ORIGINAL PAPER

Quantification of source-to-site distance uncertainty in ground motion models

Saman Yaghmaei-Sabegh[1] · Mehdi Ebrahimi-Aghabagher[1]

Abstract

Among different input data, source-to-site distance plays a major role in the results of ground motion models (GMMs). In order to determine the source-to-site distance, geometric characteristics of the seismic source need to be specified. This can be challenging when the seismic source is not known thoroughly. Empirical relationships themselves which are used to determine the geometric characteristics of seismic sources contain large degree of uncertainty. In this paper, a simple algorithm based on Monte Carlo (MC) simulation method which quantifies the uncertainties in distance metric and geometric characteristics of seismic sources is proposed. The jointly effects of magnitude and distance are considered in the proposed algorithm for uncertainties modeling. Also, this algorithm has been used to quantified errors resulted from inputting inaccurate source-to-site distance metrics in the GMMs. NGA-West2 global GMMs and event-specific isotropic and non-isotropic GMMs are used in the analysis. The results demonstrate that the uncertainty in the measurement of different source-to-site distance definitions depends on the magnitude and on the location of site with respect to the seismic source. It is also observed that the distance measurement uncertainty has a direct effect on the outcomes of GMMs. GMMs' coefficient of variation maps demonstrate that the amount of uncertainty is higher around the fault; for large magnitudes, outcome variation of GMMs reaches as high as 40% of the average predictions. The coefficient of variation in GMM results decreases with the increase in distance metrics considered where at a distance beyond 30 km, the coefficient of variation of GMM estimates drops from 40 to 10%.

Keywords Uncertainty · Ground motion model · Monte Carlo simulation · Fault geometric characteristic

1 Introduction

Uncertainty, in different forms, is an inevitable part of engineering analysis. Uncertainties are responsible for seismic risk evaluation of structural systems and installations. The uncertainties are categorized as either aleatory or epistemic. Aleatory uncertainty, which

✉ Saman Yaghmaei-Sabegh
  s_yaghmaei@tabrizu.ac.ir

[1] Department of Civil Engineering, University of Tabriz, Tabriz, Iran

is not reducible, is defined randomly or as a natural congenital variable. Epistemic uncertainty stems from a lack of knowledge or incomplete information and, unlike aleatory uncertainty, is reducible. In fact, epistemic uncertainty can be reduced by increasing data and information. With all this, distinguishing between aleatory and epistemic uncertainty is not a straightforward task and from a scientific point of view, uncertainty type identification is the main capability of the analyst in reducing of uncertainties (Der Kiureghian and Ditlevsen 2009). Epistemic uncertainties associated with probability seismic hazard analysis (PSHA) can be classified into two categories: (i) earthquake rupture prediction and (ii) ground motion characteristics (Bradley 2009).

In the first category where uncertainties are related to earthquake prediction, estimating the occurrence frequency of earthquake for a fault/faults of a region and its magnitude is very difficult and dependent on assumptions (Murray and Segall 2002). In conventional PSHA, earthquake occurrence is considered randomly with respect to time. Assuming random occurrence, the probability of an earthquake in the specified time intervals is constant and does not change over time which is inconsistent with elastic rebound theory (Reid 1911). The uncertainty of recurrence model can be categorized into the uncertainty in magnitude and time of occurrence (Kramer 1996). In determining the magnitude of an earthquake, Gutenberg and Richter (1944) obtained a linear relationship between the magnitude and total number of earthquakes in any given region and the time period of at least that magnitude, which is known as Gutenberg–Richter law. Then this relationship was transformed into the modified Gutenberg–Richter law by controlling the magnitude of earthquakes to specific values (McGuire and Arabasz 1990). Many studies have been carried out in this subject area to reduce the uncertainties. The main focus of these studies was frequently on determination of maximum magnitudes, b-value prediction and illustration of seismicity of the study area (Smit and Kijko 2016; Kijko et al. 2016; Beirlant et al. 2018). In the case of time uncertainty, several recursive models have been proposed by researchers. For this case, Markov's recurrence model presented by Knopoff (1971) using Markov chain can be instanced. In addition, Patwardhan et al. (1980) introduced the pseudo-Markov model. Most of time-dependent recurrence models are based on the renewal process. In these models, the recurrence interval of an earthquake can have different distributions and the probability of an earthquake within a given interval depends on the elapsed time since the last event. Hagiwara (1974), Utsu (1984) and Nishenko and Buland (1987) utilized Weibull, Gamma and log-normal distributions in recursive models, respectively.

The second category of uncertainty is related to modeling ground motion uncertainty. For a specific earthquake, ground motion amplitude is described by GMMs, which is a function of earthquake magnitude, source-to-site distance, site conditions and occasionally some other parameters (Bommer et al. 2010; Wang 2011; Tsang et al. 2011; Yaghmaei-Sabegh 2012; Stewart et al. 2015; Douglas 2017). GMMs are obtained using statistical regression on data recorded at the stations. Despite the expansion of the global seismic network, there is no adequate data on near field and large magnitude earthquakes which necessitate the presentation of new models according to updated data (Beven et al. 2018). Due to the assumptions and simplifications in mathematical modeling, GMMs contain epistemic uncertainty. The uncertainty can be observed in the model's inaccurate form, selection of particular data, error in input data, and statistical error in the estimation of constants (Foulser-Piggott 2014). Youngs (2006) and Stafford et al. (2009) showed that the uncertainty caused by the inaccurate form and the selection of particular data can be calculated by the logic tree. Using the logic tree, several predictions of seismic intensity measurements are made through a set of GMMs, for a given scenario. It is also shown that although the use of a set of GMMs could estimate epistemic uncertainties (Yuongs 2006; Stafford et al. 2009),

yet part of the variables remain as epistemic uncertainty (Foulser-Piggott 2014). Foulser-Piggott (2014) provides a quantitative approach for modeling of the input data uncertainty in the median prediction of GMMs by adopting the MC approach for the three variables: moment magnitude (M_w), shear wave velocity at a depth of 30 meters ($V_{S_{30}}$), and the closest distance to fault rupture plane (R_{rup}).

Source-to-site distance measurement is difficult in areas with low seismicity or where the dimensions and geometry of the faults are unclear. In these regions such as Australia, Peninsular Malaysia, and the island of Sri Lanka, using predefined areal source zones is still the standard of practice. The number of the historical earthquake and seismicity data are sparse; therefore, the amount of information available is not enough to have individual faults to be modeled as precisely in a PSHA (Lam et al. 2016; Abrahamson 2006). Therefore, in this paper, a method has been developed to quantify the distance uncertainties where the geometry of the faults is unknown. A simple algorithm based on the MC simulation method which quantifies the uncertainties in distance metric (R_{epi}, R_{rup}, and R_{jb}) and geometric characteristics of seismic sources is proposed and errors in results of the GMMs from these uncertainties have been calculated. Briefly, this paper main focus is the distance metrics uncertainty measurement and its quantification. Various factors engender source-to-site distance measurement uncertainty, which include the following:

1. The difference in fault rupture geometry.
2. When using a GMM, depending on the size of the project site, a question may arise about the selection of a point from which the distance to the source ought to be measured. This is more common in large-scale projects located at a close distance to a fault. For instance, the Oakland bridge project in San Francisco is several kilometers long. Hence, the earthquake waves along the course of this project will have different characteristics at the beginning and the end of the bridge (Moss 2009).
3. Different definitions for distance measurement are used in GMMs. Among the most common definitions, the closest horizontal distance to the vertical projection of fault rupture (R_{jb}), the closest distance to fault rupture plane (R_{rup}), the closest distance to the epicenter (R_{epi}), and the distance to the hypocenter (R_{hypo}) can be mentioned (Abrahamson and Shedlock 1997). There is also no agreement on the most appropriate definition of distance measurement, which indicates uncertainty about the comparative accuracy of these definitions (Moss 2009).

