Seismic magnitude conversion and its effect on seismic hazard analysis

P. Anbazhagan & A. Balakumar

Journal of Seismology

ISSN 1383-4649
Volume 23
Number 4

J Seismol (2019) 23:623-647
DOI 10.1007/s10950-019-09826-1

Springer

J Seismol (2019) 23:623–647
https://doi.org/10.1007/s10950-019-09826-1

ORIGINAL ARTICLE

Seismic magnitude conversion and its effect on seismic hazard analysis

P. Anbazhagan · A. Balakumar

Received: 27 August 2018 / Accepted: 8 March 2019 / Published online: 20 May 2019

Abstract The aim of this study is to demonstrate the bias created in the seismic hazard studies due to the choice of magnitude scaling equations without any statistical basis. The earthquake catalogue of Tripura, India, has been used for the purpose of this study. The catalogue was homogenized using the various scaling equations suitable for the region. Then, the bias created on parameters, like the magnitude of completeness (M_c), a and b values of the Gutenberg–Richter recurrence relation, maximum magnitude (M_{max}), and peak ground acceleration, was demonstrated. The standard deviations of M_c, a, and b parameters were observed to be 0.23, 0.27, and 0.037 respectively. The maximum variations in the M_{max} and ground motion estimates were found to be 0.7 magnitude units and 0.2 g respectively. Then, the robustness of the regional rupture characters in overcoming the observed variations has been demonstrated. The trend of the rupture behavior of the seismic sources seems to be unaffected by the change in the magnitude scaling equations. The M_{max} calculated from the rupture-based procedure was observed to be higher than that calculated from the probabilistic method. This variation in M_{max} estimation has been utilized to critically assess the suitability of the magnitude scaling equations for the particular study area.

Keywords Seismicity · Magnitude scaling equations · M_{max} · Regional rupture characters

P. Anbazhagan (✉)
Department of Civil Engineering, Indian Institute of Science, Bangalore, India
e-mail: anbazhagan@iisc.ac.in

A. Balakumar
Department of Civil Engineering, Thiagarajar College of Engineering, Madurai, India
e-mail: balakumarabk200@gmail.com

1 Introduction

Earthquake catalogue is one of the important data sources required for carrying out the seismic hazard assessment procedure. The earthquake catalogue provides the spatial and temporal information about the earthquakes that have occurred in the study area and will play a major role in establishing the Gutenberg–Richter (G–R) recurrence relationship, which helps in understanding the state of seismicity of the region. Therefore, the earthquake catalogue should be compiled carefully and should include both instrumental and historic earthquake events. However, the compiled earthquake catalogue can be used for performing seismic hazard analysis only after it has been checked for homogeneity, completeness, and removal of dependent events. These can be carried out by the processes of homogenization, declustering, and estimating the magnitude of completeness. Each of these processes has its own importance. This study deals with the first and foremost step of earthquake catalogue processing, i.e., homogenization, and its effects on the estimation of maximum magnitude of the potential seismic sources

and peak ground acceleration for seismically active region of Tripura, India. Homogenization is the process of converting different magnitude scales into the common reference magnitude scale. There is a primary need for homogenization since the earthquake events are often reported in different magnitude scales.

Various national agencies like Indian Meteorological Department (IMD) and Indian Seismic and GNSS Network (ISGN) and international agencies like United States Geological Survey (USGS) and European Mediterranean Seismological Centre (EMSC) monitor and record seismic activities regionally and globally. However, they express the size of the earthquake in different magnitude scales. The most commonly used magnitude is the local magnitude scale (M_L) which can be measured by using Wood Anderson Seismometer with Eigen period of $T_S = 0.8$ s, magnification of 2800, and damping factor of 0.8 (Richter 1935). The other commonly reported magnitude scales are surface wave magnitude (M_S) and body wave magnitude (m_b). All these magnitude scales use the amplitude of the recorded seismograms. The scales m_b, M_S, and M_L have the disadvantage that they tend to saturate at various higher magnitude levels. Thus, the use of these magnitude scales for large earthquakes leads to the underestimation of the size of the earthquake (Kramer 1996). Hanks and Kanamori (1979) proposed a non-saturating magnitude scale called moment magnitude (M_W) which relates to the amount of seismic energy released during the earthquake. Therefore, converting all the magnitude scales into moment magnitude scale becomes necessary in order to realistically represent the size of the earthquake events. This problem of converting all the magnitude scales into moment magnitude can be solved by the use of magnitude scaling equations.

Empirical relations governing the magnitude scale conversion had been established by various researchers by employing different regression techniques such as standard least squares regression (Scordilis 2006; Kolathayar et al. 2012), orthogonal standard regression (Baruah et al. 2012; Pandey et al. 2017), general orthogonal regression (Castellaro et al. 2006), and chi-square regression (Grünthal and Wahlström 2003; Stromeyer et al. 2004**)**. With the updated earthquake catalogues and improvements in the regression techniques, a lot of new equations has been developed. Choosing the appropriate magnitude conversion equation is of prime importance because it may induce variations in the seismic hazard parameters. The current practice regarding the usage of magnitude scaling equations is based solely on the expert judgment without any statistical basis (Sil et al. 2013; Anbazhagan et al. 2015a). Few authors develop new magnitude scaling equations considering available data pairs for the study region (Sitharam and Sil 2014; Thingbaijam et al. 2008). The paper demonstrates the bias induced by the selection of magnitude conversion equations in estimating the magnitude of completeness and the seismicity parameters (a and b values of the G–R recurrence relation) and its effect on the estimation of maximum magnitude of seismic sources and ground motion parameters. It also discusses the potential of the regional rupture characters in suppressing this bias created by the choice of magnitude scaling equations. For this purpose, the seismicity catalogue of Tripura with a seismic study area of 700 km radius has been analyzed for magnitude of completeness and the seismicity parameters as a function of various magnitude scaling equations. The effect of these changes in the estimation of the maximum magnitude of the potential seismic sources, and ultimately the ground motion estimates, has been analyzed. Then, the robustness of the regional rupture characters in the estimation of the stable maximum magnitude has been discussed. The usability of the magnitude scaling equations to the study area can be examined by fixing the regional rupture characters as the benchmark.

2 Earthquake catalogue of the study area

In the present study, the earthquake catalogue of Tripura and adjoining regions in the Northeast India was compiled. Tripura is a northeastern Indian state which has experienced the damages of big earthquakes in the past such as 1950 Shillong earthquake, 1897 Great Assam earthquake, 1934 Nepal Bihar earthquake, and 2017 Manu earthquake. It is considered as one of the most seismically active regions in India. It lies on the verge of two convergent plate boundaries: Indian–Eurasian plate boundary in the north and the Indian–Burmese plate boundary in the east. The seismic zonation map provided by IS 1893 Part -1 (2016) places the entire state under zone V with the design peak ground acceleration of 0.36 g. Various studies have been carried out to understand the seismicity of the state. Sitharam and Sil (2014) carried out a comprehensible seismic hazard analysis for the states of Tripura and Mizoram. Das et al. (2012a) have worked extensively on the homogenization of the

earthquake catalogue pertaining to Northeast India. Das et al. (2012b) have mapped the spatial and temporal variations of magnitude of completeness and seismicity parameters for Northeast India. Das et al. (2016) carried out a probabilistic seismic hazard analysis for Northeast India. Nath et al. (2005) estimated the source parameters and spectral attenuation laws for Northeastern Himalayas using the strong motion data. However, in this study, the earthquake catalogue of Tripura was analyzed with respect to different magnitude scaling equations. The earthquake catalogue was compiled for the state of Tripura and surrounding regions. A circular area of radius 700 km centered at Agartala (capital of Tripura) was considered for the study. The earthquake data such as the latitude and longitude of the epicenter, depth, magnitude, magnitude type, hour, minute, seconds, date, month, year, and reporting agency were collected from various reporting agencies around the world such as United States Geological Survey (USGS), European Mediterranean Seismological Centre (EMSC), International Seismological Centre (ISC), Advanced National Seismic System (ANSS), Incorporated Research Institutions for Seismology (IRIS), and the published data from Bora (2016). The compiled catalogue consists of 20,980 events reported in various magnitude scales such as M_L, m_b, M_S, and M_W. The repeating earthquake events in the catalogue were removed . The maps showing the major earthquakes in the study area and the geographical locations of all the earthquakes compiled in the present study have been presented in Fig. 1.

2.1 Magnitude scaling equations

Empirical relations between different magnitude scales have been developed by various researchers considering regional or global data set with the help of various regression methods. The standard least squares regression is the most common method for establishing the scaling relationships. The least squares regression estimates the regression parameters for which the sum of squares of the vertical errors is minimum. Sitharam and Sil (2014) developed the equations for the Mizoram and Tripura states using the standard least squares regression. Kolathayar et al. (2012) developed the magnitude scaling equations for the entire India using the standard least squares regression technique. However, the standard least squares method suffers from the drawback that it does not consider the errors in both the variables. Castellaro et al. (2006) proposed that the standard least squares regression cannot be used for developing the magnitude scaling equations because both the variables are subject to error terms. He suggested that the general orthogonal regression can be used instead of standard least squares regression since it considers the error variances of both variables. Pandey et al. (2017) have used generalized orthogonal regression for homogenizing the earthquake catalogue of Northeast India and suggested that the standard error of the estimates of the coefficients could be checked for the usability of the equation.

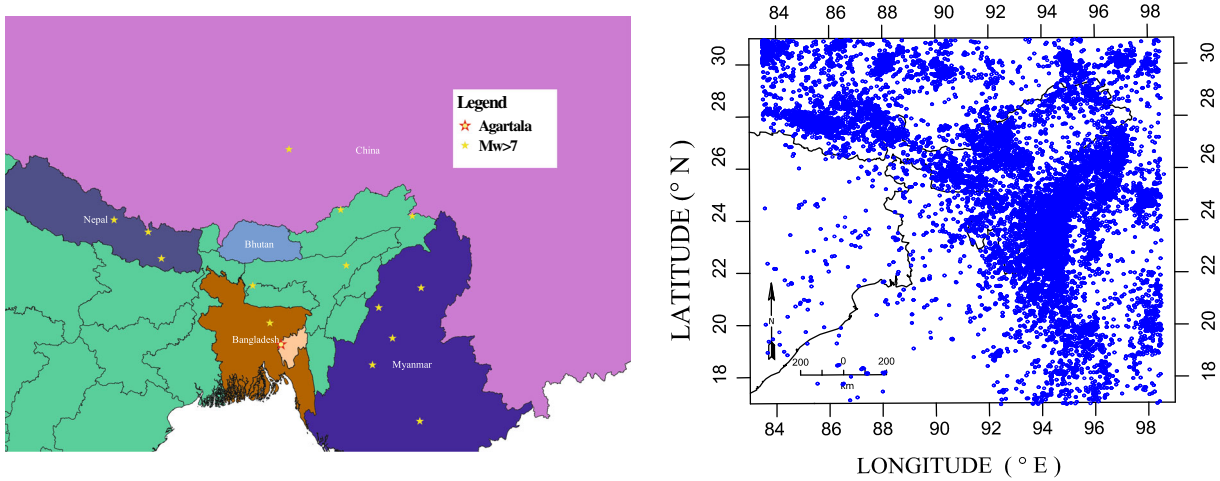

Fig. 1 **a** Map showing the major earthquakes in the study area ($M_W > 7$). **b** Map showing the location of all the compiled earthquakes in the study area

The scaling equations may be globally valid or region specific, depending on the data set used to develop it. Scordilis (2006) developed global magnitude scaling equations for converting m_b and M_S to M_W. Wason et al. (2012) emphasized on the robustness of general orthogonal regression for magnitude scaling problem and proposed global magnitude scaling equations. Lolli et al. (2014) developed magnitude scaling equations for converting teleseismic magnitudes such as m_b and $M_S$to moment magnitude. They concluded that chi-square regression technique performs better than orthogonal least squares method and mitigates the uncertainties in the equations. Das et al. (2011) have developed global regression equations using orthogonal standard regression and concluded that this regression method performs better than traditional standard least squares regression method. Gasperini et al. (2012) discussed the disagreement between the moment magnitude estimates reported by the global earthquake reporting agencies such as National Earthquake Information Center (NEIC) and Global Centroid Moment Tensor (GCMT). However, there exists the question regarding the usage of the globally valid equation for a region-specific study. A lot of researchers established the region specific magnitude scaling relationships for the Northeast India (Bora 2016; Thingbaijam et al. 2008; Baruah et al. 2012; Sitharam and Sil 2014). The Northeast India is a complex tectonic regime; thus, a single magnitude scaling equation cannot represent the entire Northeast India (Baruah et al. 2012).

With the available earthquake data from the compiled catalogue, new magnitude scaling equations were developed for the region of Tripura and surrounding region. Out of the 20,980 earthquake events in the compiled earthquake catalogue, almost 310 events were reported with more than one magnitude scale. These events with the data pairs were used for the development of region-specific magnitude scaling equations. We made use of standard least squares regression (SR) and orthogonal standard regression (OSR) for developing the equations. Since the ratio of error variances (η) is not exactly known for the study area, we assume that the errors in both the dependent and the independent variable are the same i.e., $\eta = 1$. This simplifies the general orthogonal regression (GOR) procedure to OSR procedure because OSR is the special case of GOR given $\eta = 1$. The detailed procedure of orthogonal regression was provided by Castellaro et al. (2006). The best fit lines by standard least squares regression and orthogonal standard regression for the data used in the present study have been presented in Fig. 2a–d. Simple linear functional form has been used to develop the magnitude scaling equations in this study. The functional form is given by

$$M_1 = \alpha + \beta M_2$$

where M_1 and M_2 are different magnitude scales of interest, β represents the slope, and α represents the intercept of the regression line. The coefficients obtained for different conversions obtained in the present study have been presented in Table 1.

