhutan combining distributed seismicity and finite faults. Phys Chem Earth 123:103029. https://doi.org/10.1016/j.pce.2021.103029

Görgün E (2013) Analysis of the b-values before and after the 23 October 2011 Mw 7.2 Van-Erciş, Turkey earthquake. Tectonophysics 603:213–221. https://doi.org/10.1016/j.tecto.2013.05.030

Görgün E, Zang A, Bohnhoff M, et al (2009) Analysis of Izmit aftershocks 25 days before the November 12th 1999 Düzce earthquake, Turkey. Tectonophysics 474:507–515. https://doi.org/10.1016/j.tecto.2009.04.027

Goswami HC, Sarmah SK (1982) Probabilistic earthquake expectancy in the northeast Indian region. Bull Seismol Soc Am 72:999–1009

Grünthal G, Stromeyer D, Bosse C, et al (2018) The probabilistic seismic hazard assessment of Germany—version 2016, considering the range of epistemic uncertainties and aleatory variability. Bull Earthq Eng 16:4339–4395. https://doi.org/10.1007/s10518-018-0315-y

GSI, NRSC (2012) National geomorphological and lineament mapping on 1: 50,000 Scale. Nat Resour Census Proj Natl Remote Sens Centre, ISRO, Hyderabad

Guha-Sapir D, Vos F (2011) Human Casualties in Earthquakes Advances in Natural and Technological Hazards Research. In: In Human casualties in earthquakes. Springer, Dordrecht., pp 13–24

Gupta HK, Rajendran K, Singh HN (1986) Seismicity of Northeast India region: PART I: The database. J Geol Soc India 28:345–365

Gupta HK, Rastogi BK, Mohan I, et al (1998) An investigation into the Latur earthquake of September 29 , 1993 in southern India. Tectonophysics 287:299–318. https://doi.org/10.1016/S0040-1951(98)80075-9

Gupta ID (2010) Response spectral attenuation relations for in-slab earthquakes in Indo-Burmese subduction zone. Soil Dyn Earthq Eng 30:368–377. https://doi.org/10.1016/j.soildyn.2009.12.009

Gupta ID (2006) Delineation of probable seismic sources in India and neighbourhood by a comprehensive analysis of seismotectonic characteristics of the region. Soil Dyn Earthq Eng 26:766–790. https://doi.org/10.1016/j.soildyn.2005.12.007

Gupta ID (2002) The state of the art in seismic hazard analysis. ISET J Earthq Technol 39:311–346

Gutenberg B, Richter CF (1944) Frequency of Earthquakes in California. Bull Seismol Soc Am 185–188

Guttorp P (1987) On Least-Squares Estimation of b Values. Bull Seismol Soc Am 77:2115–2124

Halder S, Mondal JK, A. K (2019) Empirical correlations for strain dependent linear 1D ground response analysis for Guwahati, India based on detailed analyses. In: 7ICEGE. Roma

Haloi J, Sil A (2015) Seismic site classification of a bridge site over River Barak on Silchar bypass road. Int J Adv Earth Sci Eng 4:275–282

Hanks TC, McGuire RK (1981) The character of high-frequency strong ground motion. Bull Seismol Soc Am 71:2071–2095. https://doi.org/10.1785/BSSA0710062071

Harinarayan, N.H., Kumar A (2020) Ground motion prediction equation for north India, applicable for different site classes. Soil Dyn Earthq Eng 139:106425. https://doi.org/10.1016/j.soildyn.2020.106425

Hartzell S. (1978) Earthquake aftershocks as Green's functions. Geophys Res Lett 5:1–4

Hashemi M, Karimi HA (2016) Seismic source modeling by clustering earthquakes and predicting earthquake magnitudes. Lect Notes Inst Comput Sci Soc Telecommun Eng LNICST 166:468–478. https://doi.org/10.1007/978-3-319-33681-7_39

Hashemi SN, Mehdizadeh R (2015) Application of hierarchical clustering technique for numerical tectonic regionalization of the Zagros region (Iran). Earth Sci Informatics 8:367–380. https://doi.org/10.1007/s12145-014-0163-5

Hazarika P, Ravi Kumar M (2012) Seismicity and source parameters of moderate earthquakes in Sikkim Himalaya. Nat hazards 62:937–952

Heath DC, Wald DJ, Worden CB, et al (2020) A global hybrid VS 30 map with a topographic slope–based default and regional map insets. Earthq Spectra 36:1570–1584

Heaton TH, Hartzell SH (1989) Estimation of strong ground motions from hypothetical earthquakes on the Cascadia subduction zone, Pacific Northwest. Pure Appl Geophys PAGEOPH 129:131–201. https://doi.org/10.1007/BF00874626

Hetényi G, Roux-mallouf R Le, Berthet T, et al (2016) Joint approach combining damage and paleoseismology observations constrains the 1714 A.D. Bhutan earthquake at magnitude 8 ± 0.5. 695–702. https://doi.org/10.1002/2016GL071033.Received