MC simulation method as a powerful uncertainty modeling tool is a computational algorithm that, by the random selection of input parameters, solves the problem and stores the results. Ultimately, using the stored outcomes, the overall result is reported by the method. Production of random numbers, based on the probability distribution assigned for integrating various uncertainties, is one of the main tasks of the MC method (Yazdani et al. 2012). The main advantages of the MC method are its ability to control the uncertainty of the parameters involved in the PSHA model and to clearly identify the effects of each parameter in the final result (Musson 1999). Also, this method has a simple concept and has been used by several researchers to evaluate PSHA results (Yazdani et al. 2017; Pavel and Vacareanu 2017; Assatourians and Atkinson 2013).

In the following, a simple MC-based simulation method is presented to evaluate the distance measurement uncertainties in Sect. 2. To quantify the uncertainties of the defined input data and the assumptions made, including fault geometry and the considered sites, we propounded an example in Sect. 3 of the paper and the GMMs that used in this paper
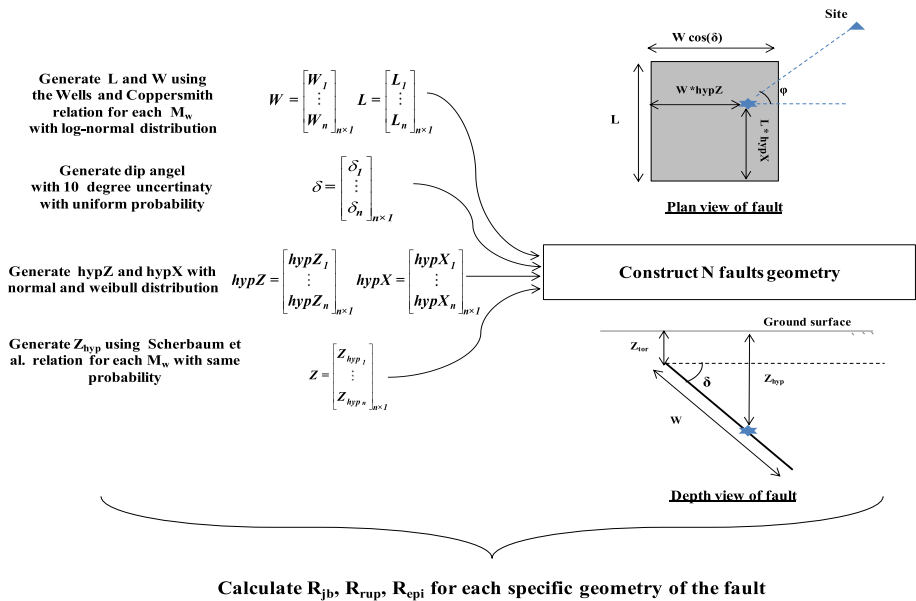
Fig. 1 Source-to-site distance simulation process

are introduced in this section. Section 4 consists of two subsections; the outcomes of three sites and the standard deviation maps are presented in the first and second subsections, respectively. Simulation results of the sites are presented for small, medium and large magnitude earthquake. In Sect. 5, we utilized the simulated data of Sect. 4 and investigated their impacts upon the GMMs. Section 5 consists of two subsections; the output results of different GMMs are obtained for each point and subsequently, their distributions are depicted for the three sites. Finally, Sect. 6 summarizes the outcomes of the paper.

2 Source-to-site distance simulation algorithm

The geometric of fault is an important factor for determining the source-to-site distance that is usually not known. The uncertainty of the fault position and the determination of the input parameters of the GMM are adopted as the epistemic uncertainty which influences the results of PSHA analysis. Among the effective geometrical factors which influence measuring the source-to-site distance with respect to the definition of distance, the following factors can be mentioned:

1. Start point of rupture on the fault plane (hypocenter location).
2. Length and width of fault rupture.
3. Rupture depth.

In the proposed algorithm, uncertainty of the various definitions of source-to-site distance is simulated. The detail of the algorithm used in this paper is shown in Fig. 1. The proposed

algorithm is very simple and consists of several parallel loops which are considered the uncertainties of the mentioned factors to determine of distance metric.

In the first loop, uncertainty length and width of fault are considered. These two parameters are related to the moment magnitude scale by the Wells and Coppersmith (1994) model (see Eqs. 1, 2). Therefore, several numbers of the length and width are generated in this loop. The generated length and width have a log-normal distribution with standard deviations 0.16 and 0.26, respectively ($\sigma_{\log_{10} L} = 0.16$, $\sigma_{\log_{10} W} = 0.26$.

$$\log_{10} W = -1.01 + 0.32 M_w \tag{1}$$

$$\log_{10} L = -2.44 + 0.59 M_w \tag{2}$$

where L and W are length and width of the rupture fault, respectively, and M_w is moment magnitude.

In the second loop, uncertainty quantification of the fault dip (δ is considered. In this path, the uncertainty of fault dip results proposed by Kilb and Hardebeck (2006), Hayes and Wald (2009) and Foulser-Piggott (2014) was used, and dip variation was considered 10°. In this loop, the random desired number of fault dip angle with the same probability is generated. For instance, if the fault dip is 70°, the angles which are generated in this loop are between 60° and 80° (i.e., 70° $\pm$ 10°) with the same probability.

One of the effective factors in the estimation of source-to-site distance R_{epi} and R_{hyp} is the hypocenter position on the fault plane. In order to determine R_{hyp}, the realization of hypocenter proposed by Mai et al. (2005) is utilized. Mai et al. (2005) represented that the relative position of hypocenter along the fault width (hyp Z) has a normal distribution with mean of 0.5 and standard deviation of 0.23, and the relative position of hypocenter along the fault length (hyp X) follows a Weibull distribution with a scale parameter of 0.626 and shape parameter of 3.921. Regarding the length and width of the fault, the desired number of hypocenter position with the aforementioned distribution is generated. In addition to the above realizations, the depth of hypocenter (Z_{hyp}) should be determined. The depth of hypocenter according to Eq. (3), which is presented by Scherbaum et al. (2004), is a function of the moment magnitude and its standard deviation is equal to 4 km. The minimum and maximum values of hypocenter depth (Z_{hyp}) are considered 4 and 16 km, respectively (Scherbaum et al. 2004).

$$Z_{hyp} = 5.63 + 0.68 M_w \tag{3}$$

With regarding standard deviation, the upper and lower limit, the desired number of Z_{hyp} is generated with the same probability.

In order to identify the realizations mentioned in the previous loop, we can also calculate the other parameters of the fault, including depth-to-top of rupture (Z_{tor}). Based on the modified Kaklamanos et al., (2011) relation, Z_{tor} can be calculated through Eq. (4):

$$Z_{tor} = \max((Z_{hyp} - hypZ. W. \sin(\delta)), 0) \tag{4}$$

After realizations of the fault geometry, we can calculate different definitions of source-to-site distance metrics.

In sum, the simulation algorithm for each moment magnitude can be carried out in the following steps

1. Loop 1: n number of length and width which is generated by Eqs. 1 and 2 ($w_{n\times1}$ and $L_{n\times1}$).
2. Loop 2: n number of the dip angle with uniform probability is generated ($\delta_{n\times1}$).

