



Regional stochastic GMPE with available recorded data for active region – Application to the Himalayan region



Ketan Bajaj, P. Anbazhagan*

Indian Institute of Science, Bangalore, India

ARTICLE INFO

Keywords:

GMPE
Fourier amplitude spectrum
Monte-Carlo
Sensitivity analysis
Himalayan region

ABSTRACT

New ground motion prediction equation for the active Himalayan region for a wide range of moment magnitude (M_w , 4–9) and distance (10–750 km) is developed. For simulating the synthetic ground motions; source, path, and site terms are derived using the Fourier amplitude spectrum of the recorded ground motion data. Uncertainty of input parameters is propagated through simulation by random sampling of the corresponding distribution of input parameters. Synthetic and recorded data are regressed using random-effect maximum likelihood regression algorithm by determining the compatible functional form. Sensitivity analysis is used in determining the impact of uncertainty of each input parameter on standard deviation of the regression residuals about the median prediction equation. Major contribution to total uncertainty is from Kappa factor in case of within-event terms and from stress drop in case of event-to-event variability. Predicted and recorded response spectra is matching within ± 1 standard deviation for the entire period range.

1. Introduction

The Himalayan region and the contiguous Sindhu-Ganga alluvial plain to the south of it have been associated with high level of seismicity. Many of the researchers [1,2] have highlighted the possibility of occurrence of the large magnitude earthquakes ($M_w > 8$) in future in the Himalayan region. Moreover, to develop the representative earthquake hazard prediction map for the Northern India, a robust ground motion prediction equation (GMPE) is needed. Hence, there is a need for coordinated scientific efforts for reliable seismic hazard prediction so that the earthquake hazard can be minimized through effective disaster planning and management, and by low-cost mitigation methods.

GMPEs play an important role in defining the conditional distribution of ground-motion amplitudes for each rupture scenario considered within the hazard calculations. For the Himalayan region, strong motion data for wide range of magnitudes and hypocentral distances are not available as of today. The lack of recorded data makes the hazard estimation more challenging especially for the Himalayan region, having such a diverse seismology [6]. However, the advancement in ground motion simulation algorithms and regression techniques help to alleviate this issue theoretically. Various researchers [e.g. 3, 4, 5] have developed GMPEs that are based on the stochastic simulations which have been calibrated using recorded weak or strong ground motion.

Proper calibration of GMPE is more challenging with the variability in the input parameters for synthetic simulation of ground motion. It is resolved by studying the impact of distribution of stochastic models' inputs parameters by estimating the between-event and within-event ground motion variability.

Based on an available recorded and simulated ground motions for the Himalayan region, various researchers have derived the region-specific GMPEs. These GMPEs were developed by Singh et al. [6], Sharma [7], Nath et al. [8], Sharma and Bungum [9], Das et al. [10], Baruah et al. [11], Sharma et al. [12], Gupta [13] and Anbazhagan et al. [14] using simulated and recorded earthquake data. In addition to these equations, National Disaster Management Authority [15], Government of India, developed region-specific GMPEs for probabilistic seismic hazard mapping of India considering only simulated data. However, most of these GMPEs are not applicable for distances (hypocentral or epicentral) larger than 100 km. While studying the iso-seismal map of large earthquakes ($M_w > 7$) in the Himalayan region, it has been observed that macroseismic intensity of V-VI is reported for distance more than 500 km. For example, for 1905 Kangra earthquake (M_w 7.8) and 1950 Assam earthquake (M_w 8.7), macroseismic intensity of V-VI is observed for distance more than 500 km. Hence, GMPEs developed for distance of less than 100 km would not be suitable for ground-motion estimation at large distances. Moreover, the input parameters for simulating the synthetic ground motions are either

* Corresponding author.

E-mail addresses: ketanbajaj@iisc.ac.in (K. Bajaj), anbazhagan@iisc.ac.in (P. Anbazhagan).

assumed or used from the other active regions in developing the Himalayan GMPE. However, these parameters are revisited and derived in the present study.

The main aim of the present study is to develop a robust GMPE for the Himalayan region by using different sets of seismological parameters. Fourier Amplitude Spectrum (FAS) of the recorded ground motions have been studied, and region-dependent source and site parameters i.e. stress drop, anelastic attenuation, kappa and corner frequency have been derived explicitly for earthquakes in the Himalayan region. These parameters have been further used for stochastic simulation of the ground motion database. The simulation of PGA and spectral acceleration values at period 0.01–10 s was performed for a wide M_w range (4–9) and hypocentral distance range (10–750 km) for rock condition having shear wave velocity up to 30 m depth (V_{s30}) of 2000 m/s. The functional form suitable for the region that represents magnitude and hypocentral distance scaling was used for deriving a new GMPE. Further, a discussion has been made on the region-based input parameters and captures the probability distribution of these parameters. The impact of these distributions were studied for different frequencies by varying the magnitude-distance pair. All the input parameters were considered as random variables, and the reduction of uncertainty in the simulation was accomplished through random sampling of the input parameters distribution. A Monte Carlo scheme has been used for determining the confidence interval and standard error in regression parameters. Sensitivity analysis was performed to assess the impact of the uncertainty of each input parameter on the final GMPE uncertainty. Finally, the new GMPE was tested with the recorded strong motion data.

2. Study area and instrumented ground motion data

The study area mainly lies in the Himalayan region that includes Kashmir Himalayan (KH), Kumaon-Garhwal Himalayan (KGH), Bihar-Nepal Himalaya (BNH) and Northeastern part of Himalayan region (NEH). The Himalayan belt is about 2900 km long and it is the highest

mountain chain in the world. It is also one of the most challenging tectonic settings in geoscience due to its complex seismotectonic setting and frequent seismogenesis. Continuous collision of Indian plate with the Asian plate at a rate of 15–20 mm/year [1] results in building of high stress in the Indian plate. The Himalayan geodynamics is the cause of the largest earthquakes (1897 Assam, 1905 Kangra, 1934 Bihar-Nepal, 1950 Assam, 2011 Sikkim, 2015 Nepal earthquake) in the foothills and north of the Main Boundary Thrust (MBT). These earthquakes were well studied by various authors [2,16,17] and it was commented that these earthquakes were failed to rupture to the surface and probably transferred the stress towards the Himalayan foothills. Bilham et al. [1] concluded the potential slip in the range of 9–14 m with the expected earthquake as large as M_w 8.9 in the KH region. BNH has a dense cluster of seismicity near to the Main Central Thrust (MCT) and the Himalayan Frontal Thrust (HFT). It is also close to the epicenter of 1934 Nepal earthquake and associated with the stress concentration zones [18]. The possibility of the occurrence of the largest earthquake in NEH is predicted by Srivastava et al. [19]. Studying the micro-seismicity, paleoseismicity, GPS and variation in local tectonics, Srivastava et al. [19] demarked the whole Himalayan region into 10 seismic gaps with the possibility of occurrence of moderate to large earthquake. Hence, while developing a new GMPE, large earthquakes need to be simulated in these seismic gaps, so that hazard estimation from any future earthquake ($M_w > 8$) can be evaluated. Large number of active faults and cluster of small to large events ($4 < M_w < 7.0$) occurring in the Himalayan region demands the development of new GMPE to estimate the hazard in the highly populated contiguous Sindhu-Ganga alluvial plain.

Recorded data used in this study consists of 78 strong to moderate earthquakes that occurred in the Himalayan region from 1988 to 2015 with a M_w of 4.5–7.8 and a hypocentral distance (R) between 10 and 500 km. The strong-motion data was collected from the strong motion instrumentation network of Indian Institute of Technology, Roorkee (IITR) (<http://www.pesmos.in/>, last accessed March 2017) and also from Virtual Data Center (VDC) (<https://strongmotioncenter.org/vdc/scripts/default.plx>, last accessed December 2017). These stations cover

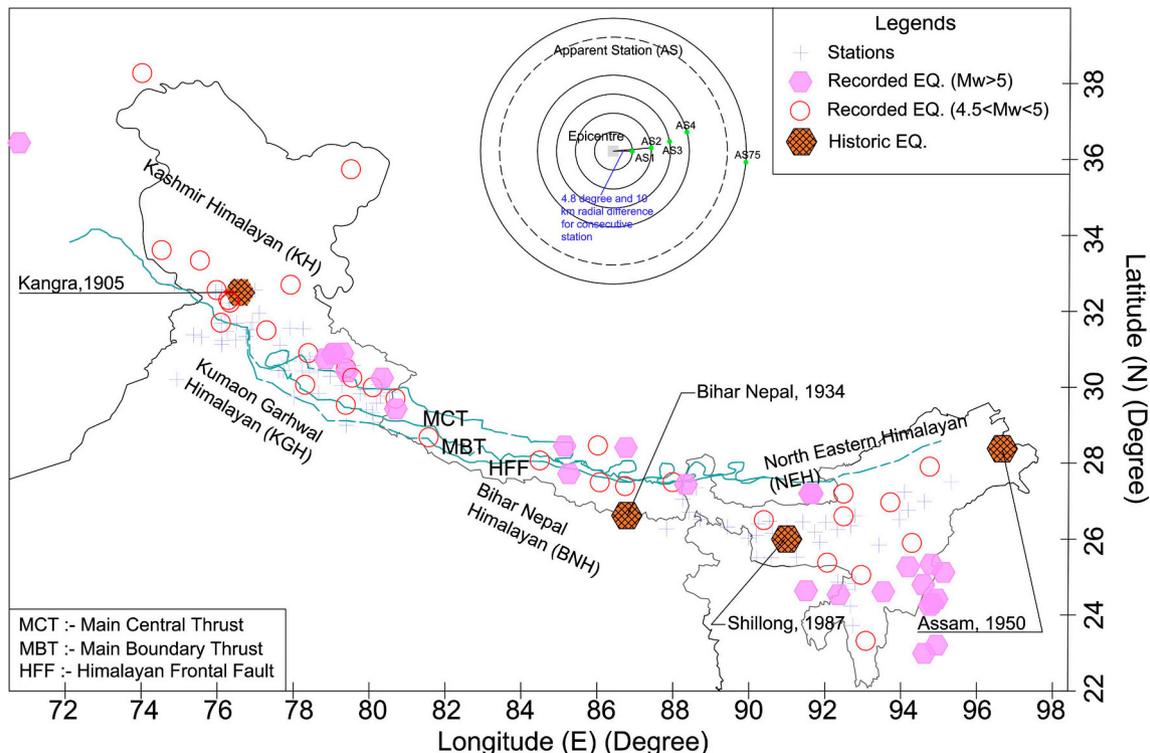


Fig. 1. Location of the recording stations (plus sign) and epicenter of the recorded earthquakes used in the study including historic earthquake occurred in the Himalayan region. The assigning of apparent station for simulating the ground motion is given in the top right, the explanation is given in the text.

the Indian Himalayan range from Jammu and Kashmir to Meghalaya (marked as plus sign in Fig. 1). Detailed description of these strong motion accelerographs and data processing of the waveforms are given in Kumar et al. [50]. Out of the total 520 ground motion recordings, 252 were collected from the IITR, 68 ground motions recorded before 2005 were collected from VDC and the rest 200 from Indian seismic and GNSS network (<http://www.isgn.gov.in/ISGN/>, last accessed December 2018). Out of 520 recordings, 241 are rock recordings ($V_{S30} \geq 760$ m/s) and the remaining 279 are soil recordings [52]. Only rock recordings are used in the present study. The processing of the strong motion data involved baseline correction, instrumental scaling, and frequency filtering. The strong motion database was processed according to the procedure suggested by Boore and Bommer [20] and Joshi et al. [21]. The detail of the instruments can be referred from Kumar et al. [50] and Indian Seismic and GNSS network website (<http://www.isgn.gov.in/ISGN/>). The obtained database was corrected for baseline, instrumental response and amplification of recorder (Dr Dipankar Saikai, personal comm. 2017). The obtained database was band passed for the frequency range of 0.5–35 Hz [21,50]. Database with signal-to-noise ratios greater than 3 were used for further analysis [53].