Observing Fig 2a–d, one can note that the slope of the OSR line is steeper compared to the SR line except in the case of M_S–M_W relation for magnitude range 6.1 to 6.9. This observation can be explained by the OSR behavior that it tries to compensate for the error in the independent variable as well as the dependent variable and, therefore, slightly inclined towards the dependent variable axis. This intrinsic behavior of the orthogonal standard regression method justifies its steeper slopes compared to the standard least-squares regression method. In this study, other magnitude scaling equations which can be applied to the Northeast India were also complied. Few researchers like Scordilis (2006), Das et al. (2011), Lolli et al. (2014), and Wason et al. (2012) have provided the conversion equations for m_b–M_W and M_S–M_W only. For those equations, the conversion of M_L–M_W is assumed to follow 1:1 line. The summary of the magnitude scaling equations compiled for the present study is given in Table 2, and the visualization of the compiled equations is presented in Fig. 3a–c.

2.2 Declustering

Various algorithms are available for the process of declustering (Gardner and Knopoff 1974; Reasenberg 1985; Uhrhammer 1986). This study utilized the algorithm proposed by Reasenberg (1985) because of the simplicity of the algorithm. This method identifies dependent events by linking the events to clusters corresponding to their interaction zones. One spatial parameter and one temporal parameter define the earthquake interaction zone. The spatial extent is based on the stress distribution near the main shock, and the temporal parameter is based on the Omori's law (Omori 1894). The interaction zones have been dynamically modelled to analyze the catalogue to find the aftershock sequences.

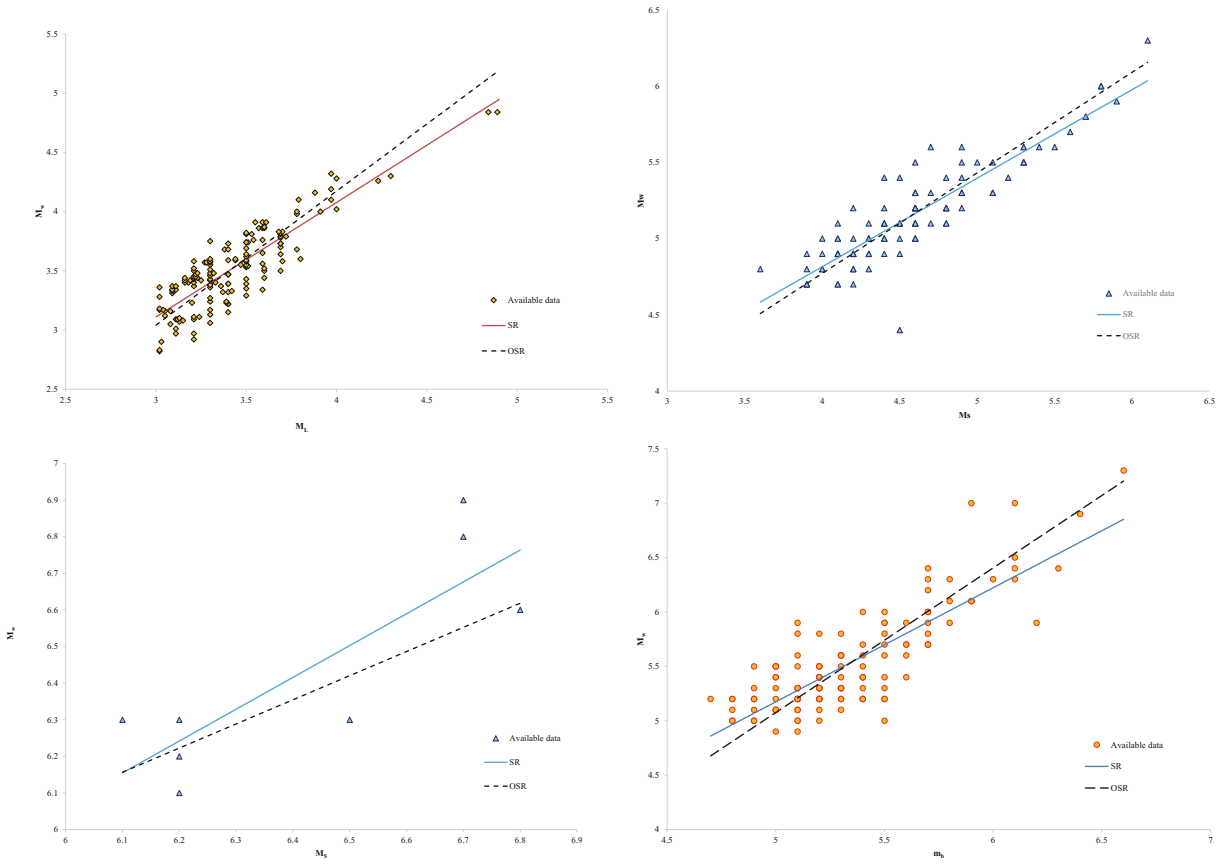

Fig. 2 **a** Relationship between M_L and M_W developed in the present study. **b** Relationship between M_S and M_W developed in the present study (Magnitude range of 3.6 to 6). **c** Relationship between M_S and M_W developed in the present study (Magnitude range of 6.1 to 6.8). **d** Relationship between m_b and M_W developed in the present study

Any earthquake that occurs within the interaction zone of a prior earthquake is to be considered as a dependent event. The parameters involved in the algorithm and the standard values are provided by Reasenberg (1985). The parameters in the algorithm and their values that will directly affect the results are presented in Table 3. The minimum (τ_{min}) and maximum (τ_{max}) look ahead time for building the clusters were fixed at 1 and 10 days respectively. The probability for observing the next earthquake (p) within the specified look ahead time is fixed at 0.95. The spatial extent (r_{fact}) within which the cluster is to be considered has been fixed at 10 km.

Table 1 Magnitude scaling equations developed in the present study

Equations	Regression technique	Intercept (α)	Slope (β)	R^2 value	No. of data points	Magnitude range
$m_b - M_W$	SR	− 0.0677 (± 0.3413)	1.0483 (± 0.0637)	0.71	127	4.7 to 6.6
	OSR	− 1.5725 (± 0.4329)	1.3296 (± 0.0808)	0.64		
$M_L - M_W$	SR	0.209 (± 0.1502)	0.967 (± 0.044)	0.75	162	3 to 4.9
	OSR	−0.3573 (± 0.1758)	1.1327 (± 0.0512)	0.73		
$M_S - M_W$	SR	2.4939	0.5805	0.8	101	3 to 6.1
	OSR	2.3172	0.6189	0.79		
$M_S - M_W$	SR	0.0283	0.9934	0.74	8	6.1 to 6.8
	OSR	− 0.1401	1.0238	0.68		

Table 2 Summary of the magnitude scaling equations compiled

Author	Regression type	Region considered	Time period	Abbreviation
Present work	SR	Tripura and surroundings	1908 to 2017	PW-SR
Present work	OSR	Tripura and surroundings	1908 to 2017	PW-OSR
Pandey et al. (2017)	GOR	Northeast India	1900–2016	PAN-GOR
Pandey et al. (2017)	OSR	Northeast India	1900–2018	PAN-OSR
Pandey et al. (2017)	SR	Northeast India	1900–2019	PAN-SR
Kolathayar et al. (2012)	SR	Entire India	250 BC to 2010	KOL-SR
Castellaro et al. (2006)	GOR	Italy	1981 to 1996	CAST-GOR
Sitharam and Sil (2014)	SR	Tripura and Mizoram	1731 to 2011	SISL-SR
Das et al. (2011)	OSR	Global	1976 to 2007	DAS-OSR
Lolli et al. (2014)	CSQ	Global	1976 to 2010	LOL-CSQ
Wason et al. (2012)	GOR	Global	1976 to 2007	WAS-GOR
Scordilis (2006)	SR	Global	1976 to 2003	SCO-SR

SR standard least squares regression, *OSR* orthogonal standard regression, *GOR* general orthogonal regression, *CSQ* chi-squared regression

These values are used in accordance with the standard values used by Reasenberg (1985).

Each of the catalogues homogenized by different scaling equations was subjected to the declustering

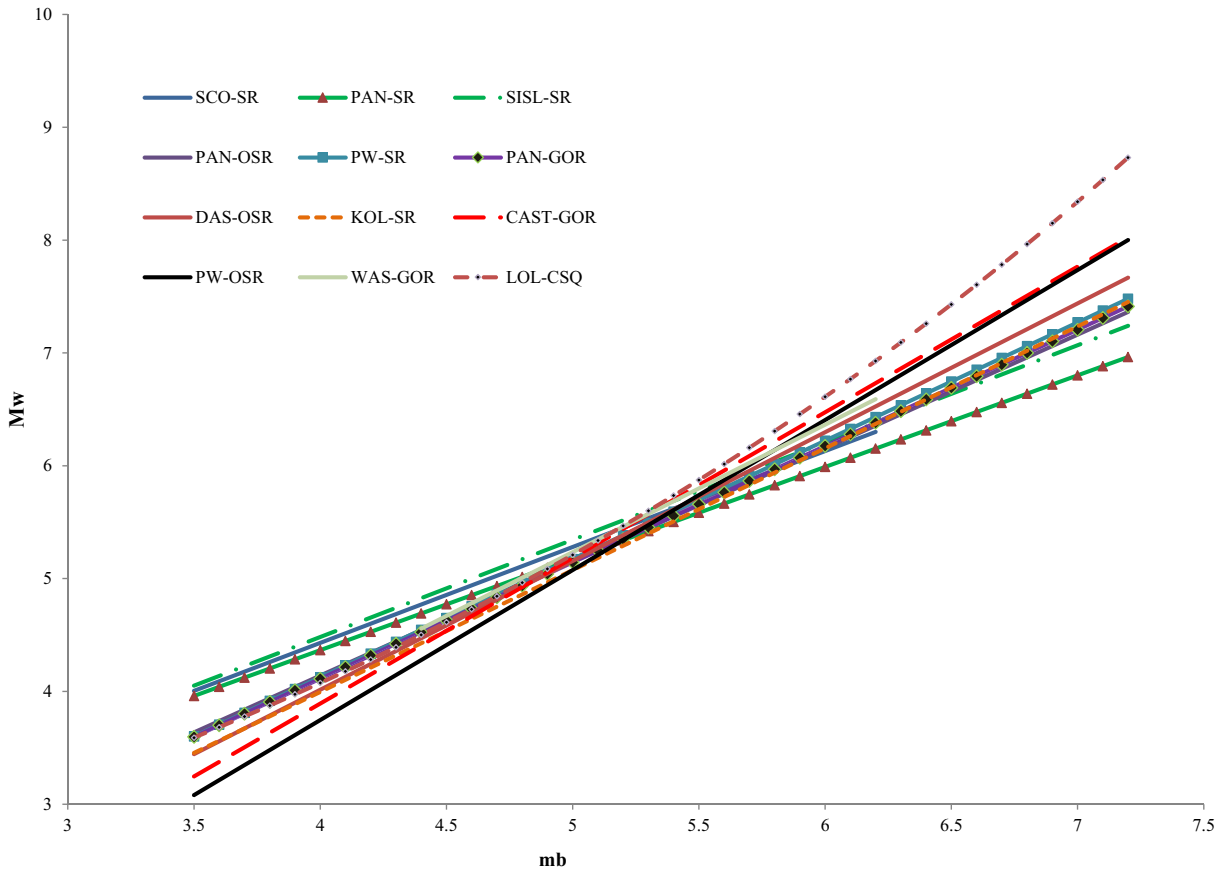

Fig. 3 **a** Comparison of the compiled magnitude scaling equations for m_b–M_W. **b** Comparison of the compiled magnitude scaling equations for M_L–M_W. **c** Comparison of the compiled magnitude scaling equations for M_S–M_W

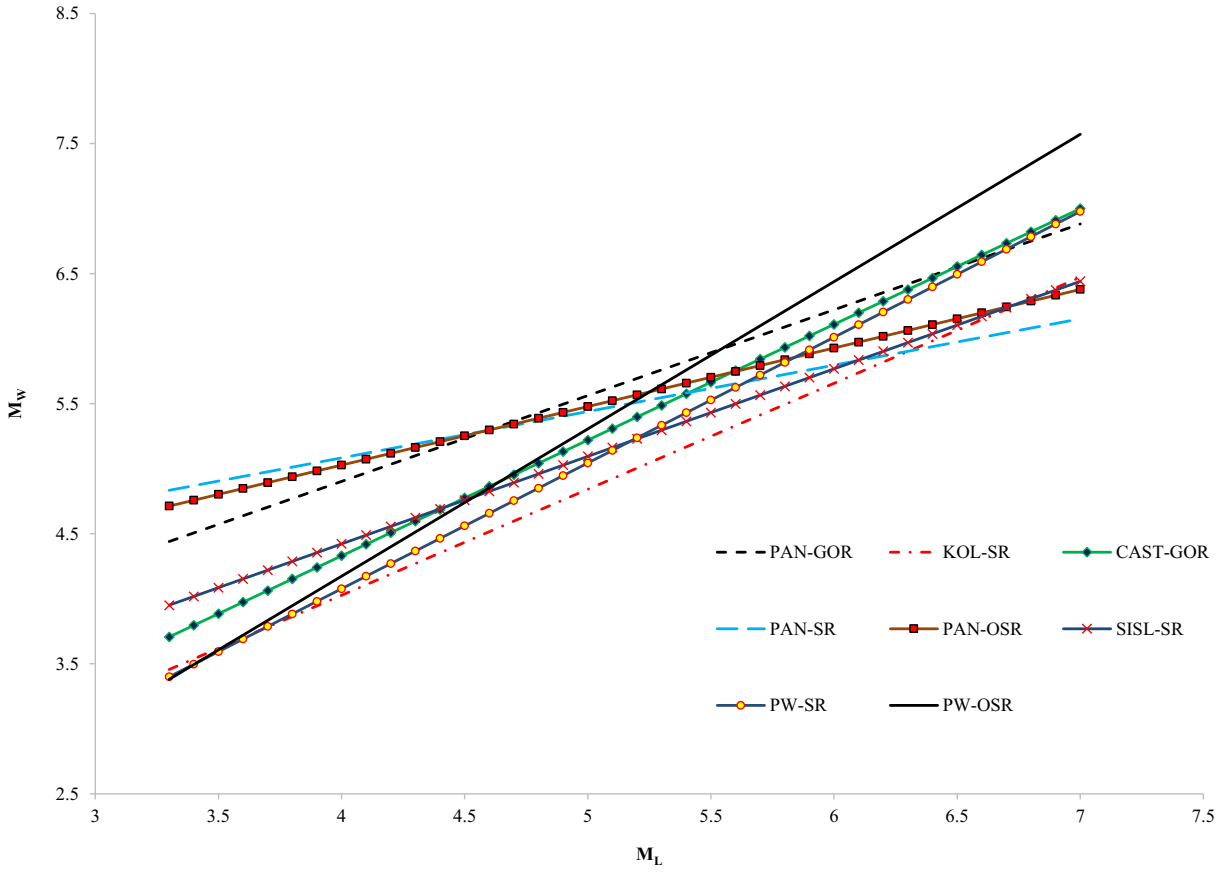

Fig. 3 (continued)

process. The results of the declustering are presented in Table 4. It can be seen that the percentage of removal of dependent events differs for each of the conversion equation. This indicates that there is a variation in the identification of foreshocks and aftershocks in each of the catalogues. It was observed that declustered catalogue of SISL-SR has almost 706 events lesser than the declustered catalogue of PW-OSR which indicates that the PW-OSR catalogue reports additional 706 events as main shocks. This could completely undermine the seismic understanding of the region under consideration. Interesting observation is that both these equations were developed for the same region. This indicates that the number of earthquake events (main shocks) significantly differs for catalogues homogenized by different equations even if they are applicable to the similar region. Even the globally applicable scaling equations show significant differences in the number of earthquakes in the declustered catalogue. This shows that the declustering process is indifferent towards the type of magnitude scaling equation (regional or global).