Fig. 2 Strike slip fault and sites configuration
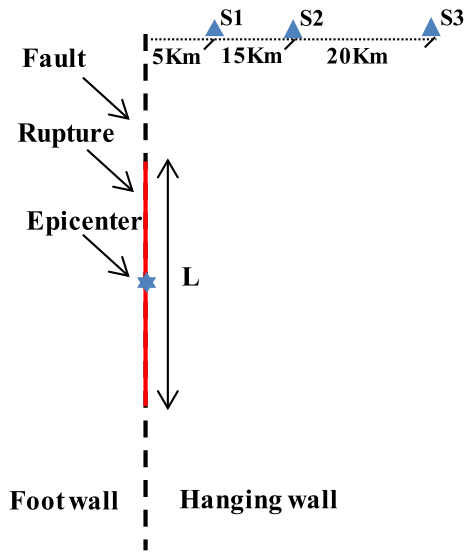
3. Loop 3: n number of hypZ and hypX regarding the probability distribution of each hypocenter is generated (hyp$X_{n \times 1}$ and hyp$Z_{n \times 1}$).

4. Loop 4: n number of Z_{hyp} and Z_{tor} regarding existing distribution is generated ($Z_{n \times 1}$, $Z_{\text{tor}_{n \times 1}}$).

By the combination of four parallel loops, n number of the fault geometry is simulated. The process of simulation is shown in Fig. 1.

3 Example set up

To quantify the epistemic uncertainty of GMM due to the uncertainty of source-to-site distance, a strike-slip fault with the length of 60 km is considered according to Fig. 2. It is assumed that the selected sites in this example have been located on the bedrock; hence not soil response is relevant. The results of the MC simulation will be presented in two parts. In the first part, the simulation results of distance determinations are presented and discussed and in the second part, simulated data in the first part is used as GMM input to determine the uncertainties in the GMM outputs.

Nine GMMs have been used in this paper: the four NGA08 GMMs, Abrahamson and Silva (2008), Boore and Atkinson (2008), Campbell and Bozorgnia (2008), and Chiou and Youngs (2008), in which they are abbreviated as AS08, BA08, CB08, and CY08, respectively. Moreover, three GMMs: Danciu and Tselentis (2007), Nowroozi (2005), Ambraseys et al. (2005), which are based on epicenter distance and abbreviated as DT07, N05, ADSS05 in this paper, are used. Finally, two event-based GMMs named isotropic (Iso) and non-isotropic (Non-iso) are used in the analysis. The main characteristics of Iso and Non-iso GMMs are that: (i) They are a random function that means they have an event-specific parameter. The GMMs are fitted to a given event based on corresponding records. (ii) The level of ground motion is not the same in all directions in the non-isotropic GMM.
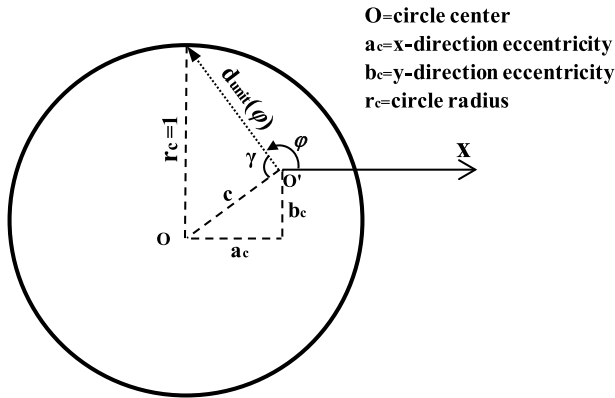
Fig. 3 Eccentric circle unit-isoline (Reconstructed from Raschke 2013)

The general form of these GMMs that are used to nonlinear regression as follows (Raschke 2013):

$$\text{Ln(PGA)} = \theta_0 - \theta_1 \ln(R) - \theta_2 R \quad \text{and} \quad R^2 = r^2 + h^2 \tag{5}$$

where PGA is peak ground motion acceleration, h is hypocenter depth, and r is the horizontal distance between the site and the epicenter. θ_0, θ_1, θ_2 are the constant coefficients obtained from a nonlinear regression. To consider the non-isotropic conditions, R is calculated utilizing Eq. 6 (Raschke 2013):

$$R = \sqrt{r^2 / d_{\text{unit}}^2(\varphi) + h^2} \tag{6}$$

The radius function $d_{\text{unit}}(\varphi)$ determines the unit-isoline with an azimuth φ of local polar coordinates. There are many types of unit-isoline that can be applied in the non-isotropic GMM which include area π equal to the unit circle of angle function. Different unit-isolines can be combined by the sum $d_{\text{unit}}^2(\varphi) = \sum a_i d_{i,\text{unit}}^2(\varphi)$ with weighting $0 \leq a_i \leq 1$ and $\sum a_i = 1$. Circle with eccentricity has been considered herein as the unit-isoline in the non-isotropic GMM which is shown in Fig. 3. If the circle with eccentricity is used, parameters of eccentricity (a_c and b_c), along with θ_0, θ_1, and θ_2, should be estimated. The main characteristics of the GMMs are listed in Table 1. The selected events to represent small, medium, and large earthquakes are Big Bear 02, California 2001 ($M_w = 4.56$), Sierra Madre, California 1991 ($M_w = 5.61$) and Loma Prieta, California 1989 ($M_w = 6.93$). For the three selected events, isotropic and non-isotropic GMMs parameters are presented in Tables 2 and 3, respectively, which obtained from the nonlinear regression.

4 Simulated results

Using the proposed algorithm, 100,000 realizations for the fault geometry are generated. According to the fault geometry, the desired distance metrics will be calculated as well. Fault rupture could occur in all parts of the fault; however, to reduce the simulated variables only one position of the fault rupture is considered so that the center of the fault

Table 1 Summary of GMMs used in this study

GMM	Abbreviation	Area	Number of records	Number of event	M_{min}–M_{max}	R_{min}–R_{max} (Km)
Danciu and Tselentis (2007)	DT07	Greece	335	151	4.5–6.9	<136
Nowroozi (2005)	N05	Iran	279	45	3–7.4	1.2–250
Ambraseys et al. (2005)	ADSS05	Europe and the Middle East	595	–	5–7.6	0–100
Abrahamson and Silva (2008)	AS08	Worldwide	2675	129	5–8.5	0–200
Boore and Atkinson (2008)	BA08	Worldwide	1574	58	5–8	0–400
Campbell and Bozorgnia (2008)	CB08	Worldwide	1561	64	4.3–7.9	0–200
Chiou and Youngs (2008)	CY08	Worldwide	1950	125	4.27–7.9	0–70
Isotropic	Iso	Event-specific	Depend on the event	1	Depend on the event	Depend on the stations distance
Non-isotropic	Non-iso	Event-specific	Depend on the event	1	Depend on the event	Depend on the stations distance

Table 2 Parameters of isotropic GMM Reconstructed from Raschke (2013)

Earthquake	θ_2	θ_1	θ_0	h	Residual variance
Big Bear-02, California 2001	0	1	−0.28	9.1	0.222
Sierra Madre, California 1991	0	1	1.747	12	0.088
Loma Prieta, California 1989	0	1	2.17	17.48	0.257

Table 3 Parameters of non-isotropic GMM Reconstructed from Raschke (2013)

Earthquake	θ_2	θ_1	θ_0	h	a_c	b_c	Residual variance
Big Bear-02, California 2001	0	1	−0.11	9.1	−0.43	−0.24	0.217
Sierra Madre, California 1991	0	1	1.59	12	−0.30	0.42	0.077
Loma Prieta, California 1989	0	1	2.16	17.48	0.13	−0.05	0.253

rupture length corresponds to the fault center. In fact, it is symmetric with regard to the fault center. The maximum rupture length due to limitations of the fault length is 60 km.

4.1 Simulation results for site S_1, S_2, and S_3

Using the simulated algorithm, the different source-to-sites distance metric distributions (R_{epi}, R_{jb}, R_{rup}) are calculated for three S_1, S_2, and S_3 sites. To quantify and compare the variation of distance distribution, the coefficient of variation is used in accordance with Eq. (7) (Everitt 1998).

$$C_v = \frac{\sigma}{\mu} \tag{7}$$

In this regard, C_v is the coefficient of variation, σ and μ are the standard deviation and the mean of the simulated data, respectively.