Hornblow S, Quigley M, Nicol A, et al (2014) Paleoseismology of the 2010 Mw 7.1 Darfield (Canterbury) earthquake source, Greendale Fault, New Zealand. Tectonophysics 637:178–190. https://doi.org/10.1016/j.tecto.2014.10.004

Hudyma M, Potvin YH (2010) An engineering approach to seismic risk management in hardrock mines. Rock Mech Rock Eng 43:891–906. https://doi.org/10.1007/s00603-009-

0070-0

Hudyma MR (2008) Analysis and interpretation of clusters of seismic events in mines. PhD Thesis. University of Western Australia

Husen S, Hardebeck J (2010) Understanding Seismicity Catalogs and their Problems. Community Online Resour Stat Seism Anal 1–35. https://doi.org/10.5078/corssa-55815573.Available

IMD Indian MeteorologicalDepartment. http://www.imd.gov.in/pages/earthquake_prelim.php. Accessed 1 May 2019

Irikura K (1983) Semi-empirical estimation of strong ground motions during large earthquakes. Bull Disaster Prev Res Institute 33:63–104

ISC Event CatalogueSearch. http://www.isc.ac.uk/iscbulletin/search/catalogue/. Accessed 1 May 2019

Ishihara K (1985) Proc. of 11th ICSMFE. In: Stability of natural deposits during earthquakes. pp 321–376

Iyengar RN, Ghosh S (2004) Microzonation of earthquake hazard in Greater Delhi area. Curr Sci 87:1193–1202

Iyengar RN, Sharma D, Siddiqui JM (1999) Earthquake history of India in medieval times. Indian J Hist Sci 34:181–237

Jacoby GC (1997) Application of tree ring analysis to paleoseismology. Rev Geophys 35:109–124

Jain KS, A.D R, Arlekar JN, Basu PC (2000) Empirical Attenuation Relationships for the Himalayan Earthquakes Based on Indian Strong Motion Data. In: Proceedings of the Sixth International Conference on Seismic Zonation.

Jain SK (1998) Indian Earthquakes : An Overview. Indian Concr J 72:

Jain SK, Pathak S (2012) Intensity Based Casualty Models : Case Study Of Bhuj and Latur Earthquake in India. In: 15th World Conference on Earthquake Engineering

Jana P, Mohapatra S, Chowdhuri SI, Chandrashekar S (2012) A report on seismic hazard assessment of Jorhat Urban Agglomeration, Assam. Geological Survey of India. Code No Sei/Ner/Rhq/2010/025.

Jarvis A, Reuter HI, Nelson A, Guevara E (2008) Hole-filled seamless SRTM data V4, International Centre for Tropical Agriculture (CIAT). http://srtm.csi.cgiar.org

Jayangondaperumal R, Wesnousky SG, Choudhuri BK (2011) Near-Surface Expression of Early to Late Holocene Displacement along the Northeastern Himalayan Frontal Thrust at Marbang Korong Creek, Arunachal Pradesh, India. Bull Seism Soc Am 101:3060–3064. https://doi.org/10.1785/0120110051

Joshi DD, John B, Kandpal GC, Pande P (2009) Paleoliquefaction Features from the Himalayan

Frontal Belt , India and its Implications to the Status of “Central Seismic Gap.” J South Asia Disaster Stud 2:139–154

Joyner WB, Boore DM (2011) Peak horizontal acceleration and velocity from strong- motion records including records from the 1979 imperial valley, california, earthquake. Bull Seismol Soc Am 71:2011–2038

Kalafat D, Görgün E (2019) Source characteristics and b-values of the Tuz Gölü Fault Zone in Central Anatolia, Turkey. J Asian Earth Sci 179:337–349. https://doi.org/10.1016/j.jseaes.2019.05.005

Kanamori H (1977) The energy release in great earthquakes. J Geophys Res 82:2981–2987

Kanamori H, Anderson DL (1975) Theoretical Basis of Some Empirical Relations in Seismology. Bull Seismol Soc Am 65:1073–1095

Kayal JR (1998) Seismotectonics of Northeast India: A Review. Mem Geol Soc India 55–68

Kayal JR (2008) Microearthquake Seismology and Seismotectonics of South Asia. Springer, India

Kayal JR, Arefiev SS, Baruah S, et al (2012) Large and great earthquakes in the Shillong plateau-Assam valley area of Northeast India Region: Pop-up and transverse tectonics. Tectonophysics 532–535:186–192. https://doi.org/10.1016/j.tecto.2012.02.007

Kayal JR, Arefiev SS, Baruah S, et al (2010) The 2009 Bhutan and Assam felt earthquakes (Mw 6.3 and 5.1) at the Kopili fault in the northeast Himalaya region. Geomatics, Nat Hazards Risk 1:273–281. https://doi.org/10.1080/19475705.2010.486561

Khattri K, Wyss M (1978) Precursory variation of seismicity rate in the Assam area, India. Geology 6:685–688. https://doi.org/10.1130/0091-7613(1978)6<685:PVOSRI>2.0.CO;2

Khattri K, Wyss M, Gaur VK, et al (1983) Local seismic activity in the region of the Assam gap, northeast India. Bull Seismol Soc Am 73:459–469