3. Seismological model parameters

Following Boore [22], the amplitude spectrum of ground motion $A(f)$ can be written as product of a source function, $\Omega(f)$, a propagation path term, $P(f)$ and a site function, $S(f)$ in the frequency domain

$$A(M_w, R, V_{S30}, f) = \Omega(M_w, f)P(R, f)S(f) \quad (1)$$

The propagation path term, $P(f)$ can be expressed as

$$P(f) = Ge^{-\frac{\pi f R}{Q(f)\beta}} \quad (2)$$

Where, G is the geometric spreading and $Q(f)$ and β are the S-wave frequency dependent quality factor and velocity of the medium respectively. Following Boore [22], a trilinear form of geometric spreading can be specified as

$$G = \begin{cases} R^{-b_1}, & R \leq R_{x1} \\ R_{x1}^{-b_1} \left(\frac{R}{R_{x1}}\right)^{-b_2}, & R > R_{x1} \wedge R \leq R_{x2} \\ R_{x1}^{-b_1} \left(\frac{R_{x2}}{R_{x1}}\right)^{-b_2} \left(\frac{R}{R_{x2}}\right)^{-b_2}, & R > R_{x2} \end{cases} \quad (3)$$

R_{x1} and R_{x2} are the two reference distances and b_1 , b_2 and b_3 are the geometric spreading coefficients for the trilinear functional form corresponding to the near-field and far-field reference distances. The quality factor $Q(f)$ can be defined as

$$Q(f) = Q_0 \left(\frac{f}{f_0}\right)^n \quad (4)$$

Here, f_0 is the reference frequency and is generally taken as 1 Hz [23] and Q_0 is the value of Q at 1 Hz, n is the frequency parameter, which is close to 1 and varies from region to region; depending on heterogeneity of the medium [23]. The near surface attenuation ($P(f)$) is given by Anderson and Hough [24] as:

$$P(f) = \exp(-\pi f \kappa) \quad (5)$$

Here, κ is the kappa factor, typically an indicator that differentiates the older/harder rock from the younger upper crustal formations. κ is low in case of older/harder rock as compared to younger upper crustal formation. Beresnev and Atkinson [25] employed finite-fault modelling for simulating strong ground motion. In this algorithm, sub-source moment depends on the sub fault dimensions that can be determined using a relation given by Beresnev and Atkinson [25]. A large uncertainty is associated with the suggested relations due to the scarcity of large earthquake recordings. To overcome this constraint, Motazedian and Atkinson [26] introduced a concept of dynamic corner frequency

and enhanced the algorithm. The improved algorithm conserves the radiated energy at a higher frequency at any sampling of subfault size thereby controlling the relative amplitude of higher versus lower frequencies [8]. The source spectrum for the n^{th} subfault is given as

$$\Omega_n(f) = CM_{0n}H_n(2\pi f)^2 \left[1 + \left[\frac{f}{f_{0n}} \right]^2 \right] \quad (6)$$

Here, C is the scaling factor, M_{0n} is the seismic moment in dyne centimeters, H_n is the scaling factor responsible for converging the energy at the high-frequency spectral level of subfault and f_{0n} is the dynamic corner frequency. The scaling factor C can be defined as follow

$$C = \frac{R_{\phi j} \sqrt{2}}{4\pi \rho \beta^3} \quad (7)$$

Here, $R_{\phi j}$, β and ρ refer to radiation pattern, shear wave velocity and average crustal density (g/cm^3), respectively. The coefficient $\sqrt{2}$ in equation (7) arises as the product of free-surface site amplification and partitioning of energy in orthogonal directions [5]. The n^{th} subfault seismic moment (M_{0n}) can be calculated using slip distribution as follow:

$$M_{0n} = \frac{M_0 D_n}{\sum D_n} \quad (8)$$

where, D_n corresponding to the relative slip weight of the n^{th} subfault, M_0 is the seismic moment, and $\sum D_n$ denotes the total slip of the fault during the entire rupture process. The dynamic corner frequency and stress drop is related to each other as per Boore [22,27].

$$f_{0n} = 4.9 \times 10^6 [N_R(t)]^{-1/3} N^{1/3} \beta (\beta \Delta \sigma / M_0)^{1/3} \quad (9)$$

Here, $N_R(t)$ is the number of rupture subfaults at a time, t . N refers to the total number of subfaults totaling to $N_R(t)$ at the end of rupture and $\Delta \sigma$ is the stress drop. The scaling function in equation (6) is defined by Motazedian and Atkinson [26] as

$$H_n = \frac{\sqrt{\sum_j \left\{ \frac{f_j^2}{1 + \left(\frac{f_j}{f_0}\right)^2} \right\}^2}}{\sqrt{\sum_j \left\{ \frac{f_j^2}{1 + \left(\frac{f_j}{f_{0n}}\right)^2} \right\}^2}} \quad (10)$$

Here, M_{0n} is the seismic moment of the n^{th} subsource, f_0 is the corner frequency of the entire fault, f_{0n} is the corner frequency of the n^{th} subsource, and j is the counter over the frequency. The detailed derivation of stress model, attenuation model and duration from FAS of the recorded waveform in the Himalayan region is explained further.

3.1. Attenuation model

The attenuation parameters, i.e., geometric spreading and anelastic attenuation using FAS were derived by dividing the whole Himalayan region into four parts, such as KH, KGH, BNH and NEH, considering the variability in seismicity [29]. In the bilinear functional form, the rate of attenuation due to geometric spreading at a moderate distance is equivalent to the rate near the source. Additionally, Atkinson (2004) observed an incremental change in the amplitude of the FAS in the medium distance range (70–140 km) caused by post-critical reflection by the Moho and Conrad discontinuities. Therefore, trilinear geometric attenuation form was adopted in this study. Using a genetic algorithm, geometric spreading was derived by considering a trilinear model (Eq. (3)) with different decay coefficients for far source and near field. Similarly, frequency dependent quality factor (Eq. (4)) was also estimated. Genetic algorithm (GA), developed by Holland (1975) and Goldberg (1989) and used by Zandieh and Pezeshk [28] for deriving the

attenuation parameters for New Madrid seismic zone was used here for determining attenuation parameters. For estimating the variability in the attenuation model, two methods can be used: one is based on a posterior covariance matrix and the other on bootstrap method [28]. In this study, bootstrap method has been used, since the modelling error was not representing the actual variability in the posterior covariance matrix. From a total of 520 recorded ground motions, 80% has been chosen from random selection of earthquake in all the four regions and shown in Fig. 1. This operation has been done 200 times for each dataset, and GA was used for each dataset as given by Bajaj et al. [29]. In the GA, the constraints were only applied to b_1 , b_2 and b_3 in trilinear-hinged functional form (equation (3)). The values used for constraints were: $0.5 \leq b_1 \leq 2.0$, $-0.5 \leq b_2 \leq 0.5$ and $0.5 \leq b_3 \leq 1.3$. The parameters R_{x1} and R_{x2} in equation (3) were not treated as variables in the GA. Instead, the GA process was repeated for all combinations of $R_{x1} = 10, 20, 30, \dots, 150$ and $R_{x2} = 70, 80, 90, \dots, 200$. The overlapping values of R_{x1} and R_{x2} were also included to evaluate the possibility of the bilinear-hinged functional form for geometric spreading. These values for constraints and parameters (R_{x1} and R_{x2}) were selected based on similar studies conducted for similar tectonic regions (i.e., active subduction zones). The range of R_{x1} and R_{x2} was tested with the recorded ground motions. By applying different combination of the parameters, source amplitude corresponding to site was estimated. For each of the constraints, the residual between the observed and the predicted spectra have been calculated. The solution having the minimum mean absolute value of the residual is the best combination. In other words, the combination of values for b_1 , b_2 , b_3 , R_{x1} and R_{x2} that resulted in same source term for stations that recorded the same event was considered. Source, site, and path terms are correlated, and there is always a trade-off between these terms. Therefore, to reduce the trade-off between these parameters, the geometric spreading is determined using GA and further quality factor is estimated empirically. One must be cautioned against using the path model derived in this study along with the source or site parameters deduced from other studies assuming different types of the path models. Further the source spectra were corrected for geometric attenuation and regressed with R to calculate the frequency dependent Quality factor [28,29]. The attenuation parameters calculated for all four regions are given in Table 1. Not significant regional dependency is observed as far as these parameters are concerned expect in case of KGH and BNH. However, Q_0 and n values have a significant variation over the four regions. The near surface attenuation for the Himalayan region has been determined using the procedure defined by Anderson and Hough [24]. The determined κ factor for the Himalayan region is 0.01 s.

3.2. Stress drop

Many studies reported that the stress drop varies with the size of the earthquake [30], whereas studies on the large earthquakes concluded that stress parameters remain constant [31]. However, the dependency of the stress drop on the size of the earthquake is highly debatable [4]. In this study, the dependency of stress drop on the M_w for the Himalayan region has been examined. The stress drop has been calculated as per the procedure defined by Boore et al. [31]. For a given earthquake and oscillator period, motions were simulated at the distance of each recording with stress drop ranging from 1.0 bars to 1000 bars with an

increment of 2. Further, residual was determined for each observation for a given stress drop. The arithmetic average of all residuals for rock stations that recorded the earthquake was computed, treating each observation as an independent variable [31]. Similar to Boore et al. [31] quadratic fit was used to the average residual, logarithmic of stress drop pairs of values, and solved for the value of stress drop that gave zero residual.

Fig. 2 (a) shows the obtained stress drop versus magnitude for the whole Himalayan region. The stress drops obtained in this study has been compared with the stress drop obtained by various authors for the same region. From this study, variability in the stress drop has been seen for $M_w < 5.5$. However, an increasing trend in the stress drop has been observed with the increase in magnitude. Drawing a reliable conclusion for larger magnitudes is difficult because the recorded data for $M_w > 5.5$ are not sufficient.

The stress drop value derived by Nath et al. [8] for historical earthquakes are plotted in Fig. 2. The calculated stress drop of earthquake in the BHN region are larger than the eastern part of the Himalaya. Further, the average value up to 5.5 M_w is also shown in Fig. 2 (a), and it observes that, earthquakes near to Kashmir and Sikkim region has the maximum stress drop value, whereas, KGH has the lowest, comparing different regions of the Himalaya. Based on the global study of large magnitudes, Allmann and Shearer [32] suggested that for the larger magnitudes the stress drop remains constant, similar assumption is considered by Drouet and Cotton [4]. Furthermore, the approximate mean value of variation of stress drop of 3 MPa and 6 MPa for interplate and intraplate region respectively, was reported by Allmann and Shearer [32]. Based on PEER NGA database, Drouet and Cotton [4] reported a stress-drop value for large events approximately varying between 0.1 and 10 MPa. Considering both studies and based on the observations of variation of stress drop with magnitude; (Fig. 2) at large magnitudes a constant stress drop has been assumed at three values that are 50, 100 and 150 bars. Finally, bilinear model for stress drop has been used with kink point at 5.5 M_w . The equations derived between stress drop and moment magnitude is shown in Fig. 2 (b). The standard deviation in the stress drop value is studied with respect to magnitude and, standard deviation of 0.2–0.5 in natural logarithm scale is obtained.