Figure 4 illustrates the values of C_v for each magnitude in the semi-logarithmic scale for the three sites, S_1, S_2 and S_3. The C_v values for R_{epi}, R_{jb}, and R_{rup} depend on the magnitude and the site position with respect to the fault. C_v coefficient of all distance measures will increase by increasing magnitude and for a given magnitude, distance increasing will reduce C_v.

For small and medium magnitude levels, C_v coefficient of the three distances metrics is closed to each other in all of the sites. The trend of changes is the same in these sites, R_{jb} and R_{rup} included the least and the largest C_v, respectively.

The C_v coefficient has a different trend in the large magnitude, and it is also sensitive to measure distance. For the S_1 site, C_v coefficient of different source-to-site distance metrics is very close to each other, and by increasing the distance (i.e., in the S_2 and S_3 sites), the C_v coefficient takes large values for the R_{epi} distance measures. The mean values estimated for R_{jb} and R_{rup} in the S_1 site (μ in Eq. 7) are lower than the mean value of R_{epi}, and due to this, their C_v coefficients are greater than R_{epi} so that by increasing the distances, the mean value

Fig. 4 C_v of different source-to-site distance metrics (R_{jb}, R_{rup}, R_{epi}) at site S_1, S_2 and S_3
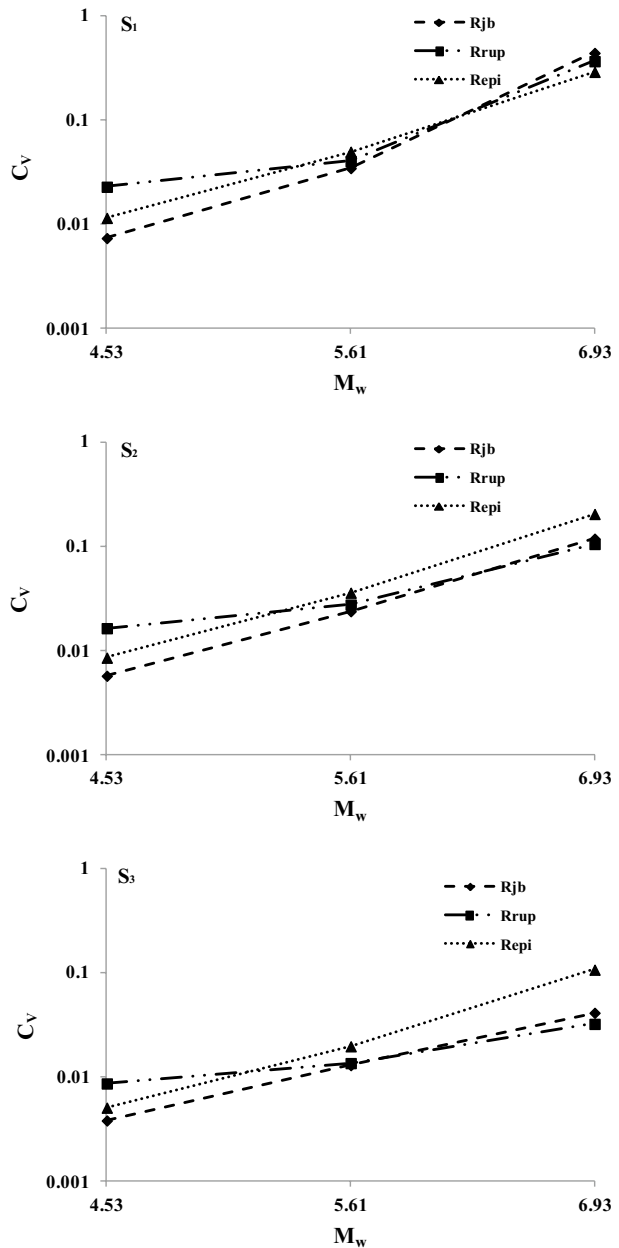
of these three distance metrics approaches each other, and the C_v coefficient of the R_{epi} is larger than others.

As mentioned before, by increasing magnitude, C_v would be increased either; however, the rate of increasing is not the same for all distance definitions. The increasing rate of R_{epi} is higher than that of other definitions with increased magnitude. By
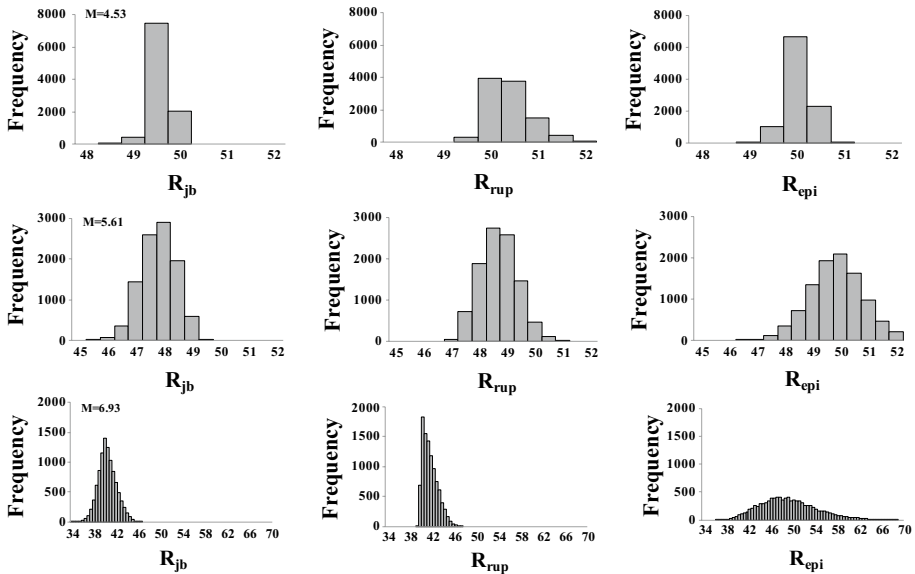
Fig. 5 Distance simulation distribution (R_{jb}, R_{rup}, R_{epi}) in different magnitude at site S_3

increasing magnitude, the length and width of the fault rupture and the points that can be the epicenter will increase; due to this process, the variation coefficient of the rupture increases as well.

The increasing process of the C_v for R_{rup} in large magnitude in contrary to R_{epi} will reduce and the reason for this is reaching the rupture level to ground's surface in large magnitude. In this case, most Z_{tor} values in the simulation algorithm are equal to zero and thus the C_v of R_{rup} is lower than other distance definitions. The increasing process of R_{jb} is also constant and C_v increases with increasing magnitude.

Figure 5 shows the distribution of simulated data for R_{jb}, R_{rup}, and R_{epi} considering different magnitudes at site S_3. In this case, R_{jb} and R_{epi} are bounded to the minimum and maximum of simulated data, respectively. The range of simulated data depends on the number of parameters involved in the definition of distance type. In the proposed algorithm, R_{jb} depends only on the length, width and dip angle of the fault; however, in addition to the mentioned parameters for R_{jb}, the epicenter location on the fault plane will additionally affect the measure of R_{epi}. Uncertainties in the length, width, dip and epicenter location on the fault plane are considered for R_{epi}. Also, length, width, dip and the Z_{tor} are the parameters that are considered in the calculation of R_{rup}. By increasing magnitude, the length and width of the rupture will increase and the uncertainty associated with the geometry of the fault increases and the extent of variation in all three types of distances increases.