Khattri KN (1987) Great earthquakes, seismicity gaps and potential for earthquakes along the Himalayan plate boundary. Tectonophysics 38:79–92

Khattri KN, Rogers AM, Perkins DM, Algermissen ST (1984) A seismic hazard map of india and adjacent areas. Tectonophysics 108:93–134

Kijko A (2004) Estimation of the maximum earthquake magnitude, mmax. Pure Appl Geophys 161:1655–1681. https://doi.org/10.1007/s00024-004-2531-4

Kijko A, Funk CW (1996) Space-time interaction amongst clusters of mining induced seismicity. Pure Appl Geophys 147:277–288. https://doi.org/10.1007/bf00877483

Kijko A, Graham G (1998) Parametric-historic procedure for probabilistic seismic hazard analysis. Part I: Estimation of maximum regional magnitude m(max). Pure Appl Geophys 152:413–442. https://doi.org/10.1007/s000240050161

Kijko A, Sellevoll M (1992) Estimation of Earthquake Hazard Parameters from Incomplete

Data Files. Part II. Incorporation of Uncertainty of Earthquake-Occurrence Model. 120–134

Kijko A, Sellevoll MA (1989) Estimation of earthquake hazard parameters from incomplete data files, part I: Utilization of extreme and complete catalogues with different threshold magnitudes. Bull Seismol Soc Am 79:645–654

Kijko A, Singh M (2011) Statistical tools for maximum possible earthquake magnitude estimation. Acta Geophys 59:674–700. https://doi.org/10.2478/s11600-011-0012-6

Knopoff L (1964) The Statistics of Earthquakes in Sourthern California. Bull Seismol Soc Am 54:1871–1873

Kolathayar S, Sitharam TG (2012) Characterization of Regional Seismic Source Zones in and around India. Seismol Res Lett 83:77–85. https://doi.org/10.1785/gssrl.83.1.77

Konstantaras AJ, Katsifarakis E, Maravelakis E, et al (2012) Intelligent Spatial-Clustering of Seismicity in the Vicinity of the Hellenic Seismic Arc. Earth Sci Res 1:1–10. https://doi.org/10.5539/esr.v1n2p1

Kramer SL (1996) Geotechnical Earthquake Engineering. Prentice Hall, Saddle River, New Jersey

Krzanowski AWJ, Lai YT (1988) A Criterion for Determining the Number of Groups in a Data Set Using Sum-of-Squares Clustering Published by : International Biometric Society Stable URL : http://www.jstor.org/stable/2531893 REFERENCES Linked references are available on JSTOR for this art. Biometrics 44:23–34. https://doi.org/10.2307/2531893

Kumar A, Baro O, Harinarayan NH (2016a) Obtaining the surface PGA from site response analyses based on globally recorded ground motions and matching with the codal values. Nat Hazards 81:543–572. https://doi.org/10.1007/s11069-015-2095-x

Kumar A, Harinarayan NH, Baro O (2017) Nonlinear soil response to ground motions during different earthquakes in Nepal, to arrive at surface response spectra. Nat Hazards 87:13–33. https://doi.org/10.1007/s11069-017-2751-4

Kumar A, Harinarayan NH, Verma V, et al (2018) Seismic Site Classification and Empirical Correlation Between Standard Penetration Test N Value and Shear Wave Velocity for Guwahati Based on Thorough Subsoil Investigation Data. Pure Appl Geophys 175:2721–2738. https://doi.org/10.1007/s00024-018-1858-1

Kumar A, Mitra S, Suresh G (2015) Seismotectonics of the eastern Himalayan and indo-burman plate boundary systems. Tectonics 34:2279–2295. https://doi.org/10.1002/2015TC003979

Kumar A, Mondal JK (2017) Newly Developed MATLAB Based Code for Equivalent Linear Site Response Analysis. Geotech Geol Eng 35:2303–2325. https://doi.org/10.1007/s10706-017-0246-4

Kumar A, Suman H (2020) Design Response Spectra and Site Coefficients for Various Seismic Site Classes of Guwahati, India, Based on Extensive Ground Response Analyses. Geotech Geol Eng 38:6255–6280. https://doi.org/10.1007/s10706-020-01434-y

Kumar D, Reddy D V, Pandey AK (2016b) Paleoseismic investigations in the Kopili Fault Zone of North East India : Evidences from liquefaction chronology. Tectonophysics 674:65–75. https://doi.org/10.1016/j.tecto.2016.02.016

Kumar S, Wesnousky SG, Jayangondaperumal R, et al (2010) Paleoseismological evidence of surface faulting along the northeastern Himalayan front, India: Timing, size, and spatial extent of great earthquakes. J Geophys Res 115:1–20. https://doi.org/10.1029/2009JB006789

Kumar S, Wesnousky SG, Rockwell TK, et al (2006) Paleoseismic evidence of great surface rupture earthquakes along the Indian Himalaya. J Geophys Res 111:1–19. https://doi.org/10.1029/2004JB003309

Kumar V, Kumar D, Chopra S (2019) Source Parameters and Scaling Relations for Moderate Size Earthquakes in North–East India Region. Pure Appl Geophys 176:45–64