2.3 Seismicity parameters

Every catalogue has to be checked for consistency and homogeneity because the events recorded in the catalogue comes from spatially unevenly distributed seismic stations. The minimum magnitude for which a seismicity catalogue can be considered as complete is known as magnitude of completeness (M_c). It can be estimated using the trend of the frequency magnitude relation. Frequency magnitude distribution (FMD) of any catalogue should follow the well-known Gutenberg–Richter (G–R) relation (Gutenberg and Richter 1944) given by,

$$\log(N) = a + b\,M$$

where N is the number of earthquakes having magnitude more than M and a,b are real integers which represent the intercept and the slope respectively. When looked

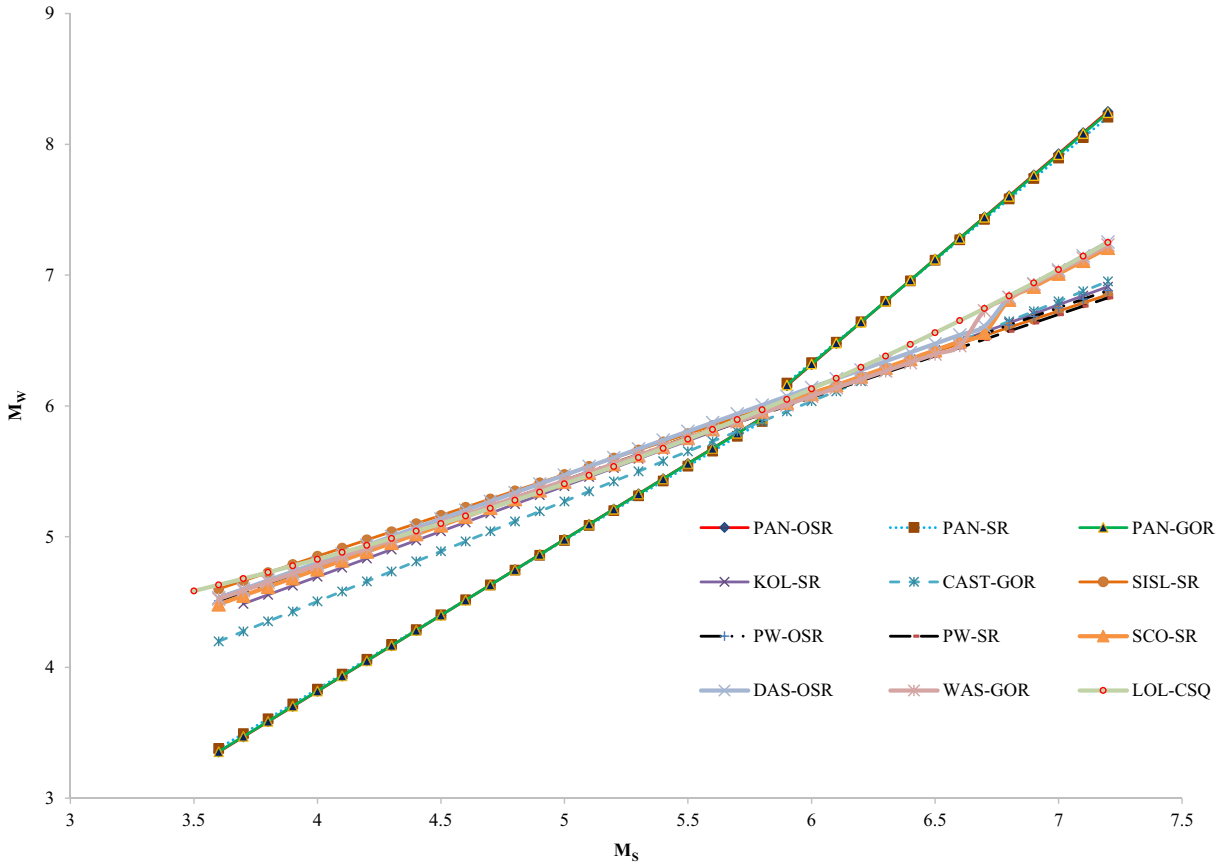

Fig. 3 (continued)

closely, these parameters can reveal a lot more information about the seismicity of the region. *a* value is a measure of the number of earthquakes for zero magnitude which indicates the seismicity of the region, whereas *b* value presents the relative number of larger earthquake events with respect to the smaller earthquake events which provides insight into the relative frequency of occurrence of larger earthquake events with respect to

Table 3 Parameters used in the declustering algorithm

Parameter	Definition	Value used
τ_{min}	The min value of look ahead time for building clusters	1 day
τ_{max}	The max value of look ahead time for building clusters	10 days
$p1$	The probability of detecting the next clustered event used to compute the look-ahead time	0.95
r_{fact}	The radii surrounding the events within which the cluster is considered	10 km
Epicenter error	Horizontal location error	1.5 km
Depth error	Vertical location error	2 km

Table 4 Percentage removal of dependent events for each of the catalogue

Scaling equations	No. of events after declustering	Percentage of events removed (%)
PAN-OSR	10,726	48.9
PAN-SR	10,491	50.0
PAN-GOR	10,732	48.8
PW-OSR	10,937	47.87
PW-SR	10,630	49.33
SISL-SR	10,231	51.23
KOL-SR	10,813	48.46
CAST-GOR	10,799	48.53
DAS-OSR	10,852	48.28
LOL-CSQ	10,584	49.55
WAS-GOR	10,379	50.52
SCO-SR	10,694	49.02

the smaller earthquake events. The point where the frequency magnitude curve deviates from the ideal G–R curve is considered as the magnitude of completeness.

This study makes use of the goodness of fit method proposed by Woessner and Wiemer (2005) for estimating the magnitude of completeness. This method makes use of the observed data and the synthetic data which is modelled by the G–R power law. The magnitude of completeness is considered as the minimum magnitude at which the goodness of fit exceeds 95%, or the percentage of residuals becomes less than 5%. This indicates that about 95% of the observed data can be modelled by the G–R power law. Using this method, the magnitudes of completeness for the set of 12 catalogues were estimated. Table 5 shows the estimated magnitude of completeness and the G–R relation parameters of the catalogues homogenized by different magnitude conversion equations. M_c varies for each catalogue with the standard deviation of 0.23. The difference in the order of 0.23 U in the M_c can cause considerable variations in subsequent analysis of maximum magnitude estimation and earthquake prediction models. The *a* and *b* parameters of the G–R relation also shows significant deviation. The bias in the *b* value can propagate as a bias in the maximum magnitude estimation and probabilistic seismic hazard analysis (PSHA) studies. The estimation of maximum magnitude by Kijko and Sellevoll (1989) depends on the *b* parameter of the G–R relation. The *b* parameter is also used in the PSHA studies to quantify the magnitude uncertainty.

Table 5 Variations in the seismicity parameters for each conversion equation

Catalogue homogenized by	M_c	*a*	*b*
PAN-OSR	2.4	6.23	0.68
PAN-SR	2.6	6.52	0.72
PAN-GOR	2.4	6.25	0.68
PW-OSR	2.1	5.91	0.63
PW-SR	2.4	6.23	0.68
SISL-SR	2.8	6.71	0.74
KOL-SR	2.3	6.15	0.67
CAST-GOR	2.3	6.03	0.64
DAS-OSR	2.2	5.97	0.64
LOL-CSQ	2.1	5.89	0.63
WAS-GOR	2.1	5.96	0.65
SCO-SR	2.0	5.84	0.62

Figure 4 depicts the variation of the M_c and the seismicity parameters for different magnitude scaling equation. It can be seen that the trend of the variation remains the same for different parameters. Whenever the magnitude of completeness (M_c) increases, the slope (*b* value) of the G–R relation becomes steeper and hence the intercept (*a* value) also increases. This observation can be explained by the indirect relation between the magnitude of completeness and the seismicity parameters. When the minimum magnitude of completeness increases, it means that there are a greater number of larger magnitude events compared to smaller magnitude events in the catalogue. This automatically increases the relative occurrence of larger magnitude events with respect to the lower magnitude events; in other words, the slope of the G–R relation becomes steeper. Since the slope has become steeper and it is decreasing function, it ultimately leads to increase in the intercept of the line as well. This could be the explanation for similar trend of variation in magnitude of completeness and the seismicity parameters. However, this is not true for the ideal cases. If the slope of the Frequency Magnitude Distribution (FMD) (Frequency Magnitude Distribution is the plot between the magnitude and the logarithm of the cumulative number of earthquakes in the catalogue) is regular, then the M_c can increase even if the seismicity parameters remain the same. Therefore, the FMDs of the catalogues are presented in Fig. 5a. From Fig. 5a, it is evident that the slope of FMDs is not regular and ultimately depends on the magnitude of completeness. This depicts the dependence of seismicity parameters on the M_c value in the present study. In order to understand the effect of the number of earthquakes on the *b* value assessment, the FMDs of the catalogues homogenized by various magnitude scaling equations are superimposed with the FMD of the catalogue with the greatest number of mainshocks in Fig. 5b. Also, the maps showing the locations of the mainshocks in each of the catalogues are presented in Fig. 5c.

3 Estimation of maximum magnitude (M_{max})

M_{max} can be defined as the maximum possible magnitude of earthquake that could occur in a seismic source. Estimation of the M_{max} forms an indispensable part of the seismic studies. The ground motion parameters such as peak ground acceleration (PGA) or peak spectral acceleration (PSA) directly depends on the maximum possible

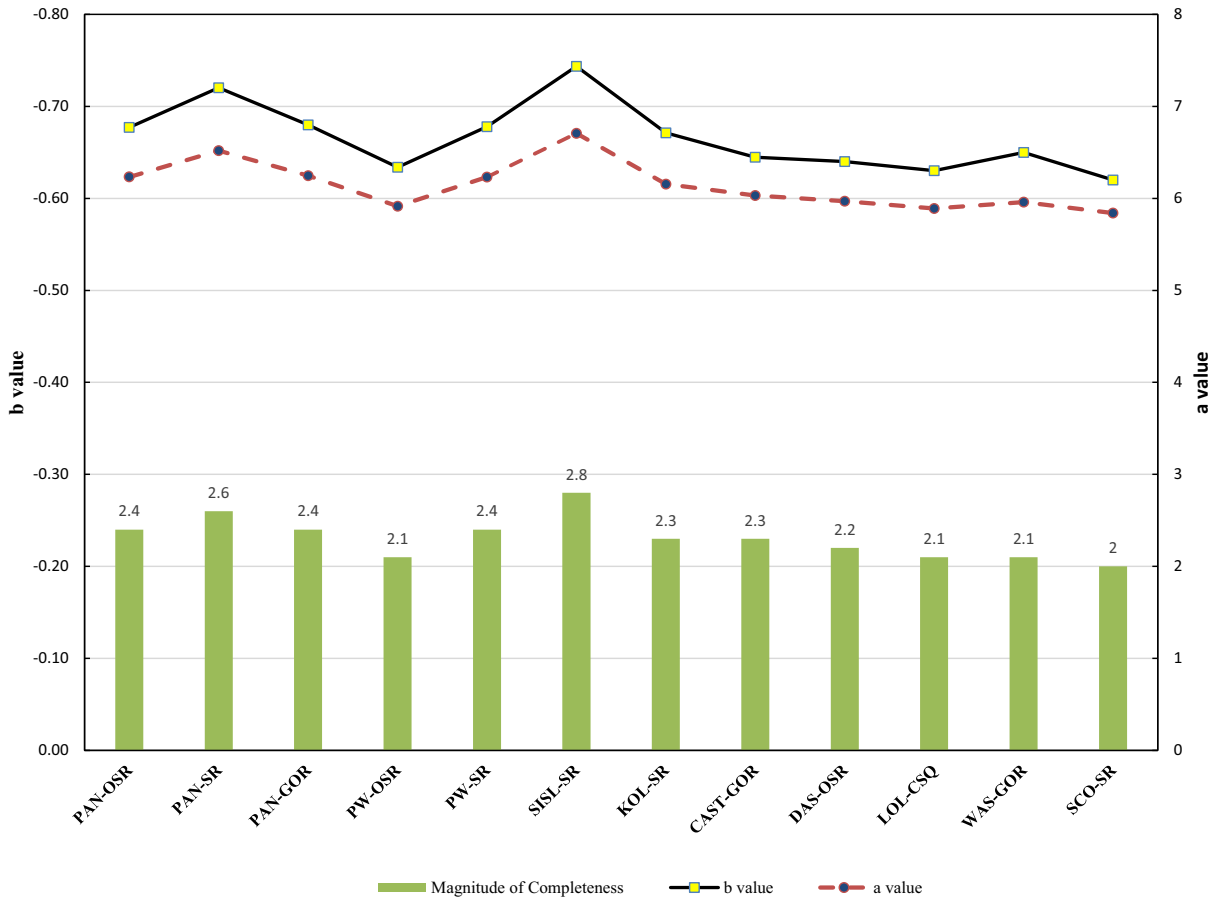

Fig. 4 Variations of M_c and the seismicity parameters for different magnitude scaling equations

magnitude of all the seismic sources. The M_{max} value can be computed deterministically (Anderson et al. 1996; Jin and Aki 1988), probabilistically (Kijko and Sellevoll 1989; Kijko and Singh 2011), or considering the regional rupture characteristics (Anbazhagan et al. 2015b). The deterministic methods are based on the empirical relationship between magnitude and various tectonic parameters, whereas the probabilistic approach is based on the Gutenberg–Richter frequency magnitude relation. The probabilistic method proposed by Kijko and Sellevoll (1989) completely depends on the seismicity parameters such as magnitude of completeness and a and b parameters of the G–R recurrence relation. However, in most of the cases, M_{max} is estimated from the maximum observed magnitude (m_{max}^{obs}) by giving a finite increment depending on the seismic activity of the region. Osher (1996) has developed the formulation for the estimation of confidence limits for the maximum magnitudes based on the information on the uncertainty of magnitude estimations of earthquakes.