4.2 Distance uncertainty map

The proposed algorithm is accomplished for all sites around the fault. The standard deviation values associated with three distance metrics (R_{epi}, R_{rup}, and R_{jb}) at the Mw6.93 event are illustrated in Fig. 6. For this case, the standard deviation trend depends on the distance
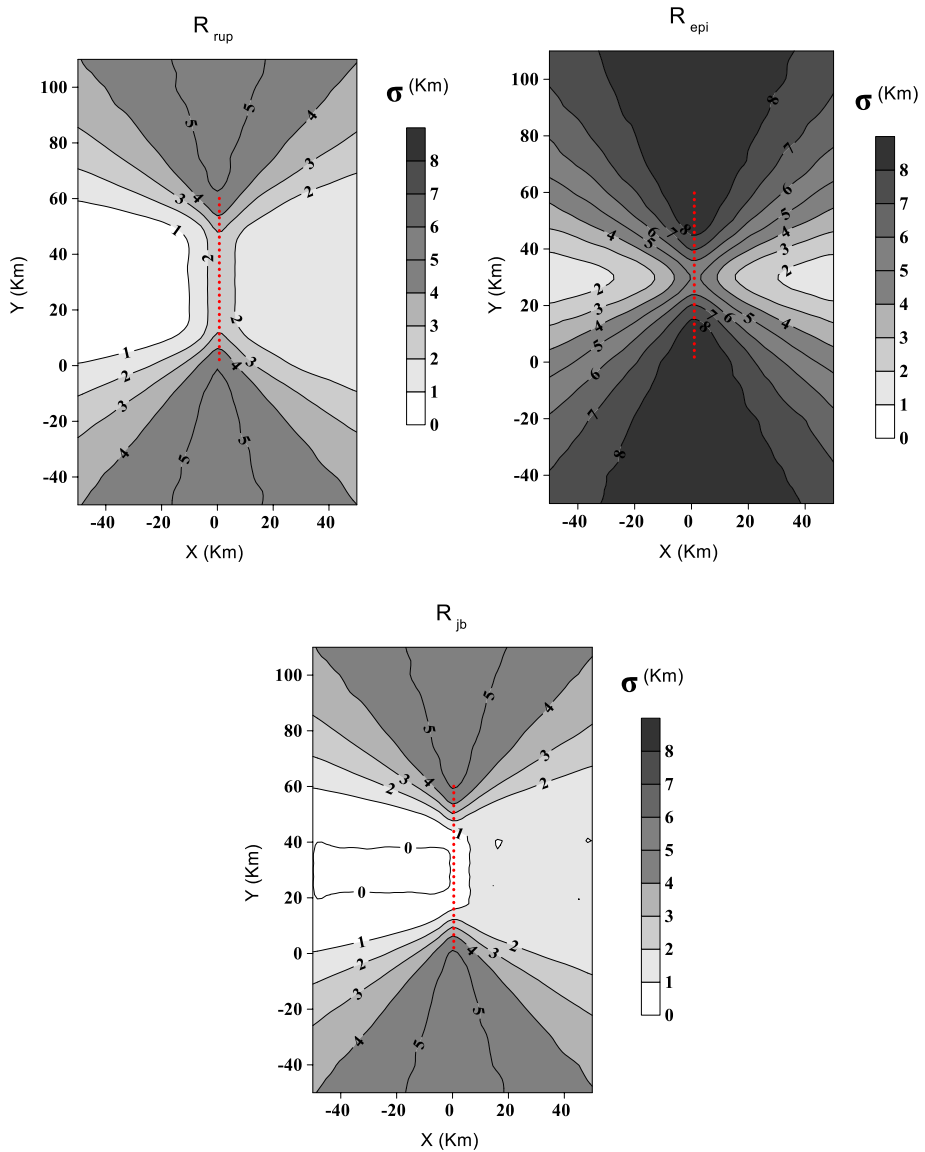
Fig. 6 The standard deviations contour of different source-to-site distance metrics for the $M_w 6.93$ event

definitions (R_{epi}, R_{jb}, R_{jb}) and the main assumption of this example as well. (The rupture has symmetric evolution with respect to fault plane.) Therefore, the standard deviation of simulated data is the lowest for all distance definitions in the middle of the fault.

The maximum variation of the standard deviation is related to R_{epi}; by changing the angle to the fault center, the standard deviation values are changed and the highest standard deviation is about 9 km, which is along the fault strike.

The variation of the standard deviation of two distances (R_{rup} and R_{jb}) is similar to another, and its lowest amount occurs in the middle of the fault. Because in the defined scenario, fault rupture is located in the middle of the fault and due to Eq. 2, the mean length of the rupture with the magnitude scale of 6.93 is about 45 km; thus, in the middle of fault region, the amount of R_{jb} is constant. The small standard deviation difference in the hanging wall and footwall is due to the 10° of measurement error on the fault dip that is considered randomly in the simulated algorithm. Because the dip variation has no effect on the footwall part, the standard deviation of R_{jb} is equal to zero. The maximum standard deviation for R_{jb} is calculated on the top and bottom of the fault strike. Because rupture level has reached the ground surface, contour variation in R_{rup} is similar to R_{jb}.

5 Distance uncertainty effect on GMMs output

5.1 GMMs simulation results for site S_1, S_2 and S_3

In this section, simulated data in the previous section are used as input in GMMs. According to Eq. (7), the C_v for all GMMs is calculated for different magnitudes. Figure 7 shows the PGA distribution of the GMMs at the site S_3. According to Table 4, Non-iso, Iso, ADSS05, N05, and DT07, which are dependent on R_{epi} have the same behavior and also the suffering of their changes is the same. Maximum and minimum of C_v are DT07 ($C_v = 0.1266$) and non-isotropic ($C_v = 0.0796$) GMMs at site S_3, respectively. As shown in Fig. 7, the PGA distribution for AS08 and CY08 is different when compared with other GMMs. According to sensitivity analysis for the input values of the NGA08 GMMs, Z_{tor} uncertainty has a more effect on the outputs of these GMMs.[1] By considering a constant value for Z_{tor}, the behavior of these GMMs is similar to each other. Figure 8 shows that after eliminating the Z_{tor} uncertainty in the simulation process, the distribution of these two GMMs (AS08 and CY08) is similar to other GMMs and C_v coefficient of these GMMs is reduced.

Figure 9 shows the C_v of the NGA08 GMMs for two cases: one with Z_{tor} uncertainty and other without Z_{tor} uncertainty. For the first case, C_v of AS08 and CY08 is large for all sizes and converges to another GMMs' C_v with an increase in magnitude. This trend depends on the variation of Z_{tor} that is large for small magnitude and small in large magnitude. As shown in Fig. 9b when Z_{tor} uncertainty is excluding, results decrease in C_v for small and medium magnitude. In both cases, C_v of PGA increases with increasing the magnitude size and decreases with increasing distance from the rupture, for all GMMs.