Kundu B, Gahalaut VK (2013) Tectonic geodesy revealing geodynamic complexity of the Indo-Burmese arc region, North East India. Curr Sci 104:920–933

Lavé J, Yule D, Sapkota S, et al (2005) Evidence for a great medieval earthquake (~1100 A.D.) in the central Himalayas, Nepal. Science (80-) 307:1302–1305. https://doi.org/10.1126/science.1104804

Ledolter J (2013) Data mining and business analytics with R. John Wiley & Sons

Leśniak A, Isakow Z (2009) Space-time clustering of seismic events and hazard assessment in the Zabrze-Bielszowice coal mine, Poland. Int J Rock Mech Min Sci 46:918–928. https://doi.org/10.1016/j.ijrmms.2008.12.003

Malik JN, Sahoo AK, Shah AA, et al (2010) Paleoseismic evidence from trench investigation along Hajipur fault , Himalayan Frontal Thrust , NW Himalaya : Implications of the faulting pattern on landscape evolution and seismic hazard. J Struct Geol 32:350–361. https://doi.org/10.1016/j.jsg.2010.01.005

Martínez-Álvarez F, Gutiérrez-Avilés D, Morales-Esteban A, et al (2015) A novel method for seismogenic zoning based on triclustering: Application to the Iberian Peninsula. Entropy 17:5000–5021. https://doi.org/10.3390/e17075000

Maurin T, Rangin C (2009) Structure and kinematics of the Indo-Burmese Wedge: Recent and fast growth of the outer wedge. Tectonics 28:1–21. https://doi.org/10.1029/2008TC002276

McCalpin JP, Nelson AR (2009) Introduction to paleoseismology. Int Geophys 95:1–27. https://doi.org/10.1016/S0074-6142(96)80068-4

McCalpin JP, Nelson AR (1996) Paleoseismology. Academic Press, Inc

McGuire RK (1976) FORTRAN computer program for seismic risk analysis. USGS Open-File Rep 76:90

McKay MD, Beckman RJ, Conover WJ (2000) A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. Technometrics

42:55–61. https://doi.org/10.1080/00401706.2000.10485979

Mishra RL, Singh I, Pandey A, et al (2016) Paleoseismic evidence of a giant medieval earthquake in the eastern Himalaya. Geophys Res Lett 43:5707–5715

Mittal H, Kumar A, Ramhmachhuani R (2012) Indian National Strong Motion Instrumentation Network and Site Characterization of Its Stations. Int J Geosci 03:1151–1167. https://doi.org/10.4236/ijg.2012.326117

Molchan G, Kronrod T, Panza GF (1997) Multi-scale seismicity model for seismic risk. Bull Seismol Soc Am 87:1220–1229

Molchan GM (1990) Strategies in strong earthquake prediction. Phys Earth Planet Inter 61:84–98. https://doi.org/10.1016/0031-9201(90)90097-H

Molina S, Lindholm CD, Bungum H (2001) Probabilistic seismic hazard analysis: Zoning free versus zoning methodology. Boll di Geofis Teor ed Appl 42:19–39

Morales-Esteban A, Martínez-Álvarez F, Scitovski S, Scitovski R (2014) A fast partitioning algorithm using adaptive Mahalanobis clustering with application to seismic zoning. Comput Geosci 73:132–141. https://doi.org/10.1016/j.cageo.2014.09.003

Morino M, Maksud Kamal ASM, Muslim D, et al (2011) Seismic event of the Dauki Fault in 16th century confirmed by trench investigation at Gabrakhari Village, Haluaghat, Mymensingh, Bangladesh. J Asian Earth Sci 42:492–498. https://doi.org/10.1016/j.jseaes.2011.05.002

Motazedian D, Atkinson GM (2005) Stochastic finite-fault modeling based on a dynamic corner frequency. Bull Seismol Soc Am 95:995–1010. https://doi.org/10.1785/0120030207

Mukhopadhyay B (2011) Clusters of moderate size earthquakes along Main Central Thrust (MCT) in Himalaya. Int J Geosci 2:318–325

Mukhopadhyay B, Fnais M, Mukhopadhyay M, Dasgupta S (2010) Seismic cluster analysis for the Burmese-Andaman and West Sunda Arc: Insight into subduction kinematics and seismic potentiality. Geomatics, Nat Hazards Risk 1:283–314. https://doi.org/10.1080/19475705.2010.494014

Mulargia F, Tinti S (1985) Seismic sample areas defined from incomplete catalogues : an application to the Italian territory. Phys Earth Planet Inter 40:273–300

Nanda SJ, Panda G (2015) Design of computationally efficient density-based clustering algorithms. Data Knowl Eng 95:23–38. https://doi.org/10.1016/j.datak.2014.11.004

Nandy DR (2001) Geodynamics of Northeastern India and the Adjoining Region,. ACB Publication, Kolkata

Nath SK (2007) Seismic microzonation atlas of Guwahati region. Dep Sci Technol Gov India, New Delhi

Nath SK, Adhikari MD, Maiti SK, et al (2014) Earthquake scenario in West Bengal with emphasis on seismic hazard microzonation of the city of Kolkata, India. Nat Hazards Earth

