



**ACCEPTED ON ANNALS OF GEOPHYSICS, 62,
2019; Doi: 10.4401/ag-7886**

Incorporation of Multivariate Statistical Distribution of Magnitude-Distance and Monte-Carlo Simulation in Probabilistic Seismic Hazard Analysis

Alireza Azarbakht^{1,2} and Mohammad Ali Ebrahimi²

¹Department of Civil and Environmental Engineering, University of Strathclyde,
Glasgow, UK.

²Department of Civil Engineering, Faculty of Engineering, Arak University, Iran, P.O. Box
38156-88359.

Incorporation of Multivariate Statistical Distribution of Magnitude-Distance and Monte-Carlo Simulation in Probabilistic Seismic Hazard Analysis

Alireza Azarbakht^{1,2,*,\\$,€} and Mohammad Ali Ebrahimi^{2,¥}

¹Department of Civil and Environmental Engineering, University of Strathclyde, Glasgow, UK.

²Department of Civil Engineering, Faculty of Engineering, Arak University, Iran, P.O. Box 38156-88359.

Abstract The classical seismic hazard analysis is based on two independent simplified assumptions including the statistical distribution of magnitude (usually Gutenberg-Richter 1958) and the distance distribution (equal probability in each point of a given source). However, the interaction between the two distributions is rarely discussed in past researches. Therefore, a joint M-R distribution has been implemented in this paper in order to shed light into these simplified assumptions. The Tehran metropolis is considered as the case study since it locates in a highly active seismic region. Three seismological datasets were used in this study, i.e. the observed dataset, the simulated dataset based on the Han and Choi 2008 methodology, and the simulated dataset based on the EqHaz software platform. Then, the classical seismic hazard analysis results are compared with the results obtained based on the joint M-R distribution. The results show that the classical seismic hazard analysis is always conservative when compared with the results based on the simulated data.

Keywords: Tehran metropolis, Simulation, Earthquake Catalogue, Han and Choi, EqHaz.

Introduction

Seismic Hazard Analysis (SHA) is an important and timely issue in the performance-based earthquake engineering (McGuire, 2004), i.e. especially in high seismic regions such as

* Correspondence to: Alireza Azarbakht, Department of Civil Engineering, Faculty of Engineering, Arak University, Iran, P.O. Box 38156-88359.

‡ E-mail: a-azarbakht@araku.ac.ir

€ Principal Advisor, the manuscript idea, writing paper and drawing conclusions

¥ M.S. Graduate Student, performing the whole analysis of the manuscript within his M.S. thesis

Iran plateau (Berberian, 2005). The main aim of SHA is to forecast ground shaking characteristics such as Peak Ground Acceleration (PGA), Peak Ground Velocity (PGV), spectral ordinates and so forth (Baker, 2008). Gutenberg-Richter is one of the most important empirical-observation relationships within the classical SHA (Richter, 1958; McGuire, 2004). Several modifications were proposed in the literature in order to enhance this relationship (Cornell, 1968; Rosenblueth and Esteva, 1966; Isacks and Oliver, 1964; Bazzurro and Cornell, 1999). Various Gutenberg-Richter distributions are depending on the treatment of the Maximum Magnitude (M_{\max}) (Kagan, 2002). The correlation between magnitude and distance is, obviously, not present in this relationship. The hypothesis of this research is that if whether the joint distribution of the magnitude and distance will result in different conclusions in the case of seismic hazard when compared to the classical approach? Therefore, a joint distribution of M-R has been employed in this paper in order to assess this hypothesis. The Tehran metropolis is considered as the case study in this research since this capital is located in the highly active seismic region and the SHA results in this area will be informative for engineering and research purposes. For this aim, two catalogues are gathered including (1) the observed events which are collected from the previous studies (Shahvar and Zare, 2013; Berberian, 1994), and (2) a simulated set of events based on the Han and Choi approach (Han and Choi, 2008). All the observed and simulated events are within the 220 km rectangular area of the capital centre. The foreshocks and aftershocks are eliminated based on the Gardner and Knopoff approach (Gardner and Knopoff, 1974). Finally, 273 observed data are available for further investigations. The whole area is divided into 11 seismic sources, and then, the seismic characteristics are calculated for each seismic region, i.e. M_{\max} , β (seismicity rate) and λ_{m_0} (the rate of earthquakes with magnitudes higher than m_0) (Wells and Coppersmith, 1994; Kijko and Sellevoll, 1992). The seismic characteristics are compared with the available findings in the literature (Tavakoli and Ghafory-Ashtiany, 1999). The classical PSHA has been performed based on Equation (1), and the results are compared with the available SHA results in the region (Gholipour, et al., 2008).

$$\lambda(IM > x) = \sum_{i=1}^{n_{sources}} \lambda(M_i > m_{\min}) \int_{m_{\min}}^{m_{\max}} \int_0^{r_{\max}} P(IM > x | m, r) f_m(M) f_r(R) dm dr \quad (1)$$

where IM is the intensity measure, $\lambda(IM > x)$ is the rate of $IM > x$, M is the moment magnitude, m_{\min} is the lower bound for the moment magnitude, $\lambda(M_i > m_{\min})$ is the rate of occurrence of earthquakes greater than m_{\min} from the source, $P(IM > x | m, r)$ comes from a given Ground Motion Prediction Equation (GMPE), $f(M)$ and $f(R)$ are, respectively, the probability density functions for magnitude and distance.