In this study, an attempt has been made to study the effect of the bias in seismicity parameters on the maximum magnitude estimation by the probabilistic method proposed by Kijko and Sellevoll (1989). All the potential seismic sources in the study area has been complied and used to estimate maximum magnitude. The major seismic sources in the Northeast India include the eastern extension of the Main Boundary Thrust (MBT), Main Central Thrust (MCT), Main Frontal Thrust (MFT), Sagaing Fault, Dauki Fault, and the Arakan Yoma Subduction zone. The other tectonic features and the seismic gaps are also compiled from various literatures (Hurukawa and Maung 2011; Yin and Harrison 2000) and the Seismotectonic Atlas of India (Dasgupta et al. 2000). The faults and the seismic gaps are precisely digitized from the abovementioned published sources. Figure 6 shows the seismic sources in Northeast India which are used in this study. There are more than 500 seismic sources in the Northeast India which includes faults, lineaments, and the seismic gaps.

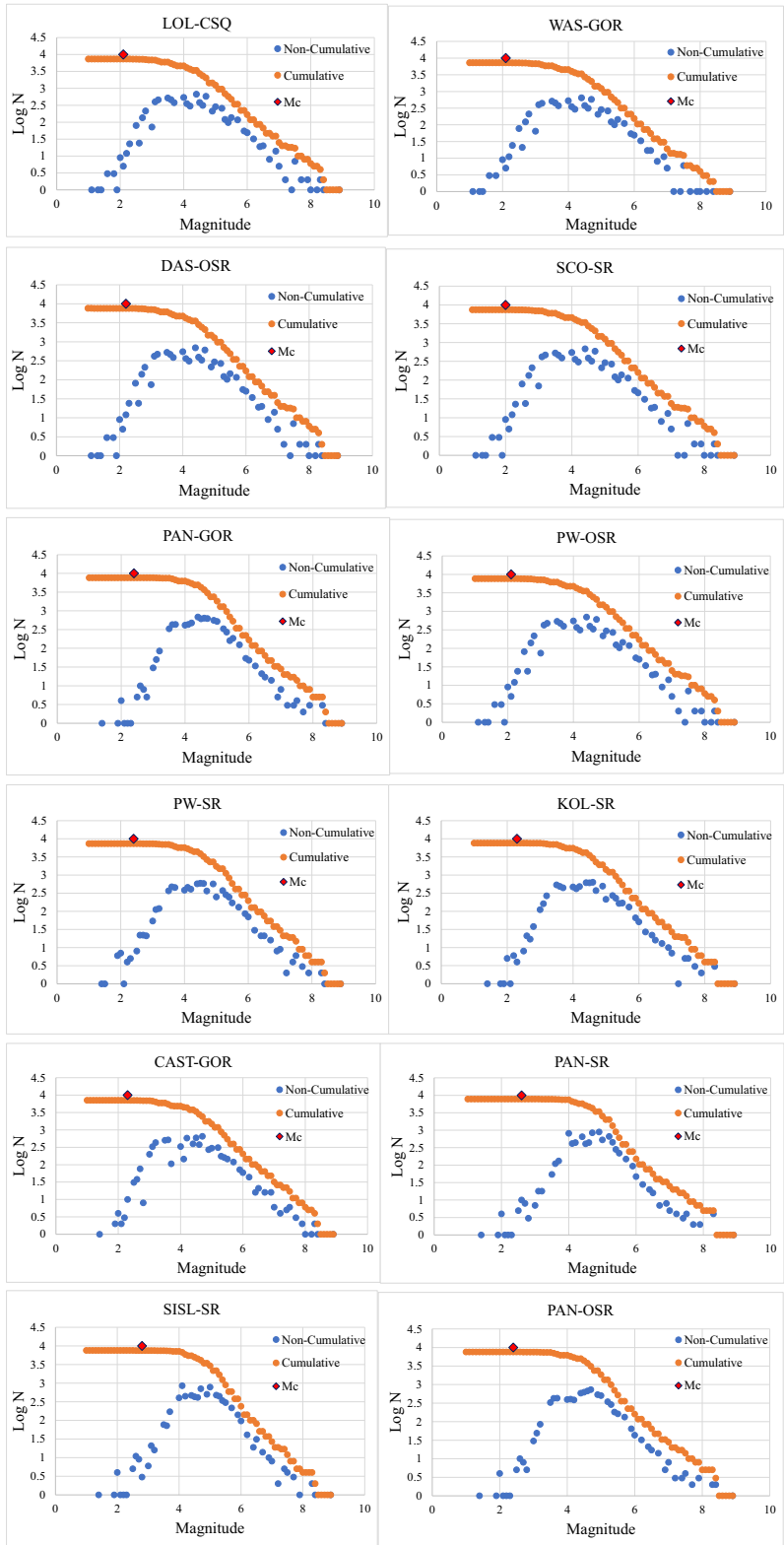

Fig. 5 **a** Frequency magnitude distribution plot for different magnitude scaling equations. **b** Comparison of the Frequency Magnitude Distribution (FMD) of various catalogues with respect to the catalogue with the greatest number of mainshocks. **c** Locations of the mainshocks in each of the declustered catalogue

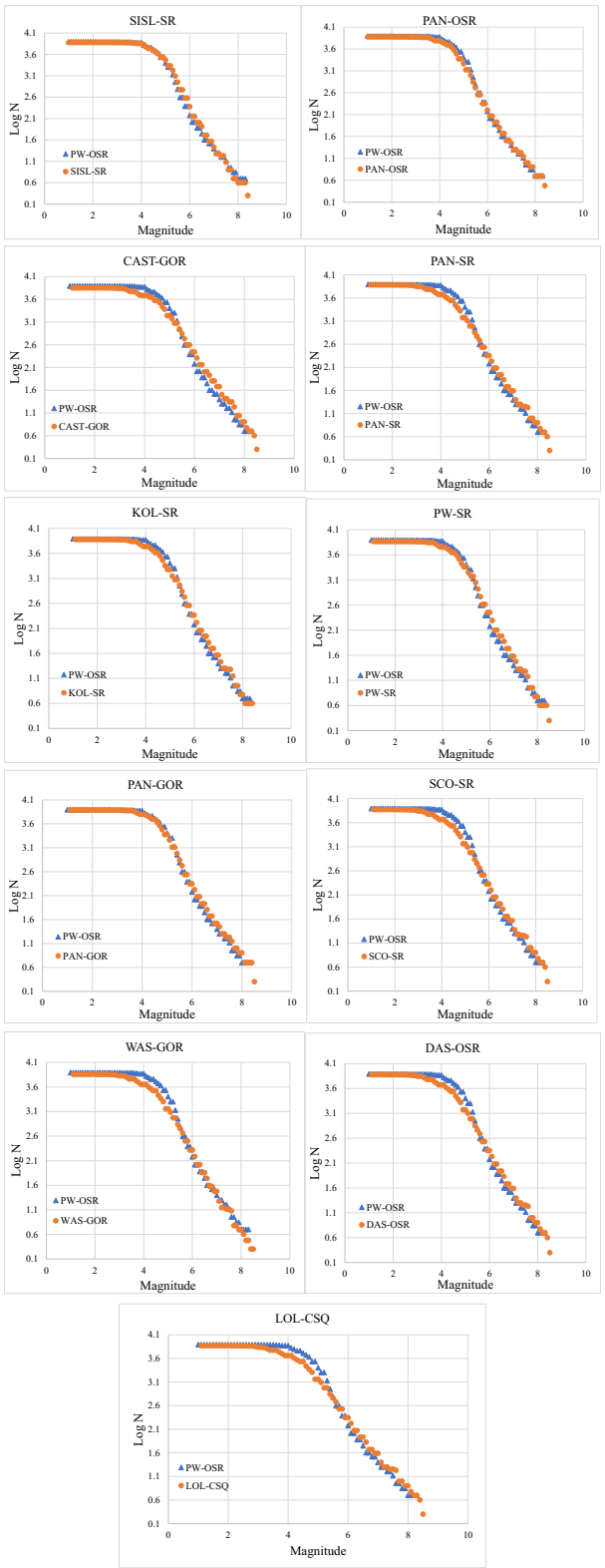

Fig. 5 (continued)

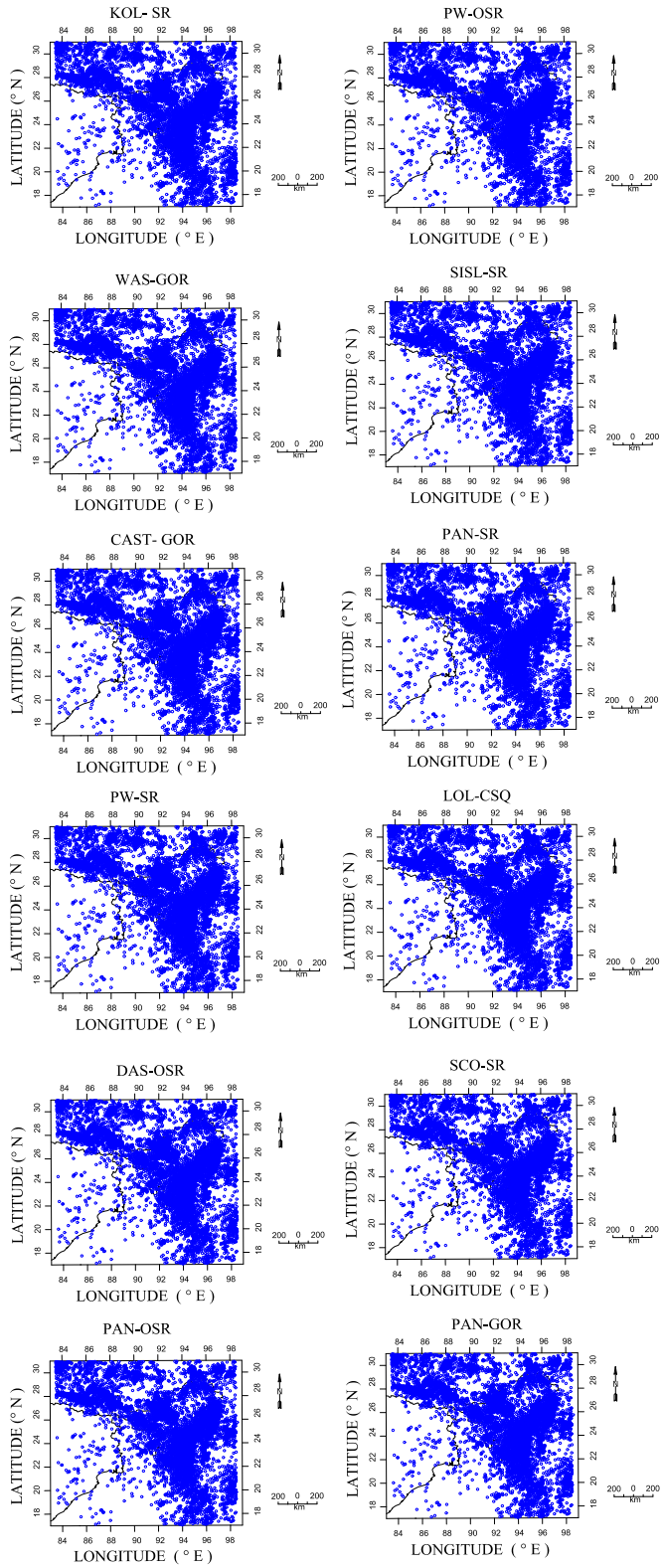

Fig. 5 (continued)

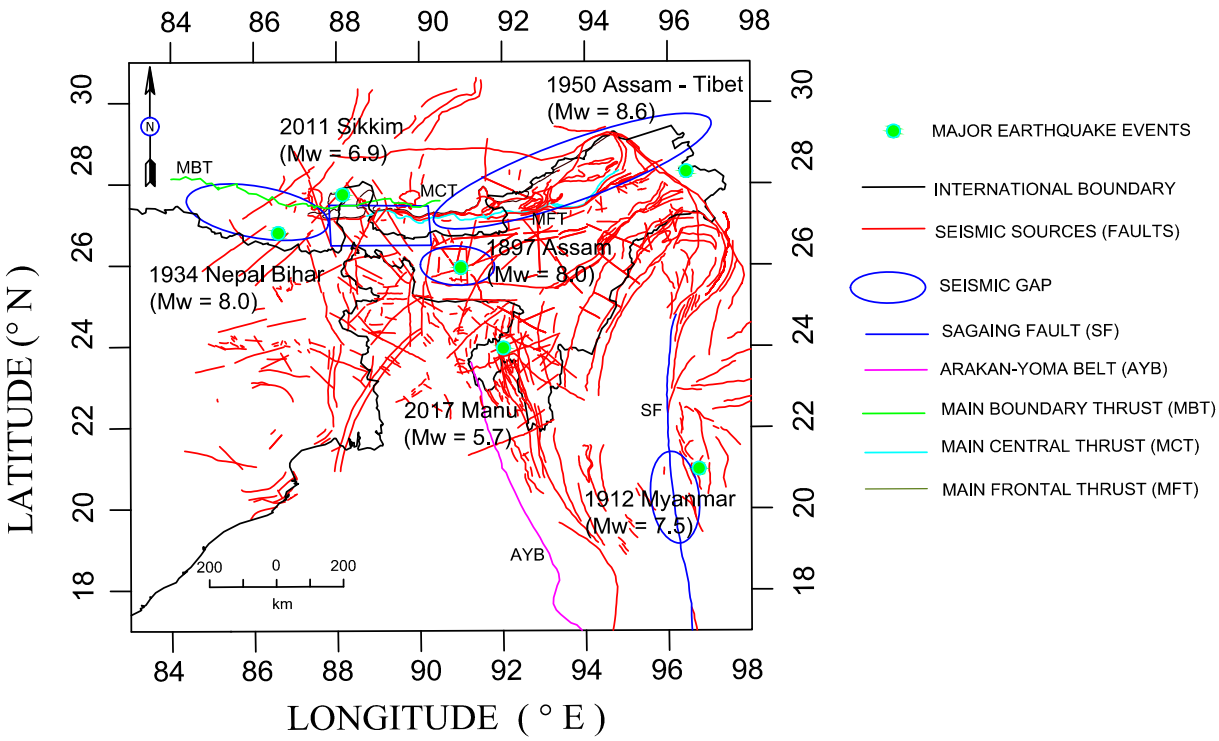

Fig. 6 Potential seismic sources in the study area and the location of major earthquake events in the past

Areal seismic sources are not considered in the study. The fault map of the study area consisting all the potential seismic sources and the seismic gaps is presented in Fig. 6.