5.2 GMMs uncertainty map

Simulated data in Sect. 4.2 are used as input for all sites around the fault. The median and standard deviation of all GMMs output are computed for all sites, and C_v is calculated using Eq. (7). Figure 10 illustrates the C_v of PGA for the Mw6.93. As shown in Fig. 10, C_v takes large values in the near-fault region and decreases by distance increasing. The C_v is large along strike in comparison with perpendicular direction to the fault. Due to the assumption behind this example (rupture is located in the middle of the fault), R_{epi} has

---

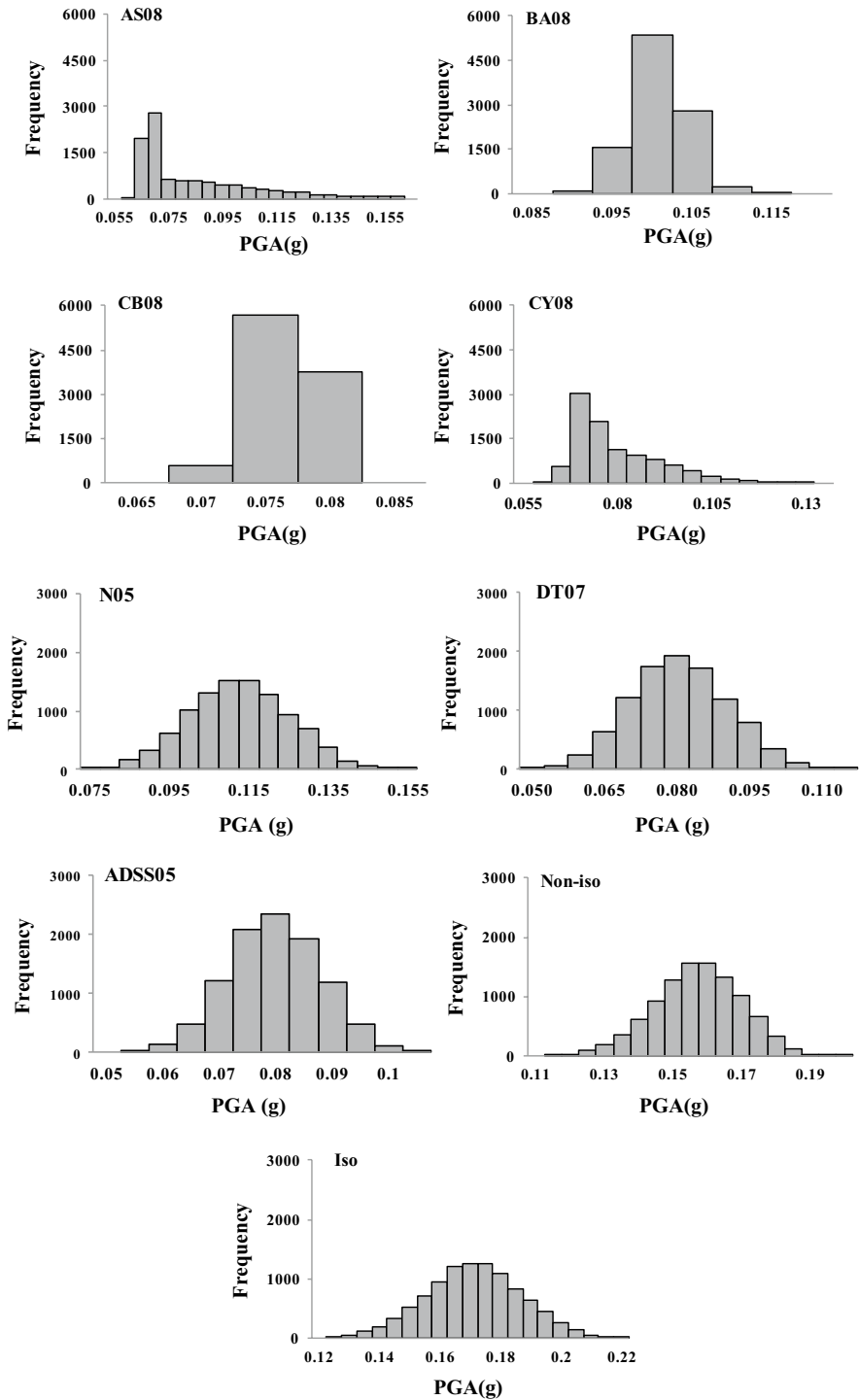[1] Sensitivity analysis not reported here for sake of brevity.

Fig. 7 GMMs outputs distribution (PGA) at site S_3 for the Mw6.93 event

Table 4 Considered different distance definition in GMMs

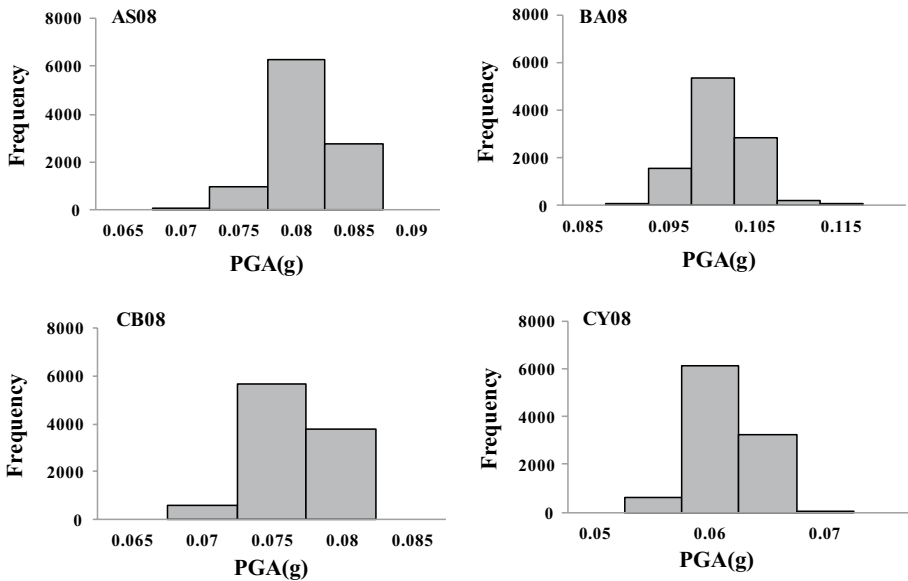| GMM	R_{rup}	R_{jb}	R_x	Z_{tor}	R_{epi}
DT07	–	–	–	–	✓
N05	–	–	–	–	✓
ADSS05	–	–	–	–	✓
AS08	✓	✓	✓	✓	–
BA08	–	✓	–	–	–
CB08	✓	✓	✓	✓	–
CY08	✓	✓	✓	✓	–
Iso	–	–	–	–	✓
Non-iso	–	–	–	–	✓

Fig. 8 GMMs outputs distribution (PGA) at site S_3 for the Mw6.93 without considering Z_{tor} uncertainty

smaller variation in this region which results small variation for PGA. The C_v of Iso and Non-iso GMMs are close to each other and three GMMs of N05, DT07 and ADSS05 are showing similar results. Maximum and minimum of C_v are N05 ($C_v = 0.4$) and Non-iso ($C_v = 0.2$) GMMs, respectively.

The C_v of NGA08 models have a different trend when compared to the other GMMs and have an asymmetrical pattern with unequal values for hanging wall and footwall. BA08 which depends on R_{jb} has a small C_v and in the perpendicular direction to rupture, R_{jb} has a small variation which results small variation in C_v. Similar to other GMMs C_v has the largest value along strike and decrease with an increasing distance. This GMM has the largest C_v of 0.32.
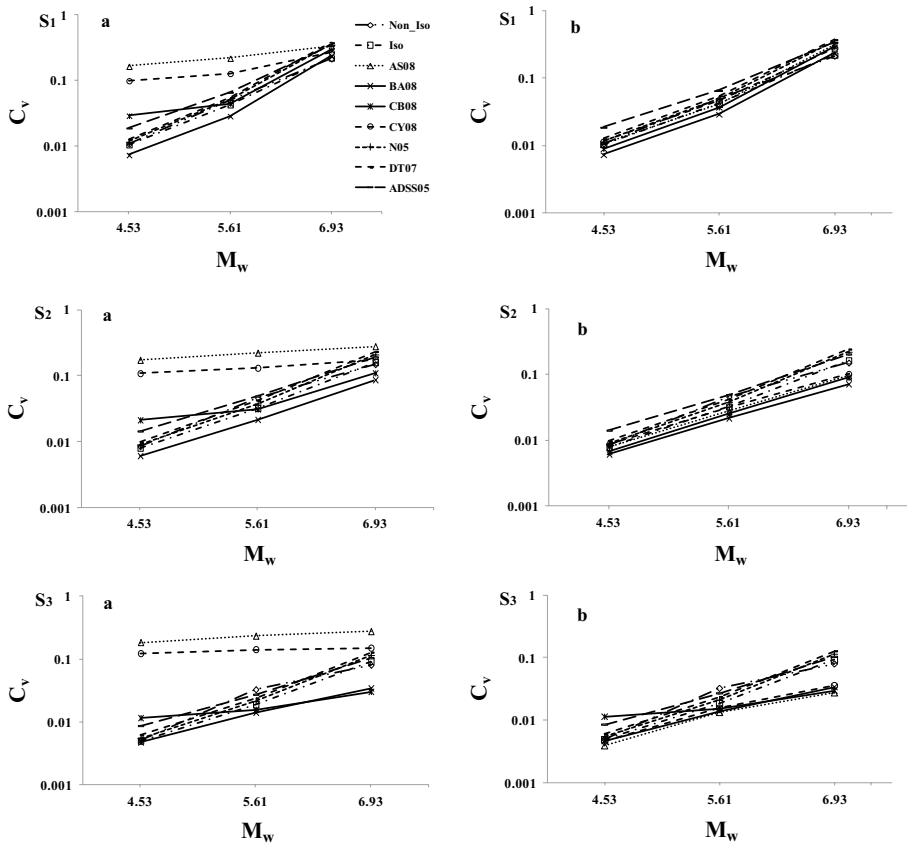
Fig. 9 C_v of GMMs in different magnitude at site S_1, S_2 and S_3 **a** with Z_{tor} uncertainty **b** without Z_{tor} uncertainty