3.3. Duration model

Another important parameter in the simulation of ground motion is the path duration function. The total duration (T_d) is the combination of the source duration (T_s) which is assumed to be the reciprocal of corner frequency [22], and path duration (T_p) which relates to propagation effects and the other effect linked to the site condition and complex source effect [22]. Generally, the observed T_d is used as the interval, between the time at which the integral reaches 5% and 95% of the maximum and 5% and 75% of the maximum. In the present study, as described in Boore and Thompson [33], the duration was computed using either acceleration or velocity database using D'_{95} . D'_{95} is defined as $D'_{95} = 2(D_{80} - D_{20})$; where D_{80} and D_{20} are the times at which the cumulative integral of acceleration squared reaches 20% and 80% of the final value, respectively [33]. To estimate the path duration (T_p) the source duration, i.e. inverse of corner frequency is subtracted from the total duration [33]. The corner frequency has been estimated by the

Table 1
Attenuation model parameter.

Region	b_1	b_2	b_3	R_{x1}	R_{x2}	Q_0	n	κ
KH	1.09 ± 0.07	-0.01	0.52 ± 0.05	50 ± 15	140 ± 25	214 ± 8.52	0.75 ± 0.12	0.018
BNH	1.11 ± 0.08	-0.02	0.45 ± 0.05	40 ± 10	160 ± 25	108 ± 10.72	0.62 ± 0.10	0.015
KGH	1.05 ± 0.07	-0.15	0.48 ± 0.05	70 ± 15	100 ± 15	76 ± 9.87	1.31 ± 0.11	0.011
NEH	1.02 ± 0.09	-0.01	0.58 ± 0.05	60 ± 10	140 ± 15	186 ± 12.31	0.86 ± 0.11	0.016

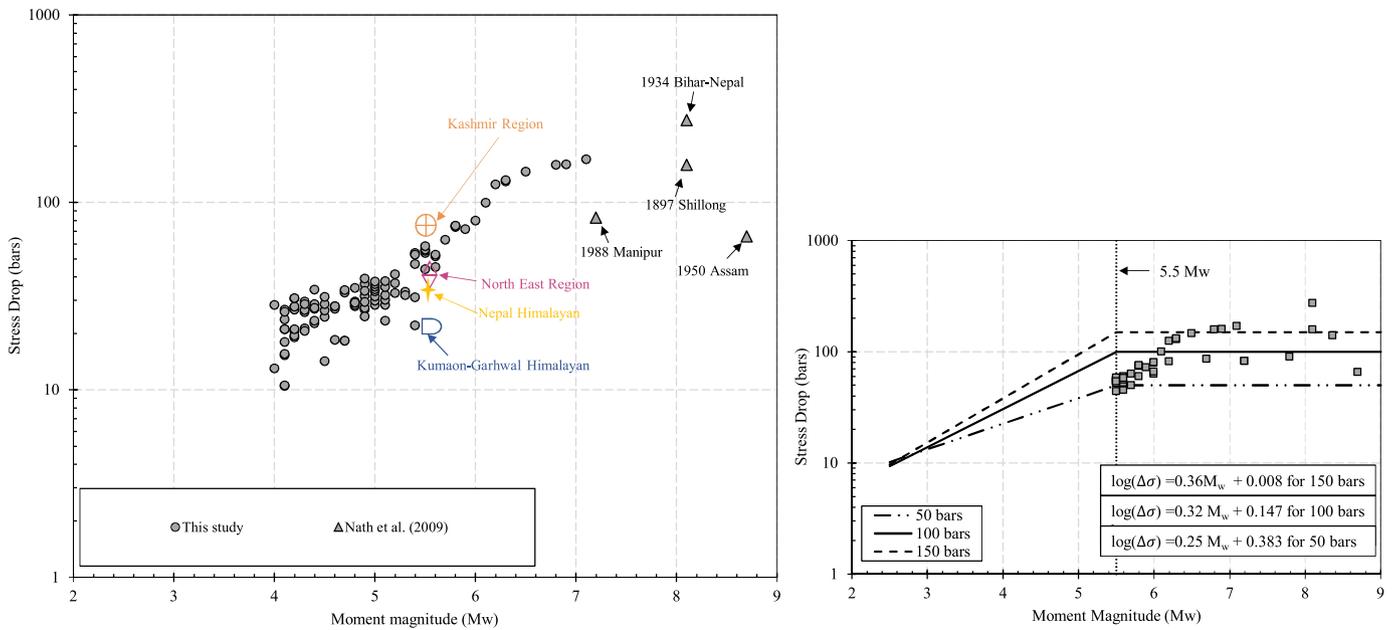


Fig. 2. (a) Inverted stress drops versus moment magnitude from the present and previous study for different regions of the Himalaya. The large symbols represent the average stress drop value for different regions up to 5.5 moment magnitude and the filled triangles represent the historic earthquakes (Nath et al., 2009), (b) Models for stress parameter derived from average stress value for small-events and three stress drop value (50, 100 and 150 bars) for whole Himalayan regions.

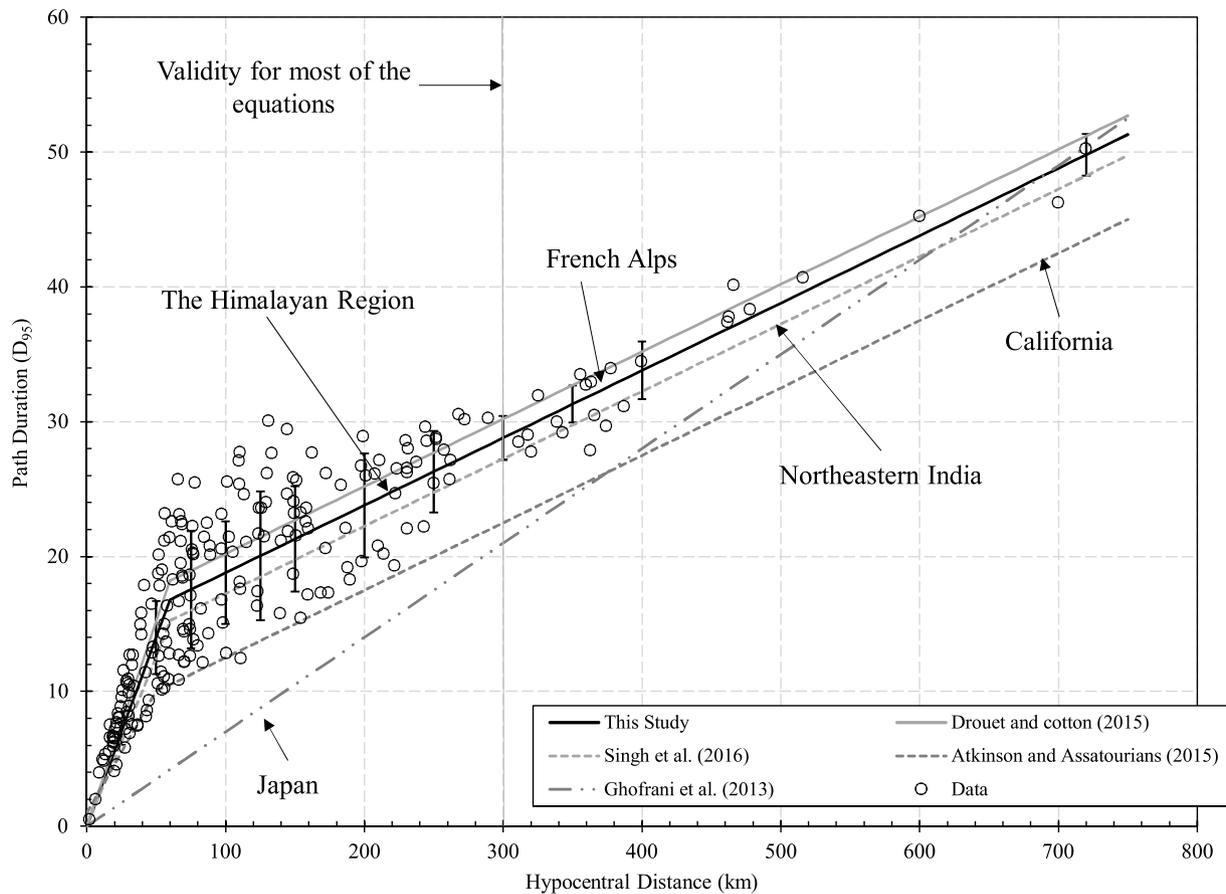


Fig. 3. Path Duration computed using the recorded data set of the Himalayan region as a function of hypocentral distance using D_{95} energy criteria. The average with error bars at different distances is shown and the two-segment linear fit of the Himalayan data is compared with the previous studies of different region but same seismotectonic.

method proposed by Andrews [51].

Fig. 3 shows the variation of path duration using D_{95} and based on the database of acceleration for the entire Himalayan region. Similar results have been studied for the velocity. As expected, velocity and acceleration lead to similar results. It was also observed from the analysis that a two-segment linear model fits well with the average data per distance bin (See Fig. 3), with a kink point at 60 km.

The whole Himalayan data recorded at rock site has been used in deriving the path duration model. Consistency between output and input path duration was checked from the stochastic simulation of the ground motion for different combinations of magnitude and distances. The input path duration (using eq. (11)) and output path duration (computed using the synthetic ground motion) was compared. It was observed from the comparison that for small magnitudes the input and output durations are almost similar but for the high magnitudes ($M_w > 5.5$), the output duration is approximately 0.94 of the input durations. Similar results were found by Edward and Fäh [3], Drouet and Cotton [4] and Singh et al. [5]. Hence, the input model was adjusted by a factor of 1.05 for D_{95} energy model [4]. The final duration model can be expressed as

$$T_p = \begin{cases} R_h \times 16.8/60, & R_h < 60 \text{ km} \\ 16.8 + 0.05 \times (R_h - 60), & R_h \geq 60 \text{ km} \end{cases} \quad (11)$$

Here, R_h is the hypocentral distance.

The uncertainty in the path duration has been carried out using the uniform distribution on the duration value at the kink point, with a standard deviation of 6.5 s and on the slope of the last segment, using a standard deviation half of its value which is consistent with Fig. 3. Fig. 3 shows the comparison of variation of duration with respect to hypocentral distance for the new equation and the existing worldwide equations. Most of the duration equations are valid up to 300 kms (see Fig. 3), however these equations have been extrapolated till 750 kms to observe the variation of existing equation with the newly derived equation. The current duration model lies in between duration model proposed by Drouet and Cotton [4] and Singh et al. [5] respectively for French Alps and Northeast India.

Table 2
Significant earthquakes in the Himalayan region considered for the study with source parameter for each event.