By analyzing the earthquake catalogue and the compiled fault details, each earthquake event was assigned to the nearest seismic source. Then, the total number of earthquakes assigned to each of the fault is computed and the maximum magnitude of the earthquakes in the fault was considered as maximum observed magnitude ($m_{\max}^{\text{obs}}$). A MATLAB code was written for this purpose as the data is too large for manual assignment. The process was repeated using the entire declustered catalogue as mentioned in Table 4. The maximum magnitude observed for each of the faults in the study area using different catalogues is presented in Table 6. To add more clarity to the faults given in Tables 6 and 7, the tectonic regime in which the fault is located has been presented in column 2. There are three main tectonic regimes in the study area—Himalayan Thrust zone (HTZ) in the north and northeastern side of the study area, Indo-Burma subduction zone (IBS) in the eastern side of the study area, and the intra-plate seismicity (IPS) in the Bengal Basin and Shillong plateau in the western side of the study area.

From Table 6, it can be seen that the standard deviation in the $m_{\max}^{\text{obs}}$ of faults could be as high as 0.5 moment magnitude units. This indicates that there could be a variation of about 5.62 times the actual seismic energy released. Not only it will affect the estimation of the ground motion parameters, but also it misrepresents the actual seismic potential of that particular seismic source. This could hinder the studies related to the identification of seismic gaps as well. However, the $M_{\max}$ estimation was also carried out to understand the propagation of this bias in the seismic hazard studies. The incremental method of $M_{\max}$ estimation will report the same amount of variation since it just gives a constant increment to $m_{\max}^{\text{obs}}$. In this study, more emphasis was placed on the $M_{\max}$ estimated by probabilistic approach. $M_{\max}$ values are estimated for the faults using the probabilistic method proposed by Kijko and Sellevoll (1989) which takes the following form:

$$M_{\max} = m_{\max}^{\text{obs}} + \frac{E_1(n_2) - E_1(n_1)}{\beta \exp(-n_2)} + m_{\min}\exp(-n)$$

where $M_{\max}$ is the maximum probable magnitude, $m_{\max}^{\text{obs}}$ is the maximum observed magnitude, and n is the number of earthquakes above $m_{\min}$

Table 6 List of faults showing significant variations in the m_{max}^{obs} estimates

Fault no.	Zone	PAN-GOR	PW-SR	PW-OSR	KOL-SR	CAST-GOR	SISL-SR	PAN-SR	PAN-OSR	DAS-OSR	LOL-CSQ	WAS-GOR	SCO-SR	Std deviation	Mean	Max diff.	Min diff.
F219	HTZ	6.3	6.3	6.3	5.3	6.4	6.4	5.3	6.3	6.4	5.4	5.8	6.2	0.5	6.0	1.1	0
MBT	HTZ	5.9	5.9	6.9	6.7	6.9	6	6.1	5.9	5.8	5.9	6	5.9	0.4	6.2	1.1	0
F132	IPS	6.7	6.7	5.8	6.1	5.9	6.7	6.5	6.7	6	6.7	6.7	6.5	0.4	6.4	0.9	0
F205	HTZ	6.4	6.4	6.5	6.3	5.3	6.4	6.3	6.4	6.3	6	5.8	6.2	0.3	6.2	1.2	0
F151	IPS	6.1	5.5	6.2	5.4	6.2	5.6	6	6.1	6	5.7	6.3	6.2	0.3	5.9	0.9	0
F438	IBS	7.6	7	7	6.9	7	7	7.6	7.6	7.2	7	7.3	7.2	0.3	7.2	0.7	0
F353	IBS	6.7	6.7	6.9	6.6	6.9	6.7	5.9	6.7	6.6	6.8	6.4	6.8	0.3	6.6	1	0
F426	IBS	5.5	6	5.5	5.4	5.4	6.1	5.5	5.5	5.4	5.6	5.4	6	0.3	5.6	0.7	0
MFT	HTZ	7.3	6.8	6.8	6.8	6.7	6.8	7.3	7.3	6.8	6.9	6.8	7.1	0.2	7.0	0.6	0
F424	IBS	5.6	5.7	5.6	5.6	5.7	6.2	6.2	5.6	5.7	5.9	5.7	6.1.	0.2	5.8	0.6	0
F164	HTZ	5.5	6	6	6	5.6	6	5.6	5.5	5.7	5.5	6	5.6	0.2	5.8	0.5	0
F163	HTZ	6.5	6.5	6.8	6.5	6.9	6.5	6.2	6.5	6.5	6.4	6.3	6.5	0.2	6.5	0.7	0
F432	IBS	5.7	5.9	5.8	5.8	5.8	6	5.3	5.7	5.9	5.7	5.7	5.8	0.2	5.8	0.7	0
F443	IBS	5.4	5.5	5.4	5.4	5.5	6	5.5	5.4	5.4	5.7	5.9	5.4	0.2	5.5	0.6	0
F334	IBS	6.1	6.2	6.4	6.1	6.5	6.5	6.1	6.1	6.3	6.2	6.2	6.2	0.2	6.2	0.4	0
F352	IBS	5.6	5.5	5.3	5.4	5.7	5.9	5.6	5.6	5.6	5.7	5.5	5.7	0.2	5.6	0.6	0
F134	IPS	5.3	5.4	5.2	5.3	5.3	5.8	5.4	5.3	5.7	5.5	5.6	5.6	0.2	5.5	0.6	0
F284	HTZ	5.5	5.7	5.4	5.5	5.7	5.9	5.7	5.5	5.4	5.7	5.4	5.6	0.2	5.6	0.5	0
F228	HTZ	5.3	5.4	5.1	5.2	5.3	5.6	5.5	5.3	5.4	6	5.5	5.7	0.2	5.4	0.9	0
F445	IBS	6	6	6.1	5.9	6.3	6.2	5.8	6	5.6	5.9	5.9	6.1	0.2	6.0	0.7	0

MBT Main Boundary Thrust, *MFT* Main Frontal Thrust, *HTZ* Himalayan Thrust zone, *IBS* Indo-Burma subduction, *IPS* intra-plate seismicity

$$\beta = -b\log(10)$$

$$n_1 = \frac{n}{\{1-\exp[-\beta(m_{max}-m_{min})]\}}$$

$$n_2 = n_1\{\exp[-\beta(m_{max}-m_{min})]\}$$

E_1 denotes the exponential integration function which is given by

$$E_1(z) = \frac{z^2 + a_1 z + a_2}{z(z^2 + b_1 z + b_2)}\exp(-z)$$

where $a_1 = 2.334733$, $a_2 = 0.250621$, $b_1 = -3.330657$, and $b_2 = -1.681534$ (Abramowitz and Stegun 1972). The value of m_{min} was taken to M_W 4.0 because this is the minimum magnitude for structural damage in most cases. From the equation, we can see that the M_{max} estimated from this method directly depends on the b parameter and the maximum observed magnitude of the seismic sources. Since there exists a significant variation in these parameters connected to the change in the magnitude scaling equation, it is obvious that there exists a certain amount of variation in the M_{max} estimates.

One of the newer methods of estimating the maximum magnitude is the use of data mining techniques and analyzing the past seismicity (Last et al. 2016). Last et al. (2016) has used various seismicity indicators (or parameters) to predict the maximum magnitude in the succeeding year. One of the simple and effective parameters used in their study was the probability of a given magnitude to exceed the threshold magnitude. If the mean of all the earthquake magnitudes occurred in the fault is taken as the threshold value, then the probability is given by the following equation:

$$P(M > M_{mean}) = 10^{-b\,(M-M_{mean})}$$

In the present study, MATLAB code was written to assign the earthquakes in the catalogue to the nearest

Table 7 List of faults showing significant variations in the M_{max} estimates using method proposed by Kijko and Sellevoll (1989)

Fault no.	Zone	PAN-GOR	PW-SR	PW-OSR	KOL-SR	CAST-GOR	SISL-SR	PAN-SR	PAN-OSR	DAS-OSR	LOL-CSQ	WAS-GOR	SCO-SR	Std deviation	Mean	Max diff	Min diff
F82	HTZ	7.7	5.9	8.1	7.7	8.2	6.9	7.7	7.7	6.5	7.5	8.1	7	0.7	7.4	2.3	0
F255	IBS	5.9	7.0	6.5	5.4	6.7	7.5	7.2	5.9	7	6.8	7.1	6.4	0.6	6.6	2.1	0
F219	HTZ	6.3	6.3	6.3	5.3	6.4	6.4	5.3	6.3	6.5	5.4	5.9	6.2	0.5	6.1	1.2	0
F343	IBS	6.5	5.9	6.5	5.4	5.9	6.9	6.4	6.2	5.8	6.3	6.4	6.7	0.4	6.2	1.5	0
MBT	HTZ	5.9	5.9	6.9	6.7	6.9	6.0	6.1	5.9	5.8	5.9	6.2	6.1	0.4	6.2	1.1	0
F105	IPS	6.2	6.3	6.2	6.2	6.3	7.5	6.3	6.2	6.4	6.7	6.5	6.7	0.4	6.5	1.3	0
F205	HTZ	6.4	6.4	6.5	6.3	5.3	6.4	6.3	6.4	6.3	6	5.9	6.4	0.3	6.2	1.2	0
F132	IPS	6.7	6.7	5.8	6.1	5.9	6.7	6.5	6.7	6	6.8	6.7	6.5	0.4	6.4	1	0
F244	HTZ	8.8	8.8	8.1	8.5	8.2	8.0	8.1	8.8	7.9	8.0	8.3	7.8	0.4	8.3	1	0
F151	IPS	6.1	5.5	6.2	5.4	6.2	5.6	6.0	6.1	6	5.7	6.3	6.2	0.3	5.9	0.9	0
F85	IPS	6.6	7.0	6.6	6.6	6.7	7.5	6.6	6.6	6.8	7.2	6.5	6.8	0.3	6.8	1	0
F438	IBS	7.6	7.0	7.0	6.9	7.0	7.0	7.6	7.6	7.2	7	7.3	7.2	0.3	7.2	0.7	0
F353	IBS	6.7	6.7	6.9	6.6	6.9	6.7	5.9	6.7	6.7	6.8	6.5	6.8	0.3	6.7	1	0
F426	IBS	5.5	6.0	5.5	5.4	5.4	6.1	5.5	5.5	5.4	5.6	5.5	6	0.3	5.6	0.7	0
MFT	HTZ	7.3	6.8	6.8	6.8	6.7	6.8	7.3	7.3	6.8	6.9	6.8	7.1	0.2	7.0	0.6	0
F424	IBS	5.6	5.7	5.6	5.6	5.7	6.2	6.2	5.6	5.8	5.9	5.7	6.1.	0.2	5.8	0.6	0
F164	HTZ	5.5	6.0	6.0	6.0	5.6	6.0	5.6	5.5	5.7	5.5	6.1	5.6	0.2	5.8	0.6	0
F163	HTZ	6.5	6.5	6.8	6.5	6.9	6.5	6.2	6.5	6.6	6.4	6.3	6.5	0.2	6.5	0.7	0
F432	IPS	5.7	5.9	5.8	5.8	5.8	6.0	5.3	5.7	5.9	5.8	5.7	6.0	0.2	5.8	0.7	0
F443	IBS	5.4	5.5	5.4	5.4	5.5	6.0	5.5	5.4	5.4	5.7	5.9	5.4	0.2	5.5	0.6	0

MBT Main Boundary Thrust, *MFT* Main Frontal Thrust, *HTZ* Himalayan Thrust zone, *IBS* Indo-Burma subduction, *IPS* intra-plate seismicity

seismic source. The mean magnitude of the earthquakes in individual faults is calculated and the probability of a given magnitude exceeding the mean magnitude is estimated. For worst-case scenario earthquake, the magnitude for which the probability of exceeding is less than 2% is considered as the maximum magnitude of the fault. The process is repeated for all the seismic sources in the study area and by changing the catalogues of various magnitude scaling equations. Since the probability function itself depends on the *b* value, the M_{max}values estimated by this method will also be dependent on the *b* value. The variations in the M_{max} estimates by Kijko and Last method are presented in Tables 7 and 8 respectively.

A maximum deviation of 0.7 magnitude units was observed for both Kijko and Last method. The reason for this deviation could be the accumulation of the bias in the *b* parameter and the maximum observed magnitude. However, since the Last method is based on data mining from the earthquake catalogue and directly dependent on the b parameter and the mean magnitude, the deviation in maximum magnitude is almost the same as of Kijko method. The deviations in the number of earthquake events could have also contributed for this bias. Since Kijko method is the widely used method for maximum magnitude estimation and the results from Kijko and Last method are almost the same, Kijko method is alone considered for the rest of the analysis.