Syst Sci 14:2549–2575. https://doi.org/10.5194/nhess-14-2549-2014

Nath SK, Raj A, Thingbaijam KKS, Kumar A (2009) ground Motion synthesis and seismic scenario in guwahati city—a stochastic approach. Seismol Res Lett 80:233–242. https://doi.org/10.1785/gssrl.80.2.233

Nath SK, Thingbaijam KKS (2012) Probabilistic seismic hazard assessment of India. Seismol Res Lett 83:135–149. https://doi.org/10.1785/gssrl.83.1.135

Nath SK, Thingbaijam KKS (2011a) Assessment of seismic site conditions: a case study from Guwahati city, Northeast India. Pure Appl Geophys 168:1645–1668

Nath SK, Thingbaijam KKS (2011b) Peak ground motion predictions in India : an appraisal for rock sites. J Seismol 15:295–315. https://doi.org/10.1007/s10950-010-9224-5

Nath SK, Vyas M, Pal I, Sengupta P (2005) A seismic hazard scenario in the Sikkim Himalaya from seismotectonics, spectral amplification, source parameterization, and spectral attenuation laws using strong motion seismometry. J Geophys Res Solid Earth 110:. https://doi.org/https://doi.org/10.1029/2004JB003199

Naylor M, Orfanogiannaki K, Harte D (2010) Exploratory data analysis: magnitude, space, and time. In: Community Online Resour. Stat. Seism. Anal.

NDMA (2010) Development of Probabilistic Seismic Hazard Map of India Technical Report

Novianti P, Setyorini D, Rafflesia U (2017) K-means cluster analysis in earthquake epicenter clustering. Int J Adv Intell Informatics 3:81–89. https://doi.org/10.26555/ijain.v3i2.100

Oldham R (1883) A catalogue of Indian earthquakes from the earliest time to the end of A.D. 1869. Mem Geol Surv India 19:163–215

Oldham RD (1899) Report on the Great Earthquake of 12 June 1897. Mem Geol Soc India 379

Pallav K, Raghu Kanth STG, Singh KD (2012a) Probabilistic seismic hazard estimation of Manipur , India. J Geophys Eng 9:516–533. https://doi.org/10.1088/1742-2132/9/5/516

Pallav K, Raghukanth S, Singh K (2012b) Probabilistic seismic hazard estimation of Manipur, India. J Geophys Eng 9:516–533. https://doi.org/10.1088/1742-2132/9/5/516

Pallav K, Raghukanth STG, Singh KD (2015) Estimation of seismic site coefficient and seismic microzonation of Imphal City, India, using the probabilistic approach. Acta Geophys 63:1339–1367

Pandey A, Jayangondaperumal R, Hetényi G, et al (2021) Establishing primary surface rupture evidence and magnitude of the 1697 CE Sadiya earthquake at the Eastern Himalayan Frontal thrust, India. Sci Rep 11:1–15. https://doi.org/10.1038/s41598-020-79571-w

Pandey AK, Chingtham P, Roy PNS, et al (2017) Homogeneous earthquake catalogue for Northeast region of India using robust statistical approaches. Geomatics, Nat Hazards Risk 8:1477–1491. https://doi.org/10.1080/19475705.2017.1345794

Pandey MR, Molnar P (1988) The distribution of intensity of the Bihar-Nepal earthquake of 15

January 1934 and bounds on the extent of the rupture zone. J Geol Soc Nepal 5:22–44

Perkins DM (1978) Acceleration hazard map sensitivity to input seismic parameters. Boll Geofis Teor Appl 20:188–196

Pettenati F, Sirovich L, Bjerrum LW (2017) Fault sources and kinematics of the 1897 M w 8.1 Assam and the 1934 M s 8.2 Nepal earthquakes retrieved by KF-NGA inversion and their seismotectonic implications. Bull Seismol Soc Am 107:2480–2489

Prakash R, Singh RK, Srivastava HN (2016) Nepal earthquake 25 April 2015: source parameters, precursory pattern and hazard assessment. Geomatics, Nat hazards risk 7:1769–1784

Priyanka RS, Jayangondaperumal R, Pandey A, et al (2017) Primary surface rupture of the 1950 Tibet-Assam great earthquake along the eastern Himalayan front, India. Sci Rep 7:5433. https://doi.org/10.1038/s41598-017-05644-y

Raghu Kanth STG, Dixit J, Dash S (2011) Ground motion for scenario earthquakes at Guwahati city. Acta Geod Geophys Hungarica 46:326–346. https://doi.org/10.1556/AGeod.46.2011.3.5

Raghu Kanth STG, Singh KD, Pallav K (2009) Deterministic Seismic Scenarios for Imphal City. Pure Appl Geophys 166:641–672. https://doi.org/10.1007/s00024-009-0460-y

Raghukanth STG, Dash SK, Dixit J (2011) Ground Motion for Scenario Earthquakes at Guwahati City. Acta Geod Geoph Hung 46:326–346