S. No.	Earthquake (EQ)	Lat.	Long.	Mw	Strike	Dip	Focal Depth	Source
1	1897 Shilong EQ	26	91	8.1	292	40	35	Nath et al. [8]
2	1905 Kangra EQ	32.5	76.6	7.8	322	55	18	Singh [6]
3	1934 Bihar Nepal EQ	26.6	86.8	8	100	30	20	Nath et al. [8]
4	1950 Assam EQ	28.38	96.68	8.7	284	45	35	Nath et al. [8]
5	1986 NE India EQ	25.42	92.08	5.4	253	20	43	Singh [6]
6	1986 Dharmasala EQ	32.18	76.29	5.4	299	19	7	CMT Harvard
7	1987 India-Burma EQ	25.27	94.2	5.9	34	32	50	Singh [6]
8	1988 India-Bangladesh	24.64	91.51	6	110	28	15	Singh [6]
9	1988 Manipur EQ	25.15	95.13	7.1	284	45	90	Nath et al. [8]
10	1991 Uttarkashi EQ	30.75	78.86	6.8	317	14	15	CMT Harvard
11	1999 Chamoli EQ	30.41	79.42	6.5	280	7	21	CMT Harvard
12	2005 Kashmir EQ	34.37	73.47	7.6	318	29	15	Raghukanth [47]
13	2005 Chamoli EQ	30.9	79.3	5.4	280	7	25	CMT Harvard
14	2007 Uttarkashi EQ	31.2	78.2	5.3	317	14	33	CMT Harvard
15	2008 Uttarakhand EQ	30.24	80.35	5.1	333	59	10	CMT Harvard
16	2009 Uttarkashi EQ	30.87	79.05	4.7	263	66	52	CMT Harvard
17	2009 Bhutan EQ	27.2	91.62	5.2	293	7	5	CMT Harvard
18	2009 Manipur EQ	25.4	94.8	5.3	261	52	10	CMT Harvard
19	2009 Myanmar EQ	24.25	94.77	5.5	241	62	115	CMT Harvard
20	2009 MIB EQ	24.31	94.84	5.5	224	66	80	CMT Harvard
21	2009 Bhutan EQ	27.2	91.63	6.1	281	6	15	CMT Harvard
22	2010 Tibet EQ	28.41	86.77	5.1	108	75	80	CMT Harvard
23	2010 Myanmar EQ	22.99	94.62	5.6	103	32	114.7	CMT Harvard
24	2010 Hindukush EQ	36.44	70.79	6.2	261	32	207.7	CMT Harvard
25	2011 India Nepal EQ	29.43	80.71	5.4	318	30	18.8	CMT Harvard
26	2011 MIB EQ	24.8	94.6	6.3	143	48	103.5	CMT Harvard
27	2011 Sikkim EQ	27.44	88.35	6.9	216	72	46	CMT Harvard

Table 3a
Parameters used for Historic earthquake.

Parameters	1897 Shillong	1905 Kangra	1934 Bihar Nepal	1950 Assam
Strike	292°	322°	100°	284°
Dip	40°	55°	30°	45°
Depth	35	18	20	35
Fault length	330	100	312	200
Fault width	150	55	80	80
Stress (bars)	159	135	275	66
Attenuation Model	NEH (Table 1)	KGH (Table 1)	BNH (Table 1)	NEH (Table 1)
Duration model	$T_p = \begin{cases} R_h \times 16.8/60, & R_h < 60 \text{ km} \\ 16.8 + 0.05 \times (R_h - 60), & R_h \geq 60 \text{ km} \end{cases}$			
References	Nath et al. [8]	Singh (1987)	Nath et al. [8]	Nath et al. [8]

4. Simulations and GMPE for the Himalayan region at hard-rock level

Initially, 23 recorded (Table 2) and 4 historic earthquakes (marked as Hexagon in Fig. 1, Table 3a) for a given magnitude and from 10 to 750 km as epicentral distance are simulated as per Motazedian and Atkinson [26] and Boore [27] using the parameters given in Tables 2 and 3. The parameters used for the simulation of the historical earthquakes is given as Table 3a. The recorded earthquakes could not cover the entire range of magnitude and distance for developing a new GMPE. Additionally, ground motions were simulated for M_w ranging from 4 to 9, using 0.1 unit step and distance range of 10–750 km. Ranges of stochastic input model parameters used in simulating ground motions are reported in Table 3b.

For simulating ground motions, a random fault orientation is defined i.e. strike and dip, with the dip constrained to be greater than 10, based on the reported dip value for recorded earthquakes. There is no clear evidence of predominant focal mechanism in the entire Himalayan region, except strike-slip and thrust faulting in the eastern part of Himalaya and Nepal region [18,34]. A fictitious fault mechanism is assigned based on the dip angle i.e. for reverse mechanism a

Table 3b
Input parameters used in the simulation.

S. No.	Parameter	Distribution	Mean	Standard Deviation
1	Attenuation Parameter	Normal	Table 1 for corresponding region	
2	Stress Drop (MPa)	Log-Normal	$\ln(\Delta\sigma)_{150} = \begin{cases} 0.36M_w + 0.008, & M_w < 5.5 \\ 150, & M_w \geq 5.5 \end{cases}$ $\ln(\Delta\sigma)_{100} = \begin{cases} 0.32M_w + 0.147, & M_w < 5.5 \\ 100, & M_w \geq 5.5 \end{cases}$ $\ln(\Delta\sigma)_{50} = \begin{cases} 0.25M_w + 0.383, & M_w < 5.5 \\ 50, & M_w \geq 5.5 \end{cases}$	0.2 to 0.5
3	Duration (s)	Normal	$T_p = \begin{cases} R_h \times 16.8/60, & R_h < 60 \text{ km} \\ 16.8 + 0.05 \times (R_h - 60), & R_h \geq 60 \text{ km} \end{cases}$	6.5
4	Focal Depth (km)	Normal	$Depth = \begin{cases} 30 \pm 10, & M_w < 6.0 \\ 50 \pm 10, & M_w \geq 6.0 \end{cases}$	
5	Fault Dimension	Normal	Blaser et al. (2010) for different fault orientation	0.3
6	V_{s30} (m/s)	-	2000	-
7	κ (s)	Normal	0.01	-

Table 4
Regression coefficients for PGA and PSA at different period along with confidence interval calculated using Monte Carlo Simulation.

Period	a_1	a_2	a_3	a_4	a_5	a_6^a	a_7^b	Φ	τ	σ
PGA	1.071 ± 0.25	-0.257 ± 0.051	-0.184 ± 0.043	-0.479 ± 0.051	0.078 ± 0.03	0.076	-0.0085	0.690	0.462	0.817
0.01	1.068 ± 0.24	-0.237 ± 0.051	-0.183 ± 0.043	-0.469 ± 0.051	0.079 ± 0.02	0.077	-0.0082	0.667	0.473	0.839
0.02	1.038 ± 0.24	-0.219 ± 0.051	-0.181 ± 0.044	-0.434 ± 0.051	0.087 ± 0.02	0.078	-0.0083	0.683	0.488	0.869
0.03	1.018 ± 0.25	-0.162 ± 0.055	-0.169 ± 0.051	-0.426 ± 0.052	0.092 ± 0.02	0.078	-0.0084	0.706	0.507	0.896
0.04	1.067 ± 0.25	-0.118 ± 0.051	-0.161 ± 0.052	-0.415 ± 0.051	0.095 ± 0.02	0.080	-0.0085	0.727	0.523	0.918
0.05	1.175 ± 0.24	-0.096 ± 0.051	-0.156 ± 0.052	-0.404 ± 0.052	0.097 ± 0.02	0.081	-0.0085	0.745	0.536	0.933
0.075	1.341 ± 0.25	-0.105 ± 0.056	-0.158 ± 0.053	-0.396 ± 0.051	0.095 ± 0.02	0.079	-0.0086	0.758	0.544	0.932
0.1	1.535 ± 0.26	-0.129 ± 0.050	-0.164 ± 0.053	-0.392 ± 0.053	0.093 ± 0.03	0.077	-0.0086	0.756	0.545	0.926
0.12	1.643 ± 0.26	-0.181 ± 0.052	-0.175 ± 0.054	-0.374 ± 0.053	0.090 ± 0.03	0.073	-0.0086	0.755	0.536	0.904
0.15	1.786 ± 0.26	-0.269 ± 0.052	-0.193 ± 0.053	-0.361 ± 0.054	0.086 ± 0.03	0.071	-0.0086	0.737	0.523	0.880
0.17	1.764 ± 0.26	-0.285 ± 0.052	-0.201 ± 0.053	-0.355 ± 0.051	0.081 ± 0.02	0.063	-0.0086	0.718	0.509	0.854
0.2	1.719 ± 0.26	-0.347 ± 0.053	-0.212 ± 0.053	-0.342 ± 0.051	0.075 ± 0.02	0.061	-0.0086	0.698	0.492	0.831
0.25	1.641 ± 0.25	-0.387 ± 0.050	-0.222 ± 0.052	-0.338 ± 0.05	0.071 ± 0.02	0.057	-0.0085	0.680	0.477	0.814
0.3	1.599 ± 0.26	-0.443 ± 0.055	-0.237 ± 0.051	-0.335 ± 0.051	0.068 ± 0.02	0.056	-0.0084	0.668	0.465	0.798
0.4	1.420 ± 0.26	-0.439 ± 0.055	-0.247 ± 0.054	-0.339 ± 0.054	0.070 ± 0.02	0.055	-0.0082	0.658	0.452	0.786
0.5	1.454 ± 0.26	-0.553 ± 0.055	-0.274 ± 0.055	-0.391 ± 0.054	0.077 ± 0.02	0.077	-0.0079	0.651	0.441	0.775
0.75	1.227 ± 0.28	-0.611 ± 0.050	-0.296 ± 0.055	-0.405 ± 0.054	0.091 ± 0.02	0.087	-0.0076	0.645	0.429	0.769
1	0.823 ± 0.28	-0.619 ± 0.050	-0.307 ± 0.055	-0.422 ± 0.054	0.109 ± 0.02	0.101	-0.0072	0.639	0.427	0.759
1.5	0.217 ± 0.28	-0.587 ± 0.055	-0.311 ± 0.061	-0.435 ± 0.055	0.125 ± 0.02	0.115	-0.0068	0.634	0.417	0.746
2	-0.433 ± 0.29	-0.543 ± 0.060	-0.305 ± 0.058	-0.462 ± 0.055	0.142 ± 0.02	0.131	-0.0065	0.629	0.401	0.744
3	-1.407 ± 0.29	-0.357 ± 0.058	-0.281 ± 0.062	-0.433 ± 0.058	0.156 ± 0.02	0.125	-0.0063	0.624	0.404	0.739
4	-2.325 ± 0.31	-0.103 ± 0.058	-0.243 ± 0.062	-0.463 ± 0.058	0.166 ± 0.02	0.131	-0.0060	0.620	0.402	0.737
5	-3.101 ± 0.31	0.056 ± 0.059	-0.222 ± 0.062	-0.482 ± 0.061	0.173 ± 0.02	0.148	-0.0059	0.618	0.402	0.737
7.5	-3.923 ± 0.33	0.251 ± 0.061	-0.191 ± 0.061	-0.498 ± 0.061	0.181 ± 0.02	0.156	-0.0059	0.616	0.404	0.735
10	-4.426 ± 0.33	0.085 ± 0.060	-0.203 ± 0.061	-0.515 ± 0.063	0.187 ± 0.02	0.197	-0.0060	0.614	0.405	0.817

^a Confidence interval in a_6 is ± 0.02 for all periods.

^b Confidence interval in a_7 is ± 0.001 for all periods.

dip lower than 40°, for strike-slip mechanism dip greater than 75° and normal mechanism, otherwise [4].

Another important factor in the synthetic simulation of ground motions is the focal depth which is estimated based on the updated catalog of earthquake events for the entire Himalayan region. On studying the focal depth for all the four regions, it was observed that for events with $M_w > 6.0$, the depth varies from 5 to 40 km and for $M_w < 6.0$, the depth varies from 1.1 to 290 km. Most of the recorded events with a hypocentral depth more than 60 km have an estimation error of about ± 20 km (USGS). Only for few events, $M_w > 6.0$, focal depth of more than 60 km is available, which is not considered in the present study. To capture the variability in the focal depth for the simulated earthquakes, a normal distribution is assumed with mean and standard deviation as 30 ± 10 for $M_w > 6.0$ and 50 ± 10 for $M_w < 6.0$ is considered.