CB08 GMM which depends on R_{rup}, R_{jb}, Z_{tor}, and fault wide has a different trend within comparison to BA08 GMM. In this GMM R_{rup} has the largest contribution to the variation in the results and in regions with smaller R_{rup} variation, the C_v is also small. C_v does not have a symmetrical trend with respect to the strike direction due to the hanging wall and footwall effect. In this GMM, the maximum C_v equal to 0.32 in along the strike.

AS08 and CY08 GMMs are similar to CB08 GMM and the considerable variation in the results are coming variety of R_{rup} and Z_{tor}. These GMMs have different trends with respect to other GMMs as C_v is greater at large distances with respect to the other GMMs. In these GMMs the maximum C_v equal to 0.32 similar to other NGA08 GMMs.
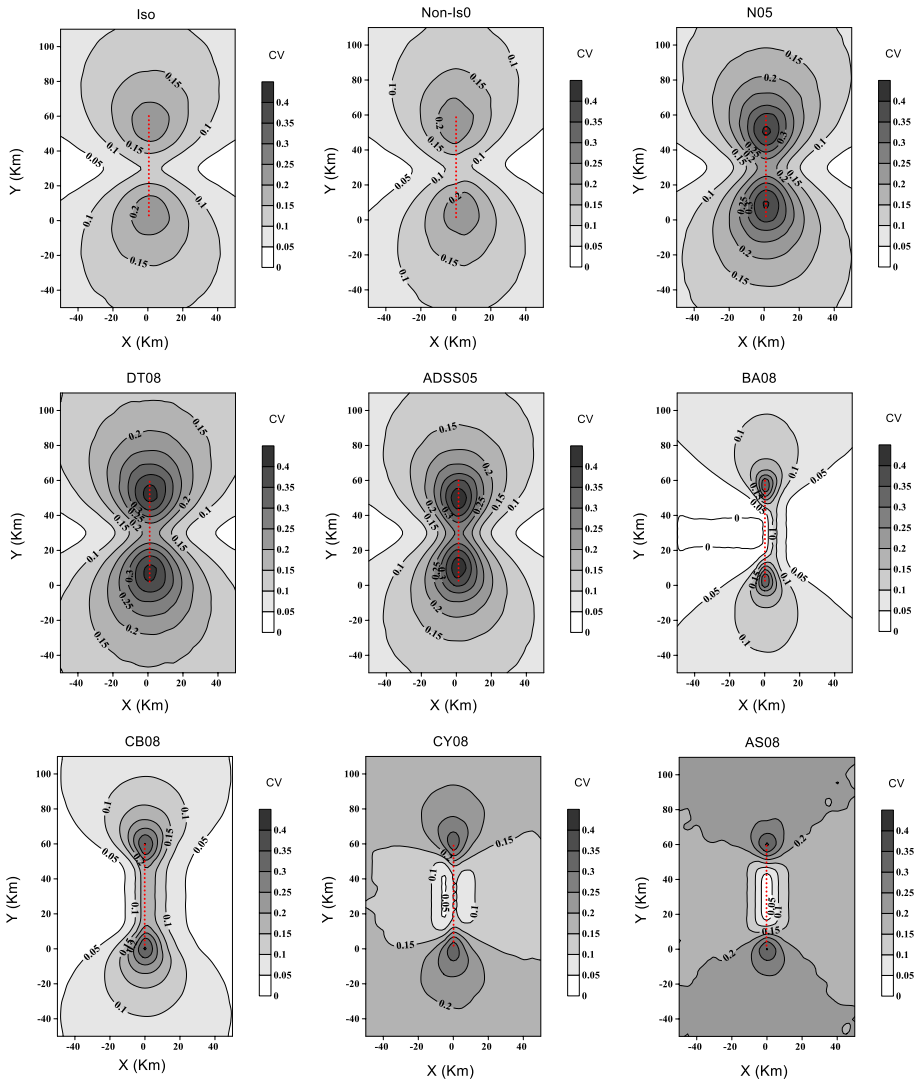
Fig. 10 The C_v contour of GMMs for the M_w6.93 event

6 Conclusions

In this paper, we investigated the uncertainty of source-to-site distance and its effects on the outputs of GMMs. A simple framework was also proposed based on MC simulation method. The results of the analysis demonstrated that the two factors of magnitude and distance to the fault affect the values of C_v. It was observed that in general, the value of this coefficient increases with the increase in magnitude and decreases with the increase in distance to the fault. Significant differences were observed in the standard deviations of different distance definitions in different directions. The highest standard deviation of all

distance definitions was along and near the fault line, whereas the least standard deviations belonged to the direction perpendicular to fault rupture.

In GMMs that are only a function of R_{epi}, the distribution of the outputs of GMMs closely resembled that of the R_{epi}. In the NGA08 GMMs whose inputs are more than one variable, the output distributions of the GMMs were not the same as inputs distribution. The BA08 model which is a function of R_{jb} had the lowest and the AS08 and CY08 GMMs had the highest C_v. The GMMs' C_v maps indicated that the variations (σ in Eq. 7) in the vicinity of the fault fluctuate between 20 and 40 percent of the average value of the GMM output, which can have a significant impact on the determination of risk level in the region. The C_v of the GMMs decreased with the increase in distance to the fault, where the amount of decrease depended on the rupture location with respect to the fault. On average, the variations were less than 10% at distances farther than 30 km.

We showed the utilization of empirical relationships to determine the dimension of rupture have uncertainty. Employing the simulation algorithm proposed in this paper, this uncertainty was quantified. It was also depicted that, depending on the GMMs, the uncertainty can alter the results up to 40%. This amount of discrepancy can theoretically reach zero with the increase in precise information on the dimensions of fault rupture. However, with the available information, the outputs of the GMMs may increase up to 40%, which can have a significant impact on the seismic hazard determination of the region. This paper explains that distance uncertainty is one of the parameters that can have a great impact on the final results of PSHA and in the design of the important structures such as power plants, bridge, schools, etc.; this uncertainty can have a significant impact on engineering decisions to ensure the safety of the structure and its economic viability.