Raghukanth STG, Somala SN (2009) Modeling of strong-motion data in Northeastern India: Q, stress drop, and site amplification. Bull Seismol Soc Am 99:705–725. https://doi.org/10.1785/0120080025

Rahman T, Nunia B (2020) A new ground motion prediction model for Northeast India (NEI) shallow crustal. IOP Conf Ser Earth Environ Sci 491:. https://doi.org/10.1088/1755-1315/491/1/012029

Rajendran CP, Rajendran K (2001) Characteristics of deformation and past seismicity associated with the 1819 Kutch earthquake, northwestern India. Bull Seismol Soc Am 91:407–426

Rajendran CP, Rajendran K, Duarah BP, et al (2004) Interpreting the style of faulting and paleoseismicity associated with the 1897 Shillong, northeast India, earthquake: Implications for regional tectonism. Tectonics 23:1–12. https://doi.org/10.1029/2003TC001605

Ramanna CK, Dodagoudar GR (2012) Probabilistic seismic hazard analysis using kernel density estimation technique for Chennai, India. Georisk 6:1–15. https://doi.org/10.1080/17499518.2010.496073

Ramkrishnan R, Sreevalsa K, Sitharam TG (2022) Strong Motion Data Based Regional Ground Motion Prediction Equations for North East India Based on Non-Linear Regression Models. J Earthq Eng 26:2927–2947. https://doi.org/10.1080/13632469.2020.1778586

Rao NP, Kumar MR (1997) Uplift and tectonics of the Shillong Plateau, Northeast India. J Phys Earth 45:167–176. https://doi.org/10.4294/jpe1952.45.167

Reddy D V, Nagabhushanam P, Kumar D, et al (2009) The great 1950 Assam Earthquake revisited : Field evidences of liquefaction and search for paleoseismic events. Tectonophysics 474:463–472. https://doi.org/10.1016/j.tecto.2009.04.024

Rehman K, Burton PW, Weatherill GA (2014) K-means cluster analysis and seismicity partitioning for Pakistan. J Seismol 18:401–419. https://doi.org/10.1007/s10950-013-9415-y

Reiter L (1990) Earthquake hazard analysis: issues and insights. New York: Columbia University Press

Rietbrock A, Strasser F, Edwards B (2013) A Stochastic Earthquake Ground-Motion Prediction Model for the United Kingdom. Bull Seismol Soc Am 103:57–77. https://doi.org/10.1785/0120110231

Roe P, Georgsen F, Abrahamsen P (2014) An Uncertainty Model for Fault Shape and Location. Math Geosci 46:957–969. https://doi.org/10.1007/s11004-014-9536-z

Rout MM, Das J, Kamal (2018) Probabilistic seismic hazard for Himalayan region using kernel estimation method (zone-free method). Nat Hazards 93:967–985. https://doi.org/10.1007/s11069-018-3336-6

Rydelek PA, Sacks IS (1989) Testing the completeness of earthquake catalogues and the hypothesis of self-similarity. Nature 337:251–253. https://doi.org/10.1038/337249a0

Sabri MSA (2002) Earthquake intensity attenuation relationship for Bangladesh and its surrounding region. University of Engineering and Technology, Dhaka

Sawires R, Peláez JA, Ibrahim HA, et al (2016) Delineation and characterization of a new seismic source model for seismic hazard studies in Egypt. Nat Hazards 80:1823–1864. https://doi.org/10.1007/s11069-015-2034-x

Saxena A, Prasad M, Gupta A, et al (2017) A review of clustering techniques and developments. Neurocomputing 267:664–681. https://doi.org/10.1016/j.neucom.2017.06.053

Schnabel PB (1973) Effects of Local Geology and Distance from Source on Earthquake Ground Motions. University of California, Berkeley

Seed H., Idriss I. (1970) Soil moduli and damping factors for dynamic response analysis. University of California, Berkeley

Shang X, Li X, Morales-Esteban A, et al (2017) K -Means Cluster for Seismicity Partitioning and Geological Structure Interpretation, with Application to the Yongshaba Mine (China). Shock Vib 2017:. https://doi.org/10.1155/2017/5913041

Shang X, Li X, Morales-Esteban A, et al (2018) Data field-based K-means clustering for spatio-temporal seismicity analysis and hazard assessment. Remote Sens 10:1–22. https://doi.org/10.3390/rs10030461

Sharma ML, Bungum H (2006) New strong ground motion spectral attenuation relations for the Himalayan region. In: 1st European Conference on Earthquake Engineering and Seismology. Switzerland

Sharma ML, Malik H (2006) Probabilistic seismic hazard analysis and estimation of spectral strong ground motion on bed rock in north east India. In: 4th International Conference on Earthquake Engineering. Taipei, Taiwan

Shi Y, Bolt BA (1982) The standard error of the magnitude frequency b value. Bull Seismol Soc Am 72:1677–1687

Shiddiqi HA, Tun PP, Kyaw TL, Ottemöller L (2018) Source study of the 24 August 2016 Mw 6.8 Chauk, Myanmar, earthquake. Seismol Res Lett 89:1773–1785