Fault dimensions (rupture length and rupture width) are estimated from the Blaser et al. [35] relationships between fault dimensions and M_w . To include the additional variability in the fault dimensions, for each simulation of ground motion, the input M_w is sampled using a normal distribution with standard deviation of 0.3 [4]. This allows

simulating fault plane with corresponding orientation of different dimensions for same magnitude. Blaser et al. [35] relationships are valid for $M_w > 4.8$. Hence, the equation [35] were extrapolated beyond their validity range, especially for small magnitudes. The extrapolation did not have a strong influence on the simulations because of very small fault extension for small magnitudes [4]. Fifty different rupture orientation are simulated for each moment magnitude with bin size 0.1 M_w .

Considering the given focal mechanism and magnitude, epicentral distance from 10 to 750 km and source-to-site azimuths from 0° to 360° are simulated. The maximum epicentral distance is selected as 750 km by studying the damage distribution map of the pre-instrumented earthquake (1934 Bihar Nepal, 1950 Assam earthquake) and recent 2015 Nepal earthquake. Past earthquake events have shown that the Himalayan region is experiencing damages in an area covering more than 600 km radius around the epicenter. Hence, a GMPE needs to be developed for a large range of hypocentral distance. The concept of apparent station (AS) has been used for simulating each event at different hypocentral distance [14]. The locations of AS have been selected in such a way that the entire hypocentral distance of 750 km can

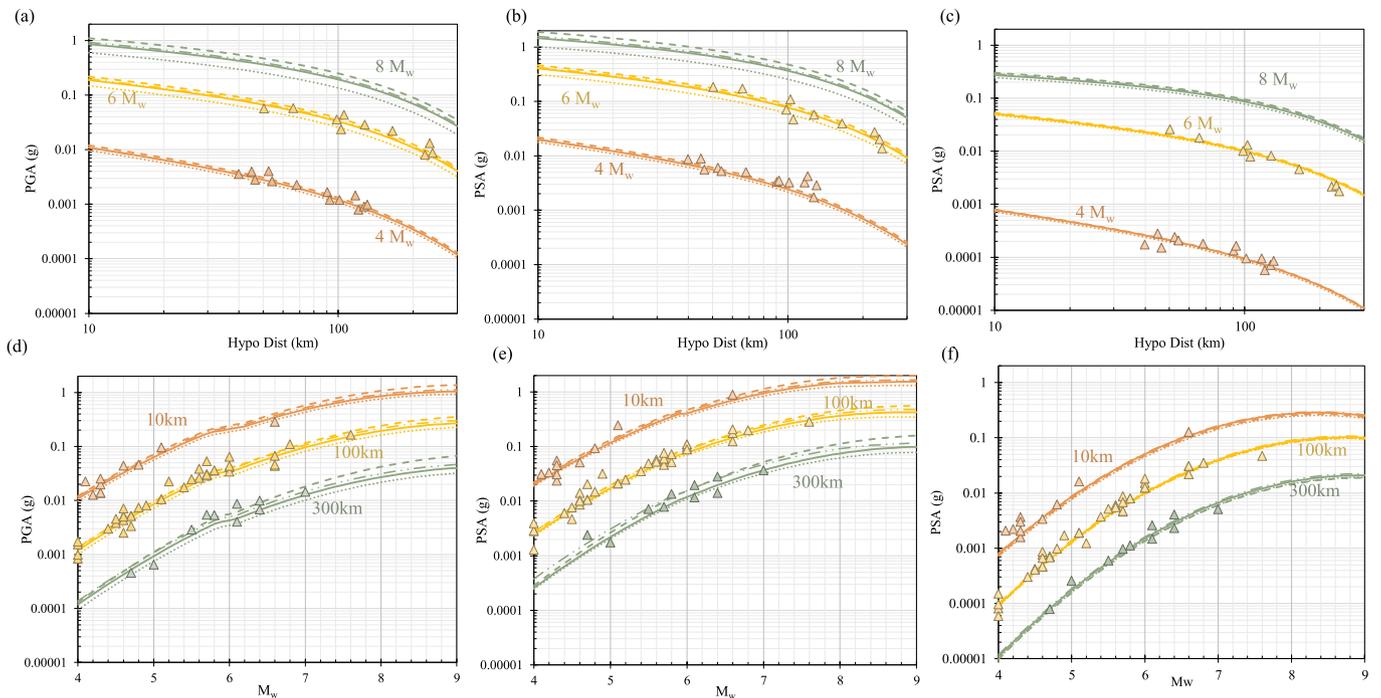


Fig. 4. Comparison of different version of stochastic ground motion prediction equation, amplitude versus distance at (a) PGA, (b) PSA at 0.2 s and (c) PSA at 1.0 s and amplitude versus magnitude (d) PGA (e) PSA at 1.0 s and (f) PSA at 1.0 s. The dotted, dashed-dotted, dashed line and solid line respectively represent the variable stress at 5, 15 MPa and the constant stress at 10 MPa. The solid triangle represents corresponding recorded data.

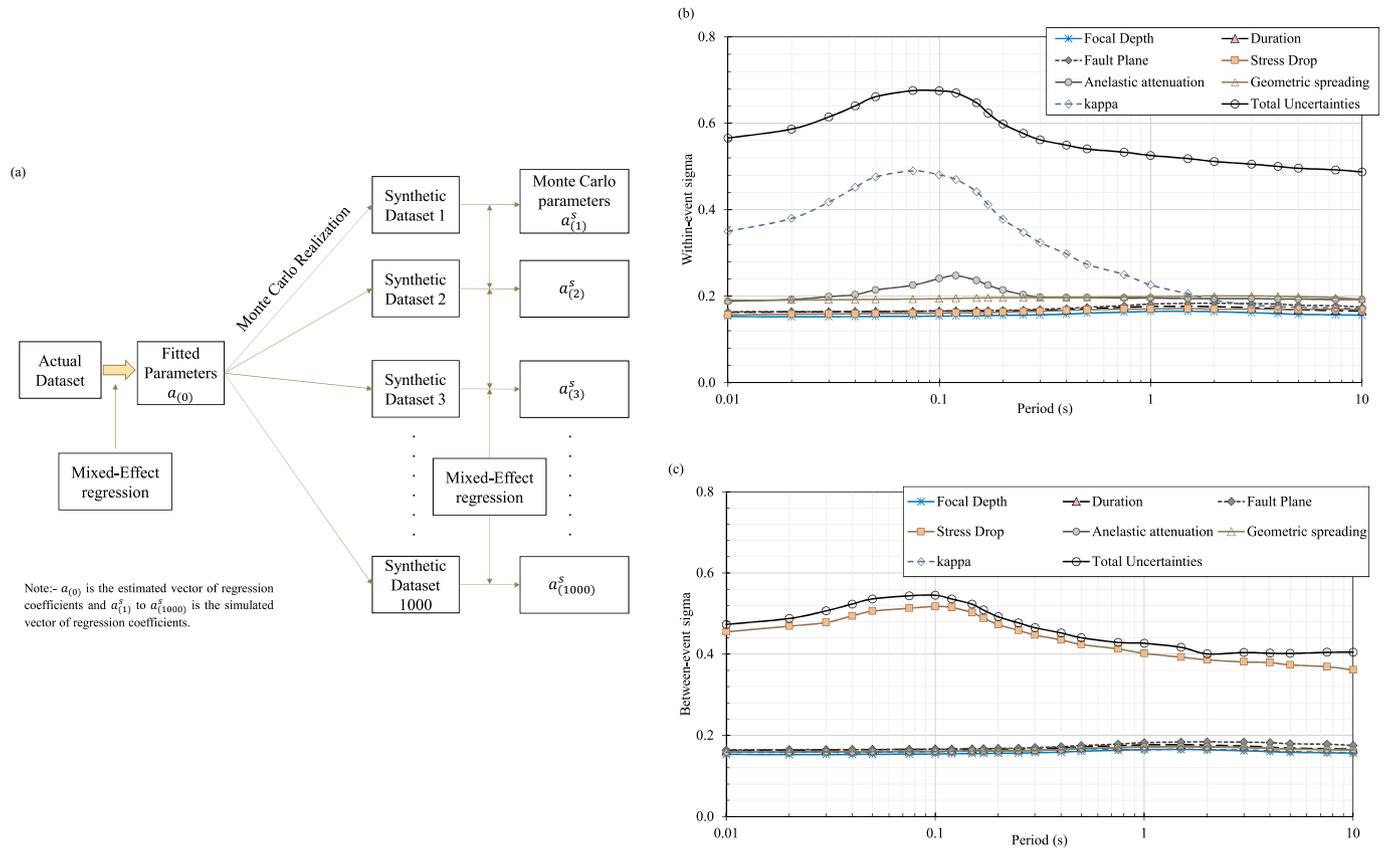


Fig. 5. (a) Monte Carlo scheme used for determining the confidence interval and standard error in regression parameters. (b) Within-event and (c) Between-event sigma obtained from considering eight models used in sensitivity analysis; same notions has been used in Fig. 5 (a) and (b).

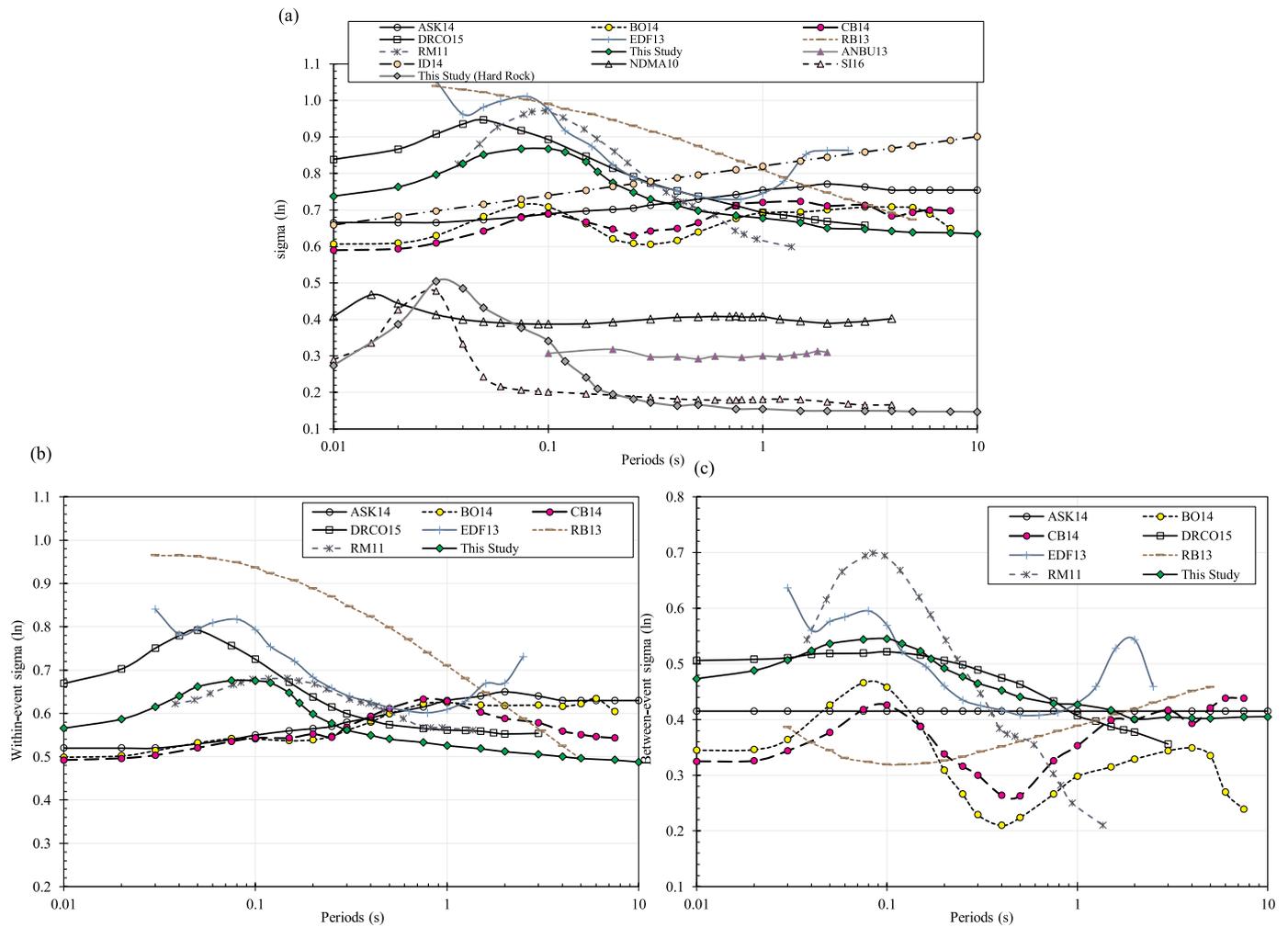


Fig. 6. Comparison of (a) Sigma; (b) Within-event and (c) Between-event standard deviations for the new Himalayan region GMPE with the NGA-West 2 GMPEs and other stochastic GMPEs. The abbreviation of the respective GMPE is given in the text.