It can be inferred from this observation that the choice of magnitude scaling equation to be used in the study could lead to a deviation of 0.7 moment magnitude units in the estimation of the maximum magnitude of the seismic sources. To get more insights into the effects of this variation, its effect on the ground motion parameter estimates was also studied. National Disaster Management Authority (NDMA 2010) has provided the ground motion prediction equations (GMPE) for different tectonic regimes in India. The functional form of the GMPE is the same for all these regions. The coefficients applicable for Northeast India, which would be more appropriate for the present study, have been used. The functional form of the GMPE is presented as follows:

Table 8 List of faults showing significant variations in the M_{max} estimates using method proposed by Last et al. (2016)

Fault no.	Zone	PAN-GOR	PW-SR	PW-OSR	KOL-SR	CAST-GOR	SISL-SR	PAN-SR	PAN-OSR	DAS-OSR	LOL-CSQ	WAS-GOR	SCO-SR	Std deviation	Mean	Max diff	Min diff
F219	HTZ	7.6	5.7	8	7.5	7.9	6.6	7.5	7.6	6.5	7.4	8	7	0.7	7.3	2.3	0
MBT	HTZ	5.8	6.9	6.4	6.6	6.7	7.4	7.2	5.6	7	6.7	7.1	6.4	0.5	6.7	1.8	0
F132	IPS	6.1	6.1	6.2	5.3	6.4	6.3	5.1	6.3	6.5	5.4	5.9	6.2	0.5	6.0	1.4	0
F205	HTZ	6.5	5.7	6.5	5.4	5.8	6.7	6.4	6.2	5.8	6.3	6.4	6.7	0.4	6.2	1.3	0
F151	IPS	5.9	5.8	6.9	6.6	6.9	6	6.4	5.9	5.7	5.8	6.1	6.1	0.4	6.2	1.2	0
F438	IBS	6	6.2	6.1	6.2	6.3	7.4	6.1	6.2	6.3	6.5	6.5	6.6	0.4	6.4	1.4	0
F353	IBS	6.3	6.2	6.3	6.2	5.2	6.3	6.3	6.4	6.1	5.8	5.9	6.1	0.3	6.1	1.2	0
F426	IBS	6.6	6.7	5.7	6.1	5.8	6.7	6.5	6.5	6	6.6	6.6	6.5	0.4	6.4	1	0
MFT	HTZ	8.5	8.6	8	8.4	8	8	7.9	8.6	7.6	7.9	8.3	7.8	0.3	8.1	1	0
F424	IBS	6	5.3	6.1	5.4	6.2	5	5.4	6.1	6	5.5	6.2	6.2	0.4	5.8	1.2	0
F164	HTZ	6.5	6.9	6.5	6.6	6.7	7.3	6.5	6.5	6.5	7.1	6.3	6.8	0.3	6.7	1	0
F163	HTZ	7.6	6.8	7	6.7	6.8	7	7.5	7.4	7.2	7	7	7.1	0.3	7.1	0.9	0
F432	IBS	6.6	6.7	6.8	6.6	6.9	6.5	5.9	6.7	6.6	6.8	6.3	6.8	0.3	6.6	1	0
F443	IBS	5.5	5.8	5.5	5.3	5.4	6.1	5.4	5.5	5.3	5.6	5.5	6.1	0.3	5.6	0.8	0
F334	IBS	7.3	6.5	6.8	6.9	6.7	6.8	6.3	7.3	6.5	6.9	6.5	7	0.3	6.8	1	0
F352	IBS	5.5	5.7	5.5	5.6	5.7	6	6.2	5.4	5.8	5.8	5.7	6.1.	0.2	5.7	0.8	0
F134	IPS	5.5	5.8	6	5.9	5.6	6.1	5.6	5.8	5.6	5.5	6	5.6	0.2	5.8	0.6	0
F284	HTZ	6.5	6.4	6.8	6.5	6.7	6.5	6.2	6.5	6.3	6.4	6.2	6.5	0.2	6.5	0.6	0
F228	HTZ	5.6	5.9	5.5	5.8	5.8	5.8	5.3	5.6	5.9	5.8	5.5	6	0.2	5.7	0.7	0
F445	IBS	5.5	5.5	5.4	5.3	5.5	5.8	5.3	5.4	5.4	5.6	5.9	5.4	0.2	5.5	0.6	0

MBT Main Boundary Thrust, *MFT* Main Frontal Thrust, *HTZ* Himalayan Thrust zone, *IBS* Indo-Burma subduction, *IPS* intra-plate seismicity

$$\ln(\text{PSA or PGA}) = c_1 + c_2M + c_3M^2 + c_4R + c_5\ln\left[R + c_6\exp^{(c_7M)}\right] + c_8\ln(R)f_o + \ln(\varepsilon)$$

where PSA is peak spectral acceleration in units of gravity, PGA is the peak ground acceleration in units of gravity, and f_o = max [ln(R/100), 0] (Fig. 7). In the GMPE, the magnitude should be in moment magnitude units and R denotes the hypocentral distance. The equation is valid for magnitude range of M_W 4.0–8.5 and the hypocentral distances less than 500 km. The coefficients for the peak ground acceleration are c_1 = − 4.2427; c_2 = 1.31; c_3 = − 0.0097; c_4 = − 0.0031; c_5 = − 1.3159; c_6 = 0.0172; c_7 = 1.0279; c_8 = 0.1083; ε = 0.4424. The abovementioned functional form considers the geometric spreading, anelastic attenuation, and magnitude saturation, and the coefficients are obtained from stratified regression.

The variation in the peak ground acceleration was analyzed using the above GMPE. The hypocentral distances ranging from 50 to 500 km in the increment of 10 km were plugged in the GMPE. The PGA values against the hypocentral distance are plotted for M_W = 4.0 and M_W = 4.7 and presented in Fig. 7 in order to visualize the effect of the actual variation of 0.7 observed in M_{max} estimate. M_W 4.0 was chosen because it is the minimum magnitude that could cause infrastructure damage. It was observed that the variation of 0.7 moment magnitude units leads to an uncertainty of 0.2 g in PGA estimates in the near field. However, as the hypocentral distance increases, the difference in the PGA estimate seems to subside. This indicates that the M_{max} estimation becomes very critical in the near-field. This observation suggests that the M_{max} values of the seismic sources should be unequivocal and certain. Estimation of M_{max} should not depend on the choice of selecting the magnitude scaling equations for homogenizing the earthquake catalogue.

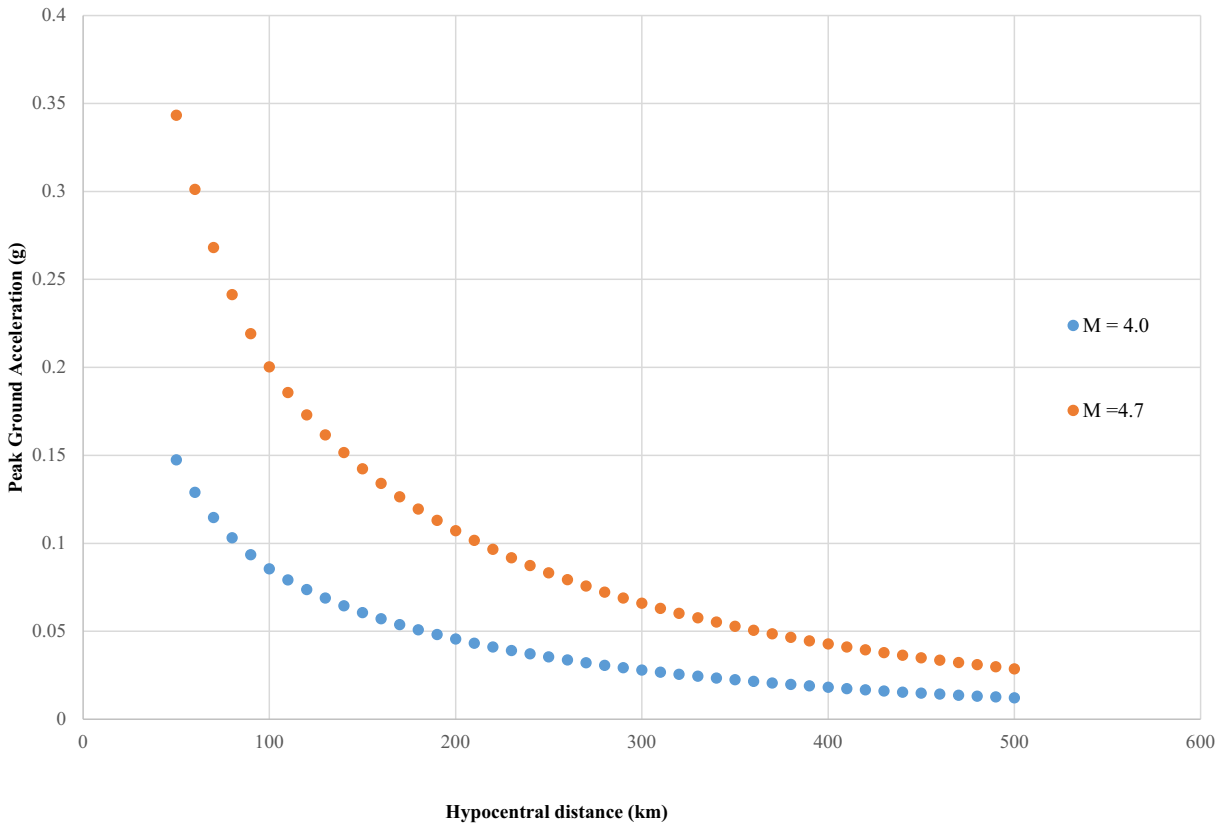

Fig. 7 Variation in the peak ground acceleration for different M_{max} estimates

4 M_{max} considering regional rupture characters

Since there is uncertainty in the M_{max} estimated from probabilistic method, there becomes a necessity to estimate a stable M_{max} that does not depend on the earthquake catalogue. Anbazhagan et al. (2015b) proposed the method for estimating the maximum magnitude that does not depend on the seismicity parameters and seismic study area. This method utilizes the parameter called percentage of fault rupture (PFR) which is the ratio of the subsurface rupture length (RLD) to the total fault length (TFL). The M_{max} value calculated from this method depends on the rupture characteristics of the seismic sources in the study area and does not depend on the magnitude of completeness or the seismicity parameters (Anbazhagan et al. 2015b). In this work, the subsurface rupture length (RLD) was calculated from the moment magnitude using the empirical relationship given by Wells and Coppersmith (1994).

$$\log(\mathrm{RLD}) = 0.59M_W - 2.44$$

RLD denotes the subsurface rupture length, and M_W is the moment magnitude. The subsurface rupture length used here indicates the maximum length to which the seismic source has ruptured in the past, indicating the rupture character of the seismic source. The seismic sources in the present study were compiled by digitizing tectonic maps. With the digitized coordinates of the seismic sources (faults), the TFL can be easily computed. With the RLD and TFL, PFR can be calculated. Then, PFR is plotted against the TFL of the seismic sources to get the regional rupture characteristics of the seismic sources.

The relation mentioned above is the general formulation that is applicable to all types of faults. According to Wells and Coppersmith (1994), there is no statistical significance at 95% confidence interval for the difference between various faulting types. Subdividing the data according to the tectonic setting does not typically improve the results (Wells and Coppersmith 1994).

They have also suggested that all-slip type equation is more appropriate for relation between subsurface rupture length and magnitude because the uncertainty is smaller in all-type regression equation than in other individual slip-type equations. Also, all the seismic sources in the seismic study area do not have a defined slip-type. Therefore, for consistency and statistical significance, the general equation has been used in the present study.

The M_{max} estimate from the region rupture character does not depend on the earthquake catalogue but is a function of the behavior of the fault rupture. In this study, effort has been made to study the stability of the maximum magnitude estimates by the regional rupture method. To have better understanding of the stability of M_{max} estimates with respect to the size of earthquake catalogue (which indirectly affects the seismicity parameters), three different catalogue sizes were exploited. One catalogue which has not been declustered, one with maximum percentage of removal from Table 4, and one with minimum percentage of removal from Table 4 were selected to study the rupture characters. The rupture character observed has been presented in Fig. 8.

From Fig 8, it is evident that the regional rupture character stays the same irrespective of the magnitude scaling equation used. So, M_{max} can be estimated as a function of the fault rupture length rather than seismicity parameters. From the sensitivity analysis of M_{max} with respect to the catalogue size, it can be said that the M_{max} calculated from the rupture-based procedure seems to be stable and is not sensitive to changes in the past seismicity or the earthquake catalogue.

With this clear understanding, the rupture method was applied to all the catalogues homogenized by various scaling equations. The unique trend between the PFR and TFL is established, indicating the character of the rupture phenomenon of the seismic sources in the study area. The same procedure was repeated by changing the magnitude scaling equations. A clear trend was obtained showing the regional rupture behavior. Figure 9 depicts the plot between PFR and TFL for different magnitude scaling equations. Interesting observation was the rupture character of the region remains unaffected by the choice of using the magnitude scaling equations. This independent behavior of the rupture of the seismic sources could be harnessed for estimating the maximum magnitude of the seismic sources.

In this study, the total fault length was first divided into distance bins based on visible trend change. The TFL was divided into five bins such as 0–50 km, 50–100 km, 100–200 km, 200–500 km, and greater than 500 km. The maximum and average PFR in each of the distance bins were calculated. Based on the average and the maximum PFR, the observed PFR was increased to a certain amount based on expert judgment. The maximum and average PFR in each PFR is given in Table 9.

In the above context, the distance bins are synonymous with the bins of fault lengths. From the table above, it can be inferred that the PFR values does not vary significantly with respect to the catalogue used. The maximum standard deviation observed was 6% corresponding to 200–500 km bin. This could mean that the deviation corresponds to the deviation of rupture length of almost 20 km which is negligible. Taking into consideration the maximum PFR values and the deterministic procedure being based on worst-case scenario, the PFR values were increased for each of the distance bins. The maximum PFR values of 90% were adopted for the faults within the length bins of 0–50 km, 50–100 km, and 100–200 km. The maximum PFR values adopted for the length bins of 200–500 km and greater than 500 km were 45% and 20% respectively. It can be observed from Table 9 that the maximum PFR observed does not vary much for the magnitude bins of 0–50 km, 50–100 km, and 100–200 km. Considering this similar behavior, equal PFR was adopted for these length bins. However, for length bins of 200–500 km and greater than 500 km, the maximum PFR observed was less compared to the other bins. Therefore, considerably lower PFR values were adopted for these bins. Since the faults in these bins have larger lengths, even a smaller PFR value could lead to larger magnitude values. These maximum PFR values were adopted for all the faults in the particular length bin to estimate the worst-case scenario earthquake. From the PFR, subsurface rupture length was calculated and ultimately the maximum magnitude was estimated using the empirical equation of Wells and Coppersmith (1994). Table 10 provides the comparison between the stable M_{max} calculated using the rupture-based procedure and the mean M_{max} obtained from the probabilistic methods.