Sil A, Sitharam TG (2013) Site Response Evaluation of Agartala City Using Geophysical and Geotechnical Data. Int J Geotech Earthq Eng 4:53–73. https://doi.org/10.4018/ijgee.2013070104

Sil A, Sitharam TG, Kolathayar S (2013) Probabilistic seismic hazard analysis of Tripura and Mizoram states. Nat Hazards 68:1089–1108. https://doi.org/10.1007/s11069-013-0678-y

Singh N, Rahman T, Wong I (2016) A New Ground - Motion Prediction Model for Northeastern India (NEI) Crustal A New Ground-Motion Prediction Model for Northeastern India (NEI) Crustal Earthquakes. Bull Seismol Soc Am 106:1282–1297. https://doi.org/10.1785/0120150180

Singh SK, Garcia D, Pacheco JF, et al (2004) Q of the Indian Shield. Bull Seismol Soc Am 94:1564–1570

Singh SK, Ordaz M, Dattatrayam RS, Gupta HK (1999) A spectral analysis of the 21 May 1997, Jabalpur, India, earthquake (Mw= 5.8) and estimation of ground motion from future earthquakes in the Indian shield region. Bull Seismol Soc Am 89:1620–1630

Sitharam TG, Anbazhagan P (2007) Seismic Hazard Analysis for the Bangalore Region. Nat Hazards 40:261–278. https://doi.org/10.1007/s11069-006-0012-z

Sitharam TG, Kolathayar S (2013) Seismic hazard analysis of India using areal sources. J Asian Earth Sci 62:647–653. https://doi.org/10.1016/j.jseaes.2012.11.013

Sitharam TG, Sil A (2014) Comprehensive seismic hazard assessment of Tripura and Mizoram states. J Earth Syst Sci 123:837–857. https://doi.org/10.1007/s12040-014-0438-8

Sloan RA, Elliott JR, Searle MP, Morley CK (2017) Active tectonics of Myanmar and the Andaman Sea. Geol Soc London, Mem 48:19–52

Srivastava HN, Verma M, Bansal BK, Sutar AK (2015) Discriminatory characteristics of seismic gaps in Himalaya. Geomatics, Nat Hazards Risk 6:224–242. https://doi.org/10.1080/19475705.2013.839483

Stepp J (1972) Analysis of completeness of the earthquake sample in the Puget sound area and its effect onstatistical estimates of earthquake hazard. In: International conference on microzonation. Seattle, USA, pp 897–910

Stewart J, Choi Y, Liu A (2001) Amplification factors for spectral acceleration in active regions. In: Annual Meeting Research Digest No. p 14

Stewart JP, Liu AH, Huang M (2000) Ground motion amplification as a function of surface geology. In: Proc. SMIP2000 Seminar on Utilization of Strong Motion Data. pp 1–22

Sukhija BSŁ, Rao MN, Reddy D V, et al (1999) Paleoliquefaction evidence and periodicity of large prehistoric earthquakes in Shillong Plateau , India. Earth Planet Sci Lett 167:269–282

Sun JI, Golesorkhi R, Seed HB (1988) Dynamic moduli and damping ratios for cohesive soils. Report no EERC 88-15. University of California Berkeley

Thingbaijam KKS, Nath SK, Yadav A, et al (2008) Recent seismicity in Northeast India and its adjoining region. J Seismol 12:107–123. https://doi.org/10.1007/s10950-007-9074-y

Thomas PJ, Reddy D V., Kumar D, et al (2007) Optical dating of liquefaction features to constrain prehistoric earthquakes in Upper Assam, NE India-some preliminary results. Quat Geochronol 2:278–283. https://doi.org/10.1016/j.quageo.2006.03.013

Toro GR (2002) Modification of the Toro et al. (1997) Attenuation Equations For Large Magnitudes And Short Distances. Risk Eng Inc

Uhrhammer RA (1976) Characteristics of Northern and Central California Seismicity. Earthq notes

USGS Search Earthquake Catalog. https://earthquake.usgs.gov/earthquakes/search/. Accessed 1 May 2019

USNRC RG-1.208 (2007) A Performance-Based Approach to Define The Site-Specific Earthquake Ground Motion A.

Vallejo LIG De, Capote R, Cabrera L, et al (2003) Paleoearthquake evidence in Tenerife (Canary Islands) and possible seismotectonic sources. Mar Geophys Res 149–160. https://doi.org/10.1007/s11001-004-5883-3

van Stiphout T, Zhuang J, Marsan D (2016) Seismicity declustering , Community Online Resource for Statistical Seismicity Analysis. https://doi.org/10.5078/corssa-52382934

Vasak P, Suorineni FT, Kaiser PK, Thibodeau D (2004) Hazard map approach using space-time clustering analysis of mining-induced microseismicity. In: Canadian Institute of Mining and Metallurgy Annual General Meeting, Edmonton

Vernant P, Bilham R, Szeliga W, et al (2014) Clockwise rotation of the Brahmaputra Valley relative to India: Tectonic convergence in the eastern Himalaya, Naga Hills, and Shillong Plateau. J Geophys Res Solid Earth 119:3782–3803. https://doi.org/10.1002/2015JB012608.Received