be captured. The ASs are established at 75 locations with azimuths covering the range of 0° to 360° around the epicenter with azimuth bin size of 4.8°. Hence, every next AS is at 10 km with an azimuth difference of 4.8° from the successive stations. For example, the first station is at 10 km with azimuth at 4.8°, the next station will be at 20 km with azimuth at 9.6° and last station will be at 750 km with azimuth at 360°. The pictorial representation of AS is shown in Fig. 1 (top). This AS is further used for simulating the ground motion for the different earthquake conditions, for the development of a new GMPE. Further all the simulated parameters are used for developing a new GMPE for the Himalayan region.

5. Regression analysis

The GMPEs relate the specific strong motion parameter of ground shaking to seismological parameter of earthquake as source, site, wave propagation path between the site and source, and geological features beneath the site. Wide varieties of models are available with different functional forms corresponding to the distance and magnitude scaling for deriving any new GMPE. However, a functional form of GMPE can be selected to any region before deriving a new robust GMPE considering available recorded data accounting the scaling corresponding to the distance and magnitude. Bajaj and Anbazhagan [36] tested the compatibility of various functional forms for distance and magnitude scaling using the mixed-effect regression of residual calculated from the functional form given in NGA-West 2 project. Based on that, within the

event and between events, uncertainties have been evaluated concerning distance and magnitude. The whole algorithm and different functional form used is explained in Bajaj and Anbazhagan [36] and best representing GMPE functional form for Himalayan region based on available data is given below:

$$\ln Y = a_1 + a_2(M - 6) + a_3(9 - M)^2 + a_4 \ln R + a_m \ln R(M - 6) + a_7 R + \sigma \quad (12)$$

Where $\ln Y$, M , R and σ are respectively logarithm of ground motion, magnitude, hypocentral distance and standard deviation and a_1 , a_2 , a_3 , a_4 , a_m and a_7 are the corresponding regression coefficients. The coefficient a_m is equal to a_3 when $M_w < 6.0$ and $R < 300$, else is equal to a_6 . For determining the regression coefficients for a new GMPE, standard nonlinear least square (LS) and random effect maximum likelihood method (ML) is used for combined recorded and simulated database. The main advantage of ML technique is that, it considers systematic differences between events through partitioning the residuals [4,37]. For the approximately even distribution of the simulated data for magnitude and distance, both ML and LS generate the same results [4]. However, ML is used as it results in both between-event and within-event uncertainty. The regression coefficients corresponding to a_1 , a_1 , a_3 , a_4 , a_5 , a_6 , a_7 and a_8 for different periods are given in Table 4.

Fig. 4 compares the various versions of the derived GMPE for the Himalayan region, i.e. the variable stress drop with hinge at 5 MPa, 10 MPa and 15 MPa (see Fig. 2) and constant stress drop at 10 MPa for a standard rock condition for all magnitudes. Comparing all GMPEs for

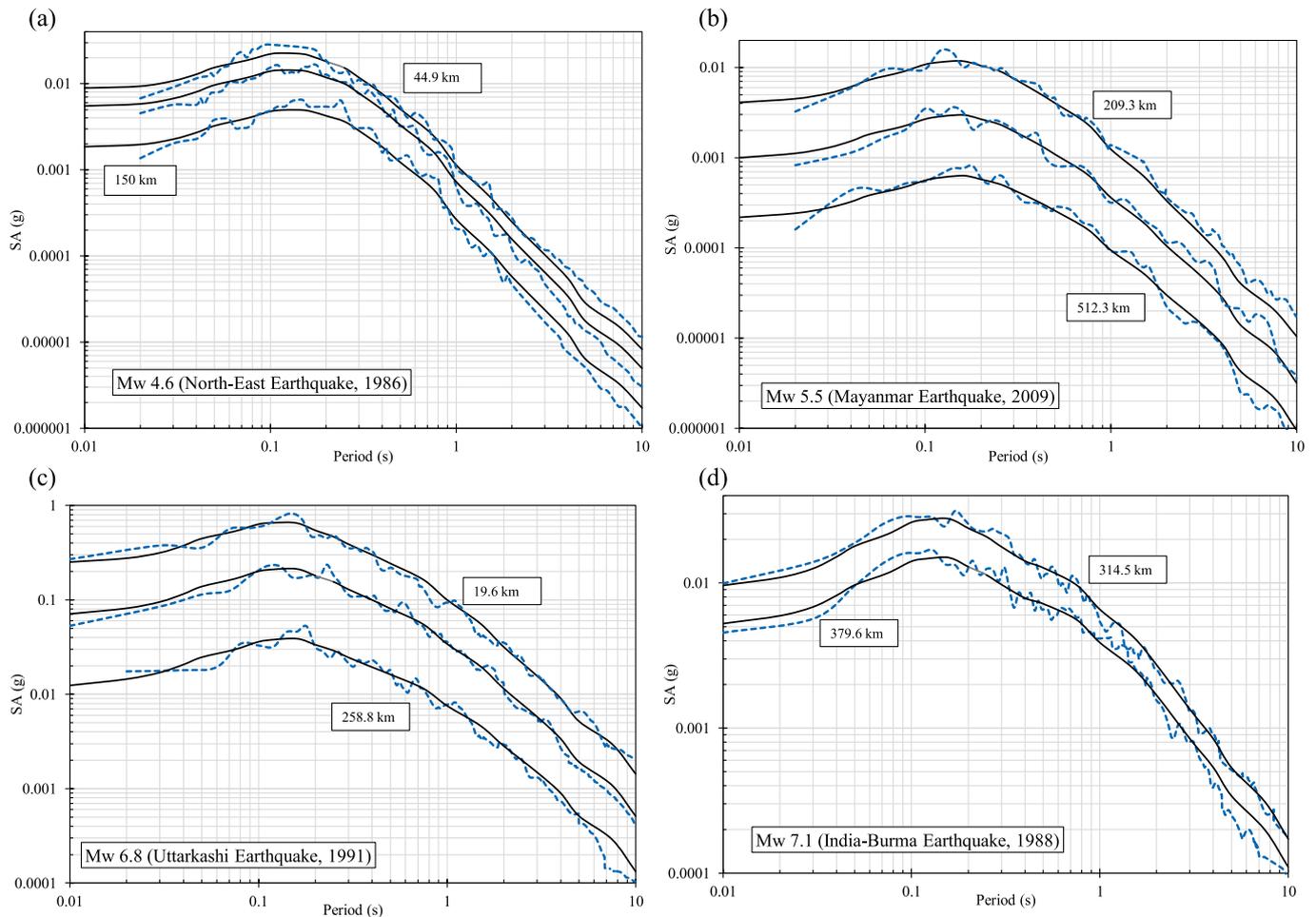


Fig. 7. Comparison of predicted and recorded response spectra (a) North-East EQ (1986); (b) Myanmar EQ (2009); (c) Uttarkashi EQ (1991) and (d) India-Burma EQ (1988) at different hypocentral distance. The solid and square dotted line respectively represents the predicted and recorded response spectra.

variable and constant stress drop, it could be seen that for the large magnitude the variation is significant, as this is the similarity between the constant and the variable stress drop. The differences are more pronounced for shorter periods as compared to longer periods. The PGA and PSA (0.2 s) values estimated from a new GMPE for different stress model are also compared with the recorded data. It is seen that the recorded data is consistent with both variable stress and constant stress drop.

6. Numerical experimentation on new GMPE

The standard error and confidence intervals are evaluated using the Monte Carlo simulation technique [38] to test the accuracy of the estimated parameters from regression. 1000 samples of the PSA were simulated using equation (12) for the given estimated regression coefficients for different sets of magnitude and hypocentral distance for this purpose. This database was simulated by bootstrapping the residuals [39]. The simulations were performed at different time periods for which regression coefficients were evaluated. For each of the simulated PSA dataset, a crossed and nested, mixed effect regression approach was used to find the regression coefficient corresponding to each of the dependent variables in equation (2). As per Efron and Tibshirani [39], the standard error of the estimated parameters is the standard deviation of the estimated parameters of the simulated database. For example, the standard error (\hat{s}_e) in regression coefficients for PSA at 0.1s for $a_1, a_2, a_3, a_4, a_5, a_6, a_7$ and a_8 is 0.21, 0.22, 0.05, 0.04, 0.02, 0.02, 0.01 and 0.00 respectively. The whole procedure is given in Fig. 5 (a), in a form of schematic diagram.

There are various methods for evaluating the bootstrap confidence interval; however, the simplest one is the standard bootstrap confidence interval that works when the estimates are normally distributed. The $100(1 - \alpha)\%$ standard bootstrap confidence interval on the estimated parameter is defined by Efron and Tibshirani [39] as $Estimate \pm z_{\alpha/2} \cdot \hat{s}_e$, where, $z_{\alpha/2}$ is the upper $100\alpha/2$ percentage point of the standard normal distribution and \hat{s}_e is the bootstrap estimate of the standard error.

The distribution of the regression coefficients is close to normal distribution; hence the above method is used for estimating the coincidence intervals. For example, the 90% confidence interval for PSA at 0.1s for $a_1, a_2, a_3, a_4, a_5, a_6, a_7$ and a_8 respectively is $1.757 \pm 0.25, 1.535 \pm 0.26, -0.129 \pm 0.050, -0.164 \pm 0.053, -0.392 \pm 0.053, 0.125 \pm 0.03, 0.077 \pm 0.02$ and -0.0086 ± 0.001 . The confidence interval for the respective period corresponding to each regression coefficient is given in Table 4.