References

Abrahamson NA (2006) Seismic hazard assessment: problems with current practice and future developments. In: First European conference on earthquake engineering and seismology, pp 3–8

Abrahamson NA, Shedlock KM (1997) Overview. Seismol Res Lett 68(1):9–23

Abrahamson N, Silva W (2008) Summary of the Abrahamson & Silva NGA ground-motion relations. Earthq Spectra 24(1):67–97

Ambraseys NN, Douglas J, Sarma SK, Smit PM (2005) Equations for the estimation of strong ground motions from shallow crustal earthquakes using data from Europe and the Middle East: horizontal peak ground acceleration and spectral acceleration. Bull Earthq Eng 3(1):1–53

Assatourians K, Atkinson GM (2013) EqHaz: an open-source probabilistic seismic-hazard code based on the Monte Carlo simulation approach. Seismol Res Lett 84(3):516–524

Beirlant J, Kijko A, Reynkens T et al (2018) Estimating the maximum possible earthquake magnitude using extreme value methodology: the Groningen case. Nat Hazards. https://doi.org/10.1007/s11069-017-3162-2

Beven K, Almeida S, Aspinall WP, Bates PD, Blazkova S, Borgomeo E, Simpson M (2018) Epistemic uncertainties and natural hazard risk assessment-Part 1: a review of different natural hazard areas. Nat Hazards Earth Syst Sci 18(10):2741–2768

Bommer JJ, Douglas J, Scherbaum F, Cotton F, Bungum H, Fäh D (2010) On the selection of ground-motion prediction equations for seismic hazard analysis. Seismol Res Lett 81(5):783–793

Boore DM, Atkinson GM (2008) Ground-motion prediction equations for the average horizontal component of PGA, PGV, and 5%-damped PSA at spectral periods between 0.01 s and 10.0 s. Earthq Spectra 24(1):99–138

Bradley BA (2009) Seismic hazard epistemic uncertainty in the San Francisco bay area and its role in performance-based assessment. Earthq Spectra 25(4):733–753

Campbell KW, Bozorgnia Y (2008) NGA ground motion model for the geometric mean horizontal component of PGA, PGV, PGD and 5% damped linear elastic response spectra for periods ranging from 0.01 to 10 s. Earthq Spectra 24(1):139–171

Chiou B-J, Youngs RR (2008) An NGA model for the average horizontal component of peak ground motion and response spectra. Earthq Spectra 24(1):173–215

Danciu L, Tselentis GA (2007) Engineering ground-motion parameters attenuation relationships for Greece. Bull Seismol Soc Am 97(1B):162–183

Der Kiureghian A, Ditlevsen O (2009) Aleatory or epistemic? Does it matter? Struct Saf 31(2):105–112

Douglas J (2017) Ground-motion prediction equations 1964–2017. Pacific Earthquake Engineering Research Center, Berkeley

Everitt B (1998) The Cambridge dictionary of statistics. Cambridge University Press, Cambridge

Foulser-Piggott R (2014) Quantifying the epistemic uncertainty in ground motion models and prediction. Soil Dyn Earthq Eng 65:256–268

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

Hagiwara Y (1974) Probability of earthquake occurrence as obtained from a Weibull distribution analysis of crustal strain. Tectonophysics 23(3):313–318

Hayes GP, Wald DJ (2009) Developing framework to constrain the geometry of the seismic rupture plane on subduction interfaces a priori–a probabilistic approach. Geophys J Int 176(3):951–964

Kaklamanos J, Baise LG, Boore DM (2011) Estimating unknown input parameters when implementing the NGA ground-motion prediction equations in engineering practice. Earthq Spectra 27(4):1219–1235

Kijko A, Smit A, Sellevoll MA (2016) Estimation of earthquake hazard parameters from incomplete data files. Bull Seismol Soc Amer 106(3):1210–1222

Kilb D, Hardebeck JL (2006) Fault parameter constraints using relocated earthquakes: a validation of first-motion focal-mechanism data. Bull Seismol Soc Am 96(3):1140–1158

Knopoff L (1971) A stochastic model for the occurrence of main-sequence earthquakes. Rev Geophys 9(1):175–188

Kramer SL (1996) Geotechnical earthquake engineering. Prentice Hall Inc, Upper Saddle River, p 653

Lam NT, Tsang HH, Lumantarna E, Wilson JL (2016) Minimum loading requirements for areas of low seismicity. Earthq Struct 11(4):539–561

Mai PM, Spudich P, Boatwright J (2005) Hypocenter locations in finite-source rupture models. Bull Seismol Soc Am 95(3):965–980

McGuire RK, Arabasz WJ (1990) An introduction to probabilistic seismic hazard analysis. In Ward SH (ed) Geotechnical and environmental geophysics, Society of Exploration Geophysicists, vol 1, pp 333–353

Moss RE (2009) Reduced uncertainty of ground motion prediction equations through Bayesian variance analysis. Civ Environ Eng 175:1–86

Murray J, Segall P (2002) Testing time-predictable earthquake recurrence by direct measurement of strain accumulation and release. Nature 419(6904):287

Musson RM (1999) Probabilistic seismic hazard maps for the North Balkan region. Ann Geophys 42(6):1109–1124

Nishenko SP, Buland R (1987) A generic recurrence interval distribution for earthquake forecasting. Bull Seismol Soc Am 77(4):1382–1399

Nowroozi AA (2005) Attenuation relations for peak horizontal and vertical accelerations of earthquake ground motion in Iran: a preliminary analysis. J Seismol Earthq Eng 7(2):109–128

Patwardhan AS, Kulkarni RB, Tocher D (1980) A semi-Markov model for characterizing recurrence of great earthquakes. Bull Seismol Soc Am 70(1):323–347

Pavel F, Vacareanu R (2017) Evaluation of the seismic hazard for 20 cities in Romania using Monte Carlo based simulations. Earthq Eng Eng Vib 16(3):513–523

Raschke M (2013) Statistical modeling of ground motion relations for seismic hazard analysis. J Seismol 17(4):1157–1182

Reid HF (1911) The elastic-rebound theory of earthquakes. Univ Calif Publ Bull Dept Geol 6(19):413–444

Scherbaum F, Schmedes J, Cotton F (2004) On the conversion of source-to-site distance measures for extended earthquake source models. Bull Seismol Soc Am 94(3):1053–1069

Smit A, Kijko A (2016) Probabilistic seismic hazard assessment from incomplete and uncertain data. In: EGU general assembly conference abstracts, vol 18, p 7417

Stafford PJ, Berrill JB, Pettinga JR (2009) New predictive equations for Arias intensity from crustal earthquakes in New Zealand. J Seismol 13(1):31–52

Stewart JP, Douglas J, Javanbarg M, Bozorgnia Y, Abrahamson NA, Boore DM et al (2015) Selection of ground motion prediction equations for the Global Earthquake Model. Earthq Spectra 31(1):19–45

Tsang H-H, Yaghmaei-Sabegh S, Anbazhagan P, Sheikh MN (2011) A checking method for probabilistic seismic-hazard assessment: case studies on three cities. Nat Hazards 58(1):67–84

Utsu T (1984) Estimation of parameters for recurrence models of earthquakes. Bull Earthq Res Inst 59:53–55

Wang Z (2011) Seismic hazard assessment: issues and alternatives. Pure appl Geophys 168(1–2):11–25

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

Yaghmaei-Sabegh S (2012) A new method for ranking and weighting of earthquake ground-motion prediction models. Soil Dyn Earthq Eng 39:78–87

Yazdani A, Shahpari A, Salimi MR (2012) The use of Monte–Carlo simulations in seismic hazard analysis in Tehran and surrounding areas. Int J Eng Trans C Aspects 25(2):159–166

Yazdani A, Mirzaei S, Dadkhah K (2017) Non-parametric seismic hazard analysis in the presence of incomplete data. J Seismol 21(1):181–192

Youngs RR (2006) Epistemic uncertainty model for use of PEER-NGA ground motion models in national hazard mapping. Powerpoint presentation

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.