Vipin KS, Sitharam TG (2013) Delineation of seismic source zones based on seismicity parameters and probabilistic evaluation of seismic hazard using logic tree approach. J Earth Syst Sci 122:661–676. https://doi.org/10.1007/s12040-013-0300-4

Vittori E, Labini SS, Leonello S (1991) Palaeoseismology : review of the state-of-the-art. Tectonophysics 193:9–32

Wang Y, Sieh K, Tun ST, et al (2014) Active tectonic and earthquake Myanmar region. J Geophys Res Solid Earth 119:3576–3822. https://doi.org/10.1002/2013JB010762

Wardlaw RL, Frohlich C, Davis SD (1990) Evaluation of precursory seismic quiescence in sixteen subduction zones using single-link cluster analysis. Pure Appl Geophys PAGEOPH 134:57–78. https://doi.org/10.1007/BF00878080

Weatherill G, Burton PW (2009) Delineation of shallow seismic source zones using K -means cluster analysis , with application to the Aegean region. Geophys J Int 176:565–588. https://doi.org/10.1111/j.1365-246X.2008.03997.x

Weichart D (1980) Estimation Of The Earthquake Recurrence Parameters For Unequal Observation Periods For Different Magnitudes. Bull Seismol Soc Am 70:1337–1346

Wells DL, Coppersmith KJ (1994) New Empirical Relationships among Magnitude, Rupture Length, Rupture Width, Rupture Area, and Surface Displacement. Bull Seismol Soc Am 84:974–1002

Wiemer S, Katsumata K (1999) Spatial variability of seismicity pa- rameters in aftershock zones. J Geophys Res 103:13135–13151

Wiemer S, McNutt S (1997) Variations in frequency-magnitude dis- tribution with depth in two volcanic areas: Mount St. Helens, Wash- ington, and Mt. Spurr, Alask. Geophys Res Lett 24:189–192

Wiemer S, Schorlemmer D (2007) ALM: An Asperity-based Likelihood Model for California. Seismol Res Lett 78:134–140

Wiemer S, Wyss M (2000) Minimum Magnitude of Completeness in Earthquake Catalogs : Examples from Alaska , the Western United States , and Japan. Bull Seismol Soc Am 90:859–869

Woo G (1996) Kernel estimation methods for seismic hazard area source modeling. Bull Seismol Soc Am 86:353–362

Wyss M, Schorlemmer D, Wiemer S (2000) Mapping asperities by minima of local recurrence time: The San Jacinto-Elsinore fault zones. J Geophys Res 105:7829–7844

Youngs RR, Chiou SJ, Silva WJ, Humphrey JR (1997) Strong Ground Motion Attenuation Relationships for Subduction Zone Earthquakes. Seismol Res Lett 68:58–73

Zamani A, Hashemi N (2004) Computer-based self-organized tectonic zoning: A tentative pattern recognition for Iran. Comput Geosci 30:705–718. https://doi.org/10.1016/j.cageo.2004.04.002

Zamani A, Khalili M, Gerami A (2011) Computer-based self-organized tectonic zoning revisited: Scientific criterion for determining the optimum number of zones. Tectonophysics 510:207–216. https://doi.org/10.1016/j.tecto.2011.07.004

Zechar JD, Jordan TH (2008) Testing alarm-based earthquake predictions. Geophys J Int 172:715–724. https://doi.org/10.1111/j.1365-246X.2007.03676.x

Zhao Y, Grujic D, Baruah S, et al (2021) Paleoseismological Findings at a New Trench Indicate the 1714 M8.1 Earthquake Ruptured the Main Frontal Thrust Over all the Bhutan Himalaya. Front Earth Sci 9:1–18. https://doi.org/10.3389/feart.2021.689457

List of Publications

Journal Papers

1. **Borah, N.**, & Kumar, A. (2022). Probabilistic seismic hazard analysis of the North-East India towards identification of contributing seismic sources. **Geomatics, Natural Hazards and Risk**, 14(1), 1-38.
2. **Borah, N.**, Kumar, A., & Dhanotiya, R. (2021). Seismic source zonation for NE India on the basis of past EQs and spatial distribution of seismicity parameters. **Journal of Seismology**, 1-24.
3. Kumar A., **Borah, N.**, Naik S.P., & Baro O. (2018), Detailed review on methodologies available to find preinstrumental missing earthquakes of the present catalogue with the relevance to seismicity assessment of the Northeast India, **Indian Geotechnical Journal**, https://doi.org/10.1007/s40098-018-0336-0.
4. **Borah, N.**, Mondal JK, Kumar, A., & Harinarayan NH. Ground Motion Prediction Equation for North-East India, applicable to different site classes, (under preparation)
5. **Borah, N.**, & Kumar, A. Probabilistic Seismic Hazard Analysis of North-East India by utilising Kernel Density method to deal with the uncertainty in location and magnitude of Past Earthquakes, (Submitted to Natural Hazards)

Conferences

1. **Borah, N.**, & Kumar, A. (2021) Assessing if Foreshocks and Aftershocks are Dependent Events of the Mainshock for North-east India. 17th WCEE, Japan.