7. Sensitivity analysis

In this study, the developed GMPE is based on the simulated ground motions which are highly influenced by the uncertainties in some of the input parameters. Hence, sensitivity analysis has been performed to test how these uncertainties influence the final GMPE uncertainty and how robustly the present GMPE models result in the presence of uncertainty. Set of seven GMPEs were computed; considering the uncertainty on each parameter, one by one (all others were set to their median values). The considered uncertainties are; focal depth uncertainty, uncertainty on duration, fault plane (random orientation and hypocentral position),

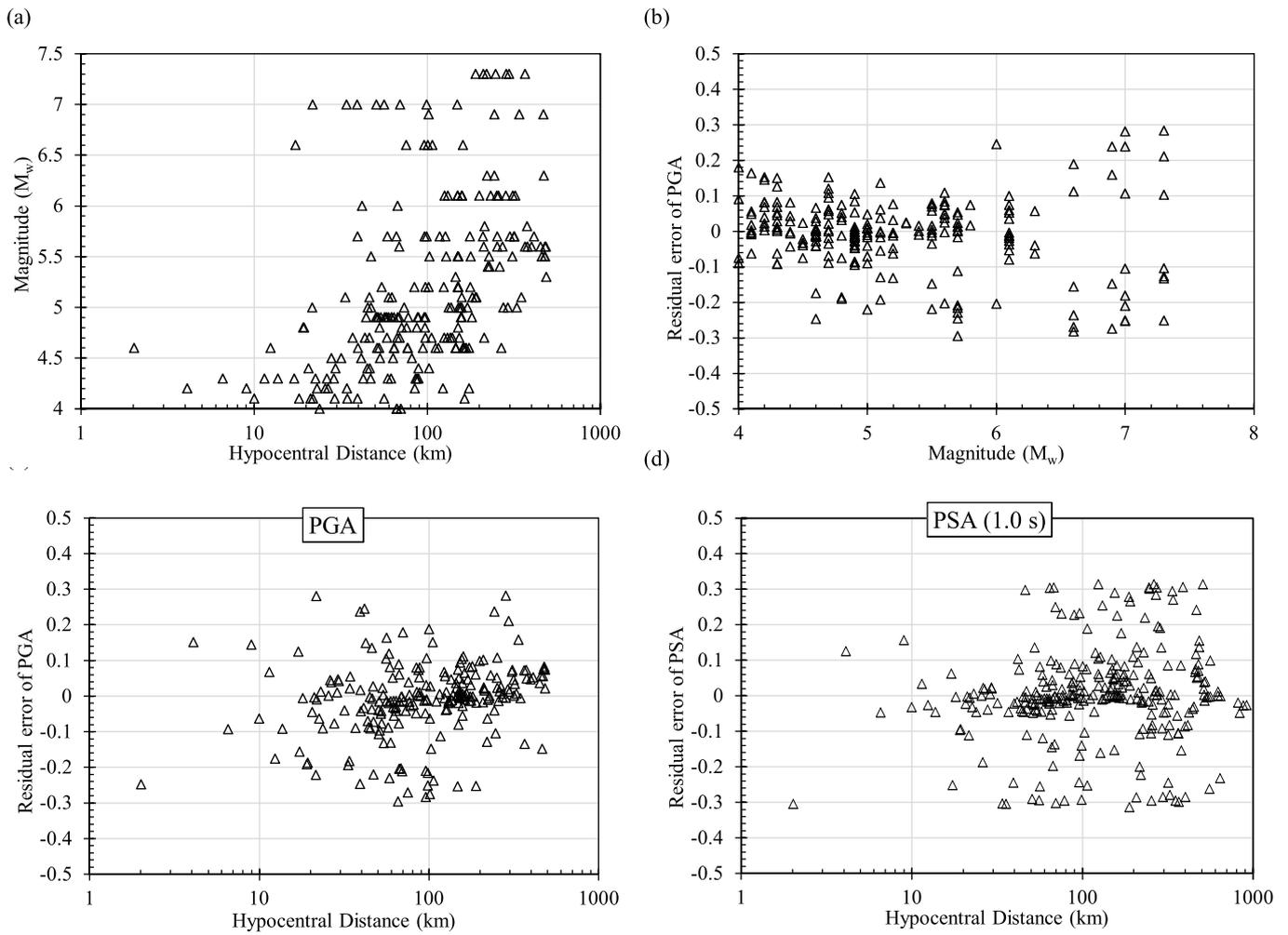


Fig. 8. (a) Magnitude versus Distance plot and distribution of residual (b) PGA versus Magnitude; (c) PGA versus Hypocentral Distance; (d) PSA (1.0s) versus Hypocentral Distance.

Table 5
Summary of GMPEs developed for the Himalayan regions.

Sl. No.	Reference	Database	(Mw)	Distance Range (Km)
1.	Singh et al. [6]	Recorded ground motions from 5 earthquakes	5.7-7.2	≤100
2.	Sharma [7]	66 recorded data from 6 earthquakes	5.5-6.8	≤150
3.	Nath et al. [46]	80 recorded events and 25 simulated events	3.0-8.5	≤100
4.	Das et al. [10]	261 recorded data from 6 earthquakes	5.5-7.2	≤300
5.	Sharma and Bungum [9]	Combined dataset consisting of 14 recorded earthquakes from India and 9 recorded earthquakes from Europe	4.6-7.6	≤200
6.	Baruah et al. [11]	82 recorded earthquakes at 8 broadband stations	2.5-5.0	≤145
7.	Nath et al. [8]	Simulated ground motions	4.8-8.1	≤100
8.	Sharma et al. [12]	Combined dataset consisting of 6 recorded earthquakes from India and 10 recorded earthquakes from Zegros region	5.2-6.9	≤100
9.	Gupta [48]	56 recorded data from 3 events	6.3-7.2	150-375
10.	NDMA [15]	1600 Simulated ground motions	4-8.5	≤500
11.	Anbazhagan et al. [14]	Simulated and recorded data of 14 earthquakes	5.3-8.7	≤300
12.	New GMPE	Simulated 5775 sets for 10,000 earthquakes	4.0-9.0	≤750

anelastic attenuation, geometric spreading, kappa factor and stress drop. Based on the analysis, it is seen that the coefficients determined by varying these parameters are almost identical and within the confidence interval as mentioned in Table 4. However, the standard deviation varies a lot with the change in each parameter (See Fig. 5 (b) and (c)). Major contribution to the total uncertainty in case of within-event terms is from κ (Fig. 5 (b)) and from stress drop in case of event-to-event variability (See Fig. 5 (c)). In addition to κ , anelastic attenuation has also peaked above 0.1s but its influence is similar to

geometric attenuation at longer periods. Comparatively, focal depth, duration, fault plane uncertainty and random variation of hypocentral location on the fault have very little impact on the total GMPE uncertainty.

The standard deviation determined in this study is compared with the standard deviation of various GMPEs that are empirical GMPE for NGA-West 2 project and stochastic for the United Kingdoms, Alps, Switzerland and the Himalayan region and given in Fig. 6. Most of the stochastic models show large deviation at low period and decreases

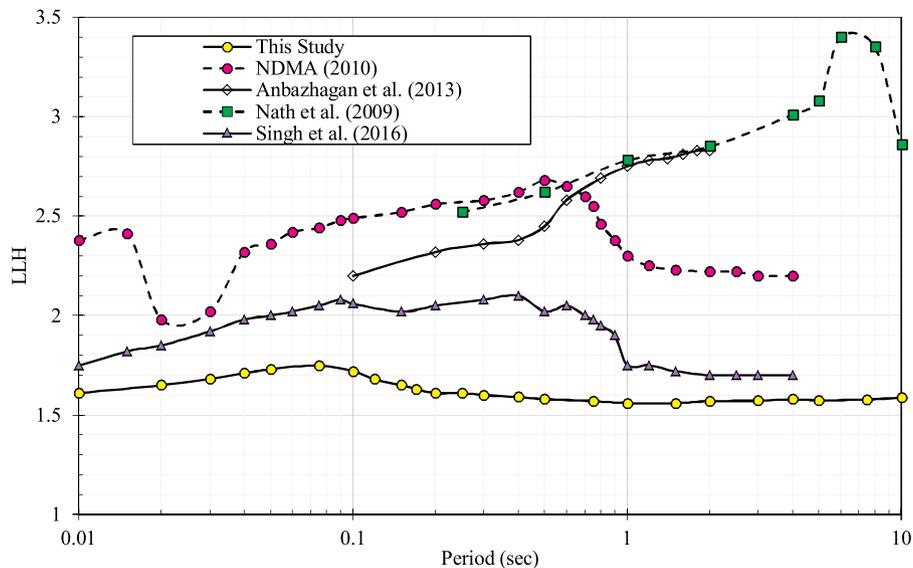


Fig. 9. LLH value calculated considering the Himalayan GMPEs at different period for the recorded data up to 2015.

with respect to periods. Boore et al. [40] and Campbell and Bozorgnia [41] (referred as BO14 and CB14) respectively, show the standard deviation as a dual peak i.e. between 0.05 and 0.1 s and after 1 s. The peak of the Edwards and Fäh [3] and Drouet and Cotton [4] (referred as EDF13 and DRCO15) and the present study is located between 0.05 s and 0.1 s. In case of NGA-West 2 GMPE, within-event sigma the peak lies between 0.5 s and 1.5 s. For all the derived Himalayan GMPEs as NDMA [15], Anbazhagan et al. [14] and Singh et al. [5] (referred as ANBU13 and SI16) the peak lies around or before 0.03s (See Fig. 6 (a)). In Fig. 6, equation Abrahamson et al. [42], Rietbrock [43], Rodriguez-Marek et al. [44] and Idriss [45] referred as AKS14, RB13, RM11 and ID14 respectively.

8. Validation and comparison of the new GMPE

The available strong motion recorded data in Himalayan region does not cover the entire range of magnitude and distance. Hence, a detailed validation of the newly derived ground motion prediction equation for all ranges of magnitude and distance is not possible. However, from Fig. 4, it is already seen that the recorded PGA and PSA (0.2 s) is matching well with the PGA and PSA calculated from the newly developed GMPE. Further, the response spectra of the North-East Earthquake (1986); Myanmar Earthquake (2009); Uttarkashi Earthquake (1991) and India-Burma Earthquake (1988), which is recorded at rock site is compared with the predicted response spectra from GMPE and given in Fig. 7. It is seen from Fig. 7 that the predicted and derived response spectra at different hypocentral distance and magnitude is matching with \pm one standard deviation with respect to mean for the entire period range. Minor differences at different periods may arise due to different uncertainties considered while simulating the stochastic strong ground motion.

Further, the error has been determined for PGA and PSA at 0.2 and 1 s with respect to magnitude and hypocentral distance. For a regression model to be unbiased, the mean of the residual is to be zero and independent of the parameters in the regression model. Further for checking the bias and average scatter, average and standard deviation of the residual's errors are also calculated. Based on Fig. 8 (a), PGA distribution of residuals with magnitude is unbiased. Similarly, the calculated residuals are plotted with respect to hypocentral distance at PSA for a period of 1.0s. From Fig. 8(b and c), it is also clear that the residuals are unbiased with respect to the hypocentral distance. This study on residuals indicates that the proposed GMPE model is unbiased

with respect to magnitude and hypocentral distance.

Further the new GMPE is compared with the existing Himalayan region GMPE using the Log-Likelihood (LLH) method for different periods. The detail of each of the existing GMPE with the new GMPE for the Himalayan region is given as Table 5. The whole procedure regarding LLH is explained in Delavaud et al. [49]. The LLH values are computed using the recorded strong motion database up to 2015. The GMPE used for the comparison are Nath et al. [8], NDMA [15], Anbazhagan et al. [14] and Singh et al. [5]. Fig. 9 shows the comparison of the GMPEs, in which LLH is plotted against the different structural period. LLH value less than 1 is obtained for all the structural periods. It is also observed that low value of LLH was obtained for PGA, but in all the existing GMPEs, it is increasing with the period.