From Table 10, it can be observed that the M_{max} values estimated from the rupture-based method are higher than the values estimated from Kijko and Last method. The faults' rupture characters are modelled based on their total fault length. It means that the faults having similar fault length to rupture synonymously. It would be unrealistic to assign a low M_{max} value to a

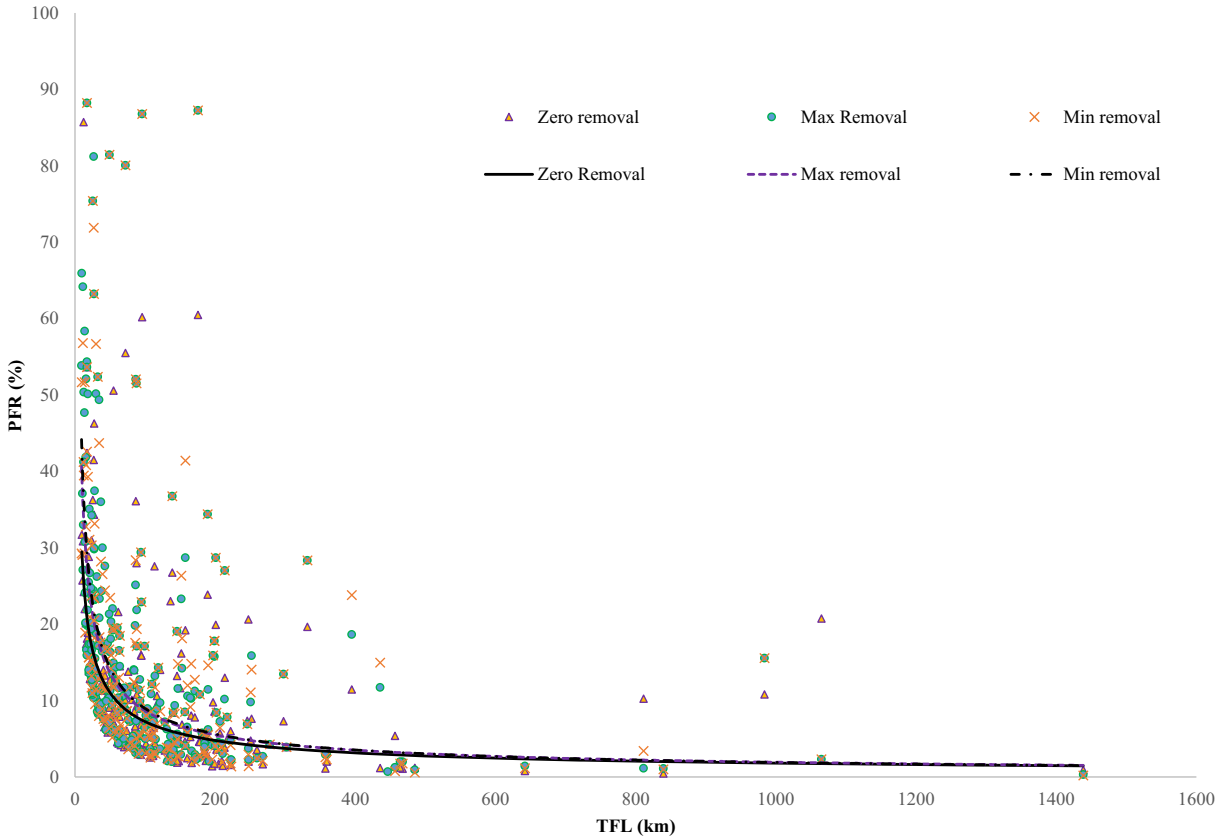

Fig. 8 Comparison of the regional rupture characters for the catalogue without declustering, catalogue with minimum removal of events (PW-OSR), and catalogue with maximum removal of events (SISL-SR)

fault having a certain fault length, whereas another fault having almost the similar length is assigned a larger M_{max}. If a seismic source of certain length can rupture to a certain extent of PFR, then all the other seismic sources having similar length has the potential to rupture to the same extent of PFR. This important attribute of the seismic sources is not captured by the earthquake catalogues and the probabilistic method of M_{max} estimation, whereas it becomes conspicuous when the rupture characters are analyzed. For this reason, the rupture-based procedure for M_{max} calculation is considered to be more robust than other methods that depend on the past seismicity and does not consider the intrinsic behavior of the fault rupture.

5 Results and discussion

The seismic hazard studies depend on the seismic data in the form of earthquake catalogue. The catalogue should be homogeneous inorder to be consistently used for the hazard analysis. However, the ambiguity on the equation to be used for homogenization could pose serious variations in the hazard estimates. The uncertainties in the estimation of G–R seismicity parameters have been studied by various researchers. Tinti and Mulargia (1985) and Rhoades (1996) have studied the effect of uncertainties in magnitude estimates on the seismicity parameters. Bender (1983) discussed the effect of unavoidable binning of magnitudes on the estimation of seismicity parameters. In the present study, the effect of homogenization of earthquake catalogue on the estimation of seismicity parameters was explored. It was observed that the deviations in the magnitude of completeness can be 0.23. For example, the catalogue homogenized by SCO-SR will report the magnitude of completeness as 2.0, whereas the one homogenized by SISL-SR will report the M_c as 2.8. The differences in a and b values for these two catalogues were observed to be 0.87 and 0.12 respectively. These differences can lead to completely different results in terms of M_{max} and ground motion estimates. This depicts the effect of

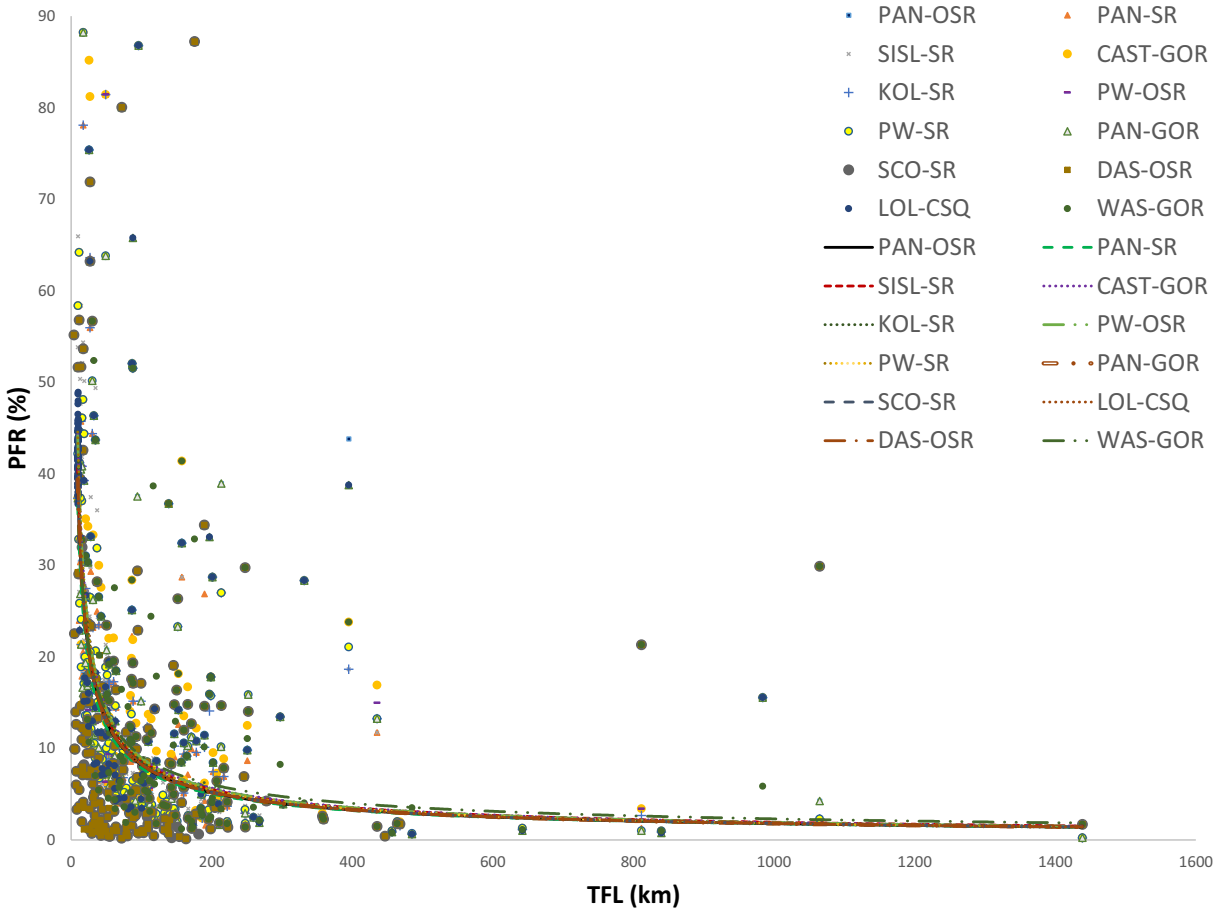

Fig. 9 Plot between PFR and TFL showing the behavior of the regional rupture phenomenon

bias in choosing the magnitude scaling equations and its effect on the seismic parameters. The reason for this variation was analyzed in this study. It was identified that the choice of magnitude scaling equations completely changes the characteristics of the earthquake catalogue and its attributes such as the magnitude of

Table 9 Maximum and average PFR values (in %) observed in different distance bins

		PAN-SR	PAN-OSR	PAN-GOR	PW-SR	PW-OSR	KOL-SR	CAST-GOR	SISL-SR	DAS-OSR	LOL-CSQ	WAS-GOR	SCO-SR
0–50 km	Max	81.44	88.21	88.21	88.21	88.21	81.44	88.21	88.21	88.21	88.21	81.44	88.21
	Mean	24.76	26.48	26.14	27.62	28.4	27.23	28.36	26.53	26.47	26.53	27.62	27.34
50–100 km	Max	86.78	86.78	86.78	86.78	86.78	86.78	86.78	86.78	86.78	85.43	86.78	86.78
	Mean	12.84	13.91	14.42	15.11	15.96	15.06	15.87	13.92	14.16	15.88	15.96	13.79
100–200 km	Max	87.24	87.24	87.24	87.24	87.24	87.24	87.24	87.24	87.24	87.24	87.24	87.24
	Mean	8.87	9.42	9.41	9.51	10.37	9.49	10.23	9.59	9.67	10.14	9.77	9.64
200–500 km	Max	38.95	43.81	38.95	28.7	28.7	28.7	28.7	28.7	28.7	28.7	28.7	28.7
	Mean	8.34	8.71	8.56	7.75	7.56	7.24	7.9	7.62	7.13	8.24	8.11	7.87
> 500 km	Max	15.56	15.56	15.56	15.56	15.56	15.56	15.56	15.56	15.56	15.56	15.56	15.56
	Mean	3.88	3.81	3.8	3.56	3.93	3.79	3.91	3.64	3.79	3.67	3.82	3.86

Table 10 Comparison of the mean M_{max} estimates from Kijko method and Last method with the estimates from rupture based method

Fault no.	Zone	Mean M_{max} from Kijko method	Mean M_{max} from Last method	M_{max} from rupture based method
F82	HTZ	7.4	7.3	7.5
F255	HTZ	6.6	6.7	6.4
F219	IPS	6.1	6.0	8.0
F343	HTZ	6.2	6.2	7.2
MBT	IPS	6.2	6.2	8.3
F105	IBS	6.5	6.4	7.8
F205	IBS	6.2	6.1	6.9
F132	IBS	6.4	6.4	7.9
F244	HTZ	8.3	8.1	7.0
F151	IBS	5.9	5.8	7.3
F85	HTZ	6.8	6.7	7.9
F438	HTZ	7.2	7.1	8.4
F353	IBS	6.7	6.6	8.4
F426	IBS	5.6	5.6	8.4
MFT	IBS	7.0	6.8	8.6
F424	IBS	5.8	5.7	7.9
F164	IPS	5.8	5.8	8.1
F163	HTZ	6.5	6.5	8.3
F432	HTZ	5.8	5.7	7.9
F443	IBS	5.5	5.5	7.1

MBT Main Boundary Thrust, *MFT* Main Frontal Thrust

completeness and the seismicity parameters. Since the current practice of seismic hazard analysis depends on the past seismicity and the compiled earthquake catalogue, the changes in the catalogue could lead to differences in hazard prediction. It was shown in the present study that the declustered earthquake catalogue was completely different for each of the magnitude scaling equations even though all the declustering parameters were kept constant. This observation seriously affects all the other estimations and predictions that depend solely on the catalogue.

The probabilistic method of extrapolating the frequency magnitude distribution to estimate the maximum magnitude of seismic sources solely depends on the attributes of the compiled earthquakes catalogue, meaning that it is sensitive to the changes in the earthquake catalogue or the past seismicity. In the present study, maximum variation of 0.7 moment magnitude units in the estimation of M_{max} was observed because of the changes in the earthquake catalogue. A change of 0.7 moment magnitude would commensurate with the variation of 11.22 times the seismic energy released. This could propagate as variation in PGA estimate up to 0.2 g even for the minimum range of magnitudes. Over prediction or under prediction of 0.2 g of peak ground acceleration can lead to catastrophic effects on structural design and economy of high-rise buildings and other important structures. Osher (1996) also studied the uncertainties in the maximum magnitude estimation because of the propagation of the uncertainties from the magnitude estimation itself. They have reported a maximum uncertainty of 0.5 U in maximum magnitude estimation because of the error in magnitude estimation of the earthquakes.