9. Conclusion

Stochastic ground motion prediction equation for the whole Himalayan region for a broad range of magnitudes (M_w 4–9) and hypocentral distances (10–750 km) was derived. The simulations were done by deriving the region-specific seismological model parameters and the uncertainty was propagated by random sampling of a corresponding distribution of the input parameters. These seismological model parameters were studied by dividing the whole Himalayan region into four parts i.e. Kashmir Himalayan, Kumaon-Garhwal Himalayan, Bihar-Nepal Himalaya and Northeastern part of Himalayan region. It was observed that these parameters are within the uncertainty of each other for the entire Himalayan region. Hence only one GMPE model was derived for the Himalayan region by adding uncertainty in the seismological model parameters. The simulated data considering these parameters was regressed using cross and nested regression approach for the defined functional form for scaling the magnitude and large distance. Monte Carlo approach was used to determine the confidence interval and standard error in regression coefficients. Further, sensitivity analysis was used to find the impact of each input parameter on the total uncertainty of the ground motion prediction equation. The defined GMPE is valid for rock level ($V_{s30} = 2000$ m/s, $\kappa = 0.01$ s). Site effect based on different seismic site classification is not considered while deriving a GMPE, as the data corresponding to different seismic site class are not extensively defined. Hence, this could be a part of future study. PGA and SA of newly developed GMPE are matching well with the recorded data for the larger and smaller hypocentral distances.

Acknowledgment

The authors would like to thank Dr. David Boore and Dr. Stéphane Drouet and two anonymous reviewers for their valuable comments during the analysis, simulation and preparing the manuscript. The authors wish to express gratitude to the Science and Engineering Research Board (SERB) of the Department of Science and Technology (DST), India for funding the project titled "Measurement of shear wave velocity at deep soil sites and site response studies", Ref: SERB/F/162/2015–2016. Authors also would thank "Board of Research in Nuclear Sciences (BRNS)", Department of Atomic Energy (DAE), Government of India for funding the project titled "Probabilistic seismic hazard analysis of Vizag and Tarapur considering regional uncertainties" (Ref No. Sanction No 36(2)/14/16/2016-BRNS).

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.soildyn.2019.105825>.

References

- [1] Bilham R, Gaur VK, Molnar P. Himalayan seismic hazard. *Science* 2001;293:1442–4.
- [2] Bilham R. Raising Kathmandu. *Nat. Geosci* 2015;8:582–4.
- [3] Edwards B, Fäh D. A stochastic ground-motion model for Switzerland. *Bull Seismol Soc Am* 2013;103:78–98.
- [4] Drouet S, Cotton F. Regional stochastic GMPEs in low-seismicity areas: scaling and Aleatory variability analysis-application to the French Alps. *Bull Seismol Soc Am* 2015;105:1883–902.
- [5] Singh NM, Rahman T, Wong IG. A new ground motion prediction equation for Northeastern India (NEI) crustal earthquakes. 2016;105:981–92. <https://doi.org/10.1785/0120150180>.
- [6] Singh RP, Aman A, Prasad YJJ. Attenuation relations for strong ground motion in the Himalayan region. *Pure Appl Geophys* 1996;147:161–80.
- [7] Sharma ML. Attenuation relationship for estimation of peak ground horizontal acceleration using data from strong motions arrays in India. *Bull Seismol Soc Am* 1998;88:1063–9.
- [8] Nath SK, Raj A, Thingbaijam KKS, Kumar A. Ground motion synthesis and seismic scenario in Guwahati city; a stochastic approach. *Seismol Res Lett* 2009;80(2):233–42.
- [9] Sharma ML, Bungum H. New strong ground motion spectral acceleration relation for the Himalayan region. *First European conference on earthquake engineering and seismology*. 2006. p. 1459.
- [10] Das S, Gupta ID, Gupta VK. A probabilistic seismic hazard analysis of Northeast India. *Earthq Spectra* 2006;22:1.
- [11] Baruah S, Gogoi NK, Erteleva Q, Aptikaev F, Kayal JR. Ground motion parameters of Shillong plateau: one of the most seismically active zones of Northeastern India. *Earthq Sci* 2009;22:283–91.
- [12] Sharma ML, Douglas J, Bungum H, Kotadia J. Ground-motion prediction equations based on data from Himalayan and Zagros regions. *J Earthq Eng* 2009;13:1191–210.
- [13] Gupta ID. Response spectral attenuation relations for inslab earthquakes in Indo-Burmese subduction zone. *Soil Dyn Earthq Eng* 2010;30:368–77.
- [14] Anbazhagan P, Kumar A, Sitharam TG. Ground motion prediction equation considering combined dataset of recorded and simulated ground motions. *Soil Dyn Earthq Eng* 2013;53:92–108. 2013.
- [15] National Disaster management authority (NDMA). *Development of probabilistic hazard of India*. India: Government of India; 2010.
- [16] Seeber H, Gornitz V. River Flow along the Himalayan arc as indicator of active tectonics. *Tectonophysics* 1983;92:355–67.
- [17] Khattri KN. Great earthquakes, seismicity gaps and potential to earthquake disaster along the Himalayan plate boundary. *Tectonophysics* 1987;138:92–9.
- [18] Torre TL de la, Monsalve G, et al. Earthquake processes of the Himalayan collision zone in eastern Nepal and the southern Tibetan Plateau. *Geophys J Int* 2007;171:718–38.
- [19] Srivastava HN, Verma M, Bansal BK, Sutar AK. Discriminatory characteristics of seismic gaps in Himalaya. *Geo Nat Haz Risk* 2015;6:224–42.
- [20] Boore DM, Bommer J. Processing of strong motion accelerograms: needs, options and consequences. *Soil Dyn Earthq Eng* 2005;25:93–115.
- [21] Joshi A, Kumar P, et al. Determination of $Q_b(f)$ in different parts of Kumaon Himalaya from the inversion of spectral acceleration data. *Pure Appl Geophys* 2012;169:1821–45.
- [22] Boore DM. Simulation of ground motion using the stochastic method. *Pure Appl Geophys* 2003;160:635–76.
- [23] Aki K, Chouet B. Origin of coda waves: source, attenuation, and scattering effects. *J Geophys Res* 1975;80:3322–42.
- [24] Anderson JG, Hough SE. A model for the Shape of the Fourier amplitude spectrum of acceleration at high frequencies. *Bull Seismol Soc Am* 1984;74:1969–93.
- [25] Beresnev I, Atkinson GM. Modeling finite fault radiation from wn spectrum. *Bull Seismol Soc Am* 1997;87:67–84.
- [26] Motazedian D, Atkinson GM. Stochastic finite-fault modeling based on a dynamic corner frequency. *Bull Seismol Soc Am* 2005;95:995–1010. 2005.
- [27] Boore DM. Comparing stochastic point-source and finite-source ground motion simulations: SMSIM and EXSIM. *Bull Seismol Soc Am* 2009;99:3202–16.
- [28] Zandieh A, Pezeshk S. Investigation of geometric spreading and quality factor in the New Madrid seismic zone. *Bull Seismol Soc Am* 2010;100:2185–95.
- [29] Bajaj K, Anbazhagan P, Moustafa SSR, Al-Arifi NSN. Attenuation characterization of the Himalayan region using spectral analysis of seismic waves. *Geophys J Int* 2019. [Under Review].
- [30] Ide S, Beroza GC, Prejean SG, Ellsworth WL. Apparent break in earthquake scaling due to path and site effects on deep borehole recordings. *Geophys Res Lett* 2003;108:2271.
- [31] Boore DM, Campbell KW, Atkinson GM. Determination of stress parameters for eight well-recorded earthquakes in eastern North America. *Bull. Seism. Soc. Am.* 2010;100:1632–45. 2010.
- [32] Allmann BP, Shearer PM. Global variations of stress drop for moderate to large earthquakes. *J Geophys Res* 2009;114:B01310. <https://doi.org/10.1029/2008JB005821>.
- [33] Boore DM, Thompson EM. Path durations for use in the stochastic-method simulation of ground motions. *Bull Seismol Soc Am* 2014;104:2541–52.
- [34] Rastogi BK. Earthquake focal mechanism and plate tectonics in the Himalayan region. *Tectonophysics* 1974;21:47–56.
- [35] Blaser L, Krüger F, Ohrnberger M, Scherbaum F. Scaling relation of earthquake source parameter estimates with special focus on Subduction environment. *Bull Seismol Soc Am* 2010;100:2914–26. 2010.
- [36] Bajaj K, Anbazhagan P. Determination of GMPE functional form for an active region with limited strong motion data: application to the Himalayan region. *J Seismol* 2018. <https://doi.org/10.1007/s10950-017-9698-5>.
- [37] Stafford PJ. Crossed and nested mixed-effects approaches for enhanced model development and removal of the ergodic assumption in empirical ground-motion models. *Bull Seismol Soc Am* 2014;104:702–19.
- [38] Press WH, Teukolsky SA, Vetterling WT, Flannery BP. *Numerical Recipes*. Cambridge: Cambridge University Press; 2007.
- [39] Efron B, Tibshirani R. *An introduction to the bootstrap*. New York: Chapman & Hall; 1993.
- [40] Boore DM, Stewart JP, Seyhan E, Atkinson GM. NGA-West2 equations for predicting PGA, PGV, and 5%-damped PSA for shallow crustal earthquakes. *Earthq Spectra* 2014;30(3):1057–85.
- [41] Campbell KW, Bozorgnia Y. NGA-West2 ground motion model for the average horizontal components of PGA, PGV, and 5%-damped linear acceleration response spectra. *Earthq Spectra* 2014;30(3):1087–115.
- [42] Abrahamson NA, Silva WJ, Kamai R. Summary of the ASK14 ground-motion relation for active crustal regions. *Earthq Spectra* 2014;30(3):1025–55.
- [43] Rietbrock A, Strasser FO, Edwards B. A stochastic earthquake ground-motion prediction model for the United Kingdom. *Bull Seismol Soc Am* 2013;103:57–77.
- [44] Rodriguez-Marek A, Montalva GA, Cotton F, Bonilla F. Analysis of single-station standard deviation using the KiK-net data. *Bull Seismol Soc Am* 2011;101:1242–58.
- [45] Idriss IM. An NGA-West2 empirical model for estimating the horizontal spectral values generated by shallow crustal earthquakes. *Earthq Spectra* 2014;30(3):1155–77. 2014.
- [46] Nath SK, Vyas M, Pal I, Sengupta P. A hazard scenario in the Sikkim Himalaya from seismotectonics spectral amplification source parameterization and spectral attenuation laws using strong motion seismometry. *J Geophys Res* 2005;110:1–24.
- [47] Raghukanth STG. Ground motion estimation during the Kashmir earthquake of 8th October 2005. *Nat Hazards* 2008;46:1.
- [48] Gupta ID. Response spectral attenuation relations for in slab earthquakes in Indo-Burmese subduction zone. *Soil Dyn Earthq Eng* 2010;30:368–77.
- [49] Delavaud E, Scherbaum F, Kühn N, Riggelsen C. Information-theoretic selection of ground-motion prediction equations for seismic hazard analysis: an applicability study using Californian data. *Bull Seismol Soc Am* 2009;99:3248–63.
- [50] Kumar A, Mittal H, Sachdeva R, Kumar A. Indian strong motion instrumentation network. *Seismol Res Lett* 2012;83:59–66.
- [51] Andrews DJ. Objective determination of source parameters and similarity of earthquakes of different size. Das S, Boatwright J, Scholz SH, editors. *Earthquake source mechanics* Washington DC: Maurice Ewing Series 6, AGU; 1986. p. 259–67.
- [52] Anbazhagan P, Srilakshmi KN, Bajaj K, Moustafa SSR, Al-Arifi NSN. Determination of seismic site classification of seismic recording stations in the Himalayan region using HVSR method. *Soil Dyn Earthq Eng* 2019;304–16.
- [53] Bajaj K and Anbazhagan P. "Comprehensive amplification estimation of the Indo Gangetic basin deep soil sites in the seismically active area" soil dynamics and earthquake Engineering (under Review).