In this study, it is proved that the regional rupture characters could be used as a robust tool for the estimation of maximum magnitude of seismic sources. This is because of the fact that the rupture behavior of the seismic sources does not depend on the past seismicity and the trend is the same irrespective of the magnitude scaling equation used. In the present study, it was observed that the M_{max} predicted using the rupture-based procedure is higher than the M_{max} predicted using the Kijko and Last methods. This behavior can be explained with the help of the plot between the PFR and TFL. All the faults falling under a particular TFL bin are expected to rupture in a similar way in terms of the percentage of fault rupture. For example, if a fault within a particular length bin can rupture up to x % of its length, then all other faults in that bin have the potential to rupture up to x % in the future even if the past seismicity does not reveal the same. This characteristic of the seismic sources was harnessed to estimate the M_{max}. However, the process of demarking the length bins becomes very critical for the rupture-based method and should be backed by expert opinion.

The regional rupture characters could also provide some valuable insights towards the applicability of the magnitude scaling equations to the study area. In this way, one might get to know whether a particular magnitude scaling equation could represent the regional rupture characters in a better way. It is obvious that region-specific magnitude scaling equations would better represent the rupture character of the study area. One way is to compare the M_{max} estimates predicted from the rupture procedure and the M_{max} estimated from the Kijko method using the catalogue homogenized using that particular magnitude scaling

equation. The difference between the M_{max} estimates from the rupture method and Kijko method using a magnitude scaling equation could be a measure of robustness of that particular magnitude scaling equation in representing the regional rupture behavior. For ideal cases, the difference should be zero which would indicate that the magnitude scaling equation is able to represent the rupture behavior of the region perfectly. This almost resembles the residual analysis between M_{max} predicted from rupture method and M_{max} from kijko method using a magnitude scaling equation. In such case, statistical tools could be used to quantify the significance of the difference. Statistics such as mean and standard deviation of the residuals could provide valuable information regarding the match between the two variables. The mean, median, and standard deviation of the residuals for each of the magnitude scaling equations were presented in Table 11. These parameters could help in interpreting both the central tendency and the spread of the residual distribution.

From Table 11, it can be observed that PAN-SR shows greater deviation in terms of both central tendency and the spread of the residual distribution. CAST-GOR performs considerably well with respect to mean but performs poorly when the standard deviation is considered. SISL-SR performs very well in terms of central tendency and has a considerable spread. However, the difference between the mean and median suggests the significant number of outliers in the distribution. The global relations such as DAS-OSR, LOL-CSQ, WAS-GOR, and SCO-SR have comparatively lesser mean which suggests the robustness of the equations. However, the uncertainties associated with their predictions are higher which is evidenced from higher standard deviations. The regionally developed equation PW-OSR has the lowest standard deviation and considerably lower means indicating the robustness of the equation. Also, the fact that the median is not very far from mean indicates that the effect of outliers in the distribution is comparatively small. The interpretation from the discussion is that the regionally developed equations for the study area (PW-OSR and SISL-SR) perform reasonably better than non-region-specific equations (CAST-GOR and KOL-SR). This suggests the importance of using region-specific magnitude scaling equations in the seismic hazard analysis. Thus, the regional rupture characters could be viewed as one of the ways of analyzing the usability of a particular magnitude scaling equation to a specific site.

Table 11 Mean, median, and standard deviation of the residuals for different magnitude scaling equations

Magnitude scaling equations	Residuals = M_{max} (rupture) − M_{max} (Kijko)		
	Standard deviation	Mean	Median
PAN-OSR	0.9336	1.4902	1.5801
PAN-SR	0.8977	1.5307	1.5751
PAN-GOR	0.9297	1.5007	1.5923
PW-OSR	0.8428	1.5122	1.5714
PW-SR	0.8925	1.5015	1.5604
SISL-SR	0.8771	1.3949	1.4905
KOL-SR	0.8627	1.5443	1.6257
CAST-GOR	0.9015	1.4722	1.5022
DAS-OSR	0.9321	1.4421	1.634
LOL-CSQ	1.0058	1.3704	1.5231
WAS-GOR	1.0739	1.2824	1.3521
SCO-SR	0.8977	1.4963	1.5557

6 Conclusion

The homogeneity of the seismic catalogue is of utmost importance for the hazard studies. With the increasing number of equations, most people choose the scaling equations based on judgment without statistical basis. However, the judgment varies for each person and hence the results are affected by the individual's judgment. In this work, it is shown that choosing magnitude scaling equations can create a bias of 0.23 U for the magnitude of completeness. In the present study, the difference in the magnitude of completeness estimate affects the estimation of a and b parameters which forms the basis for the seismic hazard studies and earthquake forecasting. The b parameter of the G–R relation plays a vital role in the estimation of maximum magnitude for each fault as well as PSHA studies for quantifying the magnitude uncertainty. The differences in the estimates of maximum magnitude and the PGA can be as high as 0.7 magnitude units and 0.2 g respectively. This could lead to significantly different results for the same region carried out by two different people. The understanding of the seismic status of the region could differ. This situation can be corrected by two ways: (1) Estimation of the maximum magnitude using the regional rupture characters which is opaque to the changes in the

earthquake catalogue or the choice of magnitude scaling equations used. The regional behavior of the rupture of faults is found to be a more effective tool in seismic hazard studies since it does not depend on the seismicity parameters. However, caution should be taken while demarking the length bins. It is suggested to distinguish the length bins after closely examining the regional rupture characters plot and also backed by expert judgments. (2) Standardizing the procedure for the selection of magnitude scaling equations. Statistical tests for model selection can be employed to select the magnitude scaling equations in a more methodical way. It is suggested to back the selection of magnitude scaling equations with statistical analysis so that this bias created by individual's perspective can be eliminated. The regional rupture character, furthermore, used to comment on the suitability of the magnitude scaling equations to the particular study area. The regional rupture characters help us to understand more about the robustness of the magnitude scaling equations in representing the actual rupture behavior of the seismic sources in the region.

7 Data and resources

All the earthquake data used in the present study was obtained from United States Geological Survey (USGS) (https://www.usgs.gov/natural-hazards/earthquake-hazards), European Mediterranean Seismological Centre (EMSC) (https://www.emsc-csem.org/Earthquake/?filter=yes), International Seismological Centre (ISC) (http://www.isc.ac.uk/iscbulletin/search/catalogue/), Advanced National Seismic System (ANSS) (https://earthquake.usgs.gov/monitoring/anss/), and Incorporated Research Institutions for Seismology (IRIS) (https://ds.iris.edu/ieb/). The details of the seismic sources used in the study were taken from SEISAT-2000 report published by Geological Survey of India (GSI).

Funding information The "Board of Research in Nuclear Sciences (BRNS)," Department of Atomic Energy (DAE), Government of India funded the project titled "Probabilistic seismic hazard analysis of Vizag and Tarapur considering regional uncertainties & Studies of Tripura Earthquake and Liquefied Soil" (Ref No. Sanction No. 36(2)/14/16/2016-BRNS-36016 dated July 1st, 2016).

References

Abramowitz M, Stegun IA (1972) Handbook of mathematical functions: with formulas, graphs, and mathematical tables, vol 55. Dover publications, New York, p 886

Anbazhagan P, Bajaj K, Patel S (2015a) Seismic hazard maps and spectrum for Patna considering region-specific seismotectonic parameters. Nat Hazards 78(2):1163–1195

Anbazhagan P, Bajaj K, Moustafa SS, Al-Arifi NS (2015b) Maximum magnitude estimation considering the regional rupture character. J Seismol 19(3):695–719

Anderson JG, Wesnousky SG, Stirling MW (1996) Earthquake size as a function of fault slip rate. Bull Seismol Soc Am 86(3):683–690

Baruah S, Baruah S, Bora PK, Duarah R, Kalita A, Biswas R, Gogoi N, Kayal JR (2012) Moment magnitude (MW) and local magnitude (ML) relationship for earthquakes in Northeast India. Pure Appl Geophys 169(11):1977–1988

Bender B (1983) Maximum likelihood estimation of b values for magnitude grouped data. Bull Seismol Soc Am 73(3):831–851

Bora DK (2016) Scaling relations of moment magnitude, local magnitude, and duration magnitude for earthquakes originated in northeast India. Earthq Sci 29(3):153–164

Castellaro S, Mulargia F, Kagan YY (2006) Regression problems for magnitudes. Geophys J Int 165(3):913–930

Das R, Wason HR, Sharma ML (2011) Global regression relations for conversion of surface wave and body wave magnitudes to moment magnitude. Nat Hazards 59(2):801–810

Das R, Wason HR, Sharma ML (2012a) Homogenization of earthquake catalog for northeast India and adjoining region. Pure Appl Geophys 169(4):725–731

Das R, Wason HR, Sharma ML (2012b) Temporal and spatial variations in the magnitude of completeness for homogenized moment magnitude catalogue for northeast India. J Earth Syst Sci 121(1):19–28

Das R, Sharma ML, Wason HR (2016) Probabilistic seismic hazard assessment for northeast India region. Pure Appl Geophys 173(8):2653–2670

Dasgupta S, Narula PL, Acharyya SK, Banerjee J (2000) Seismotectonic atlas of India and its environs. Geol Surv India

Gardner JK, Knopoff L (1974) Is the sequence of earthquakes in Southern California, with aftershocks removed, Poissonian? Bull Seismol Soc Am 64(5):1363–1367

Gasperini P, Lolli B, Vannucci G, Boschi E (2012) A comparison of moment magnitude estimates for the European—Mediterranean and Italian regions. Geophys J Int 190(3):1733–1745

Grünthal G, Wahlström R (2003) An M w based earthquake catalogue for central, northern and northwestern Europe using a hierarchy of magnitude conversions. J Seismol 7(4):507–531

Gutenberg B, Richter CF (1944) Frequency of earthquakes in California. Bull Seismol Soc Am 34(4):185–188

Hanks TC, Kanamori H (1979) A moment magnitude scale. J Geophys Res Solid Earth 84(B5):2348–2350

Hurukawa N, Maung Maung P (2011) Two seismic gaps on the Sagaing Fault, Myanmar, derived from relocation of historical earthquakes since 1918. Geophys Res Lett 38(1)

IS 1893-Part 1 (2016) Criteria for earthquake resistant design of structures: general provisions and buildings. Bureau of Indian Standards, New Delhi

Jin A, Aki K (1988) Spatial and temporal correlation between coda Q and seismicity in China. Bull Seismol Soc Am 78(2):741–769

Kijko A, Sellevoll MA (1989) Estimation of earthquake hazard parameters from incomplete data files. Part I. Utilization of extreme and complete catalogs with different threshold magnitudes. Bull Seismol Soc Am 79(3):645–654

Kijko A, Singh M (2011) Statistical tool for maximum possible earthquake magnitude estimation. Acta Geophys 59:674–700

Kolathayar S, Sitharam TG, Vipin KS (2012) Spatial variation of seismicity parameters across India and adjoining areas. Nat Hazards 60(3):1365–1379

Kramer SL (1996) Geotechnical earthquake engineering. Prentice–Hall international series in civil engineering and engineering mechanics. Prentice-Hall, New Jersey

Last M, Rabinowitz N, Leonard G (2016) Predicting the maximum earthquake magnitude from seismic data in Israel and its neighboring countries. PloS One 11(1):e0146101

Lolli B, Gasperini P, Vannucci G (2014) Empirical conversion between teleseismic magnitudes (mb and Ms) and moment magnitude (M w) at the Global, Euro-Mediterranean and Italian scale. Geophys J Int 199(2):805–828

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(B1)

NDMA (2010) Development of probabilistic seismic hazard map of India; technical report by National Disaster Management Authority, Government of India

Omori F (1894) On the after-shocks of earthquakes. J Coll Sci Imp Univ Tokyo 7:111–200

Osher B (1996) Statistical estimation of the maximum magnitude and its uncertainty from a catalogue including magnitude errors. In: Earthquake hazard and risk. Springer, Dordrecht, pp. 25-37

Pandey AK, Chingtham P, Roy PNS (2017) Homogeneous earthquake catalogue for northeast region of India using robust statistical approaches. Geomat Nat Haz Risk 8(2):1477–1491

Reasenberg P (1985) Second-order moment of central California seismicity, 1969–1982. J Geophys Res Solid Earth 90(B7): 5479–5495

Rhoades DA (1996) Estimation of the Gutenberg-Richter relation allowing for individual earthquake magnitude uncertainties. Tectonophysics 258(1–4):71–83

Richter CF (1935) An instrumental earthquake magnitude scale. Bull Seismol Soc Am 25(1):1–32

Scordilis EM (2006) Empirical global relations converting M S and m b to moment magnitude. J Seismol 10(2):225–236

Sil A, Sitharam TG, Kolathayar S (2013) Probabilistic seismic hazard analysis of Tripura and Mizoram states. Nat Hazards 68(2):1089–1108

Sitharam TG, Sil A (2014) Comprehensive seismic hazard assessment of Tripura and Mizoram states. J Earth Syst Sci 123(4): 837–857

Stromeyer D, Grünthal G, Wahlström R (2004) Chi-square regression for seismic strength parameter relations, and their uncertainties, with applications to an M w based earthquake catalogue for central, northern and northwestern Europe. J Seismol 8(1):143–153

Thingbaijam KKS, Nath SK, Yadav A, Raj A, Walling MY, Mohanty WK (2008) Recent seismicity in northeast India and its adjoining region. J Seismol 12(1):107–123

Tinti S, Mulargia F (1985) Effects of magnitude uncertainties on estimating the parameters in the Gutenberg-Richter frequency-magnitude law. Bull Seismol Soc Am 75(6):1681–1697

Uhrhammer RA (1986) Characteristics of northern and central California seismicity. Earthq Notes 57(1):21

Wason HR, Das R, Sharma ML (2012) Magnitude conversion problem using general orthogonal regression. Geophys J Int 190(2):1091–1096

Wells DL, Coppersmith KJ (1994) New empirical relationships among magnitude, rupture length, rupture width, rupture area, and surface displacement. Bull Seismol Soc Am 84(4):974–1002

Woessner J, Wiemer S (2005) Assessing the quality of earthquake catalogues: estimating the magnitude of completeness and its uncertainty. Bull Seismol Soc Am 95(2):684–698

Yin A, Harrison TM (2000) Geologic evolution of the Himalayan-Tibetan orogen. Annu Rev Earth Planet Sci 28(1):211–280