A set of events is also generated by employing the Han and Choi approach (Han and Choi, 2008) for different return periods, i.e. 10^4 , 10^5 and 10^6 years. This approach uses a random generation procedure based on the occurrence rate parameter which will be

described in details in the following sections. The other seismic characteristics, such as M_{\max} and λ_{m0} , are kept identical to the observed catalogue. Hence, the generated catalogue has a kind of inherent epistemic uncertainty which makes it different from the observed data. The PSHA approach is, then, performed again for the simulated cases and the results are compared with the classical hazard curve as shown in Figure 1. As seen in Figure 1, both of the hazard curves are close together except in the middle part which corresponds to the moderate return period values, i.e. between 10 years and 3500 years return periods which is a relatively wide range of interest in practice. As seen in Figure 1, the hazard curve based on the observed data is always higher than the hazard curve based on the simulated events which indicate that the hazard curve based on the observed catalogue is still on the safe side. It is worth mentioning that this difference is not meaningful when Equation (1) being employed, however, the difference rise up when a joint distribution between magnitude and distance is taken into consideration. This difference is the main focus of this paper and will be discussed with details in the next sections. It is worth mentioning that the author's findings show that the SHA results based on using $f_m(M).f_r(R)$ meaningfully differ from the results found on the implementation of $f_{m,r}(M, R)$ within Equation (1).

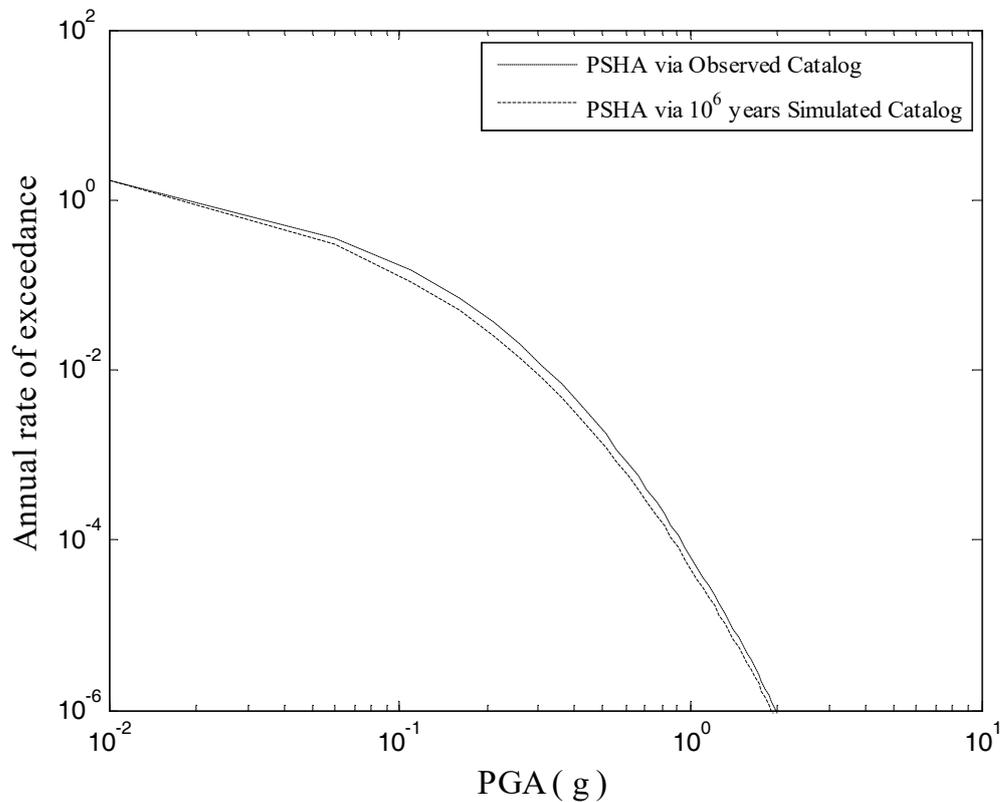


Figure 1. The hazard curves in Tehran metropolis by using the observed classical catalogue and Han and Choi simulated catalogue.

Input Data

Seismic Catalogue

The seismic catalogue was collected based on three data resources including: (1) the historical data (Berberian, 1994); (2) The data provided by (Shahvar and Zare, 2013) in which contains the moment magnitude; and (3) The BHRC data available at its website (see Data and resources).

Tehran metropolis is located in 35.69N and 51.42E (see Data and Resources section). Initially, the seismic data were gathered in the region 37N to 43.3N and 49.7E to 53.1E in which provides a rectangular area with the dimension equal to 150 km around Tehran metropolis. The possible active faults, as well as the distribution of the collected seismic data, are shown in Figure 2 (see also Data and Resources section). Furthermore, it was found that a set of active faults are located in the borders of the considered region. Therefore, the part was extended to 33.8N to 37.5N and 49.2E to 53.6E which provides a rectangular area with the dimension equal to 220 km (see Figure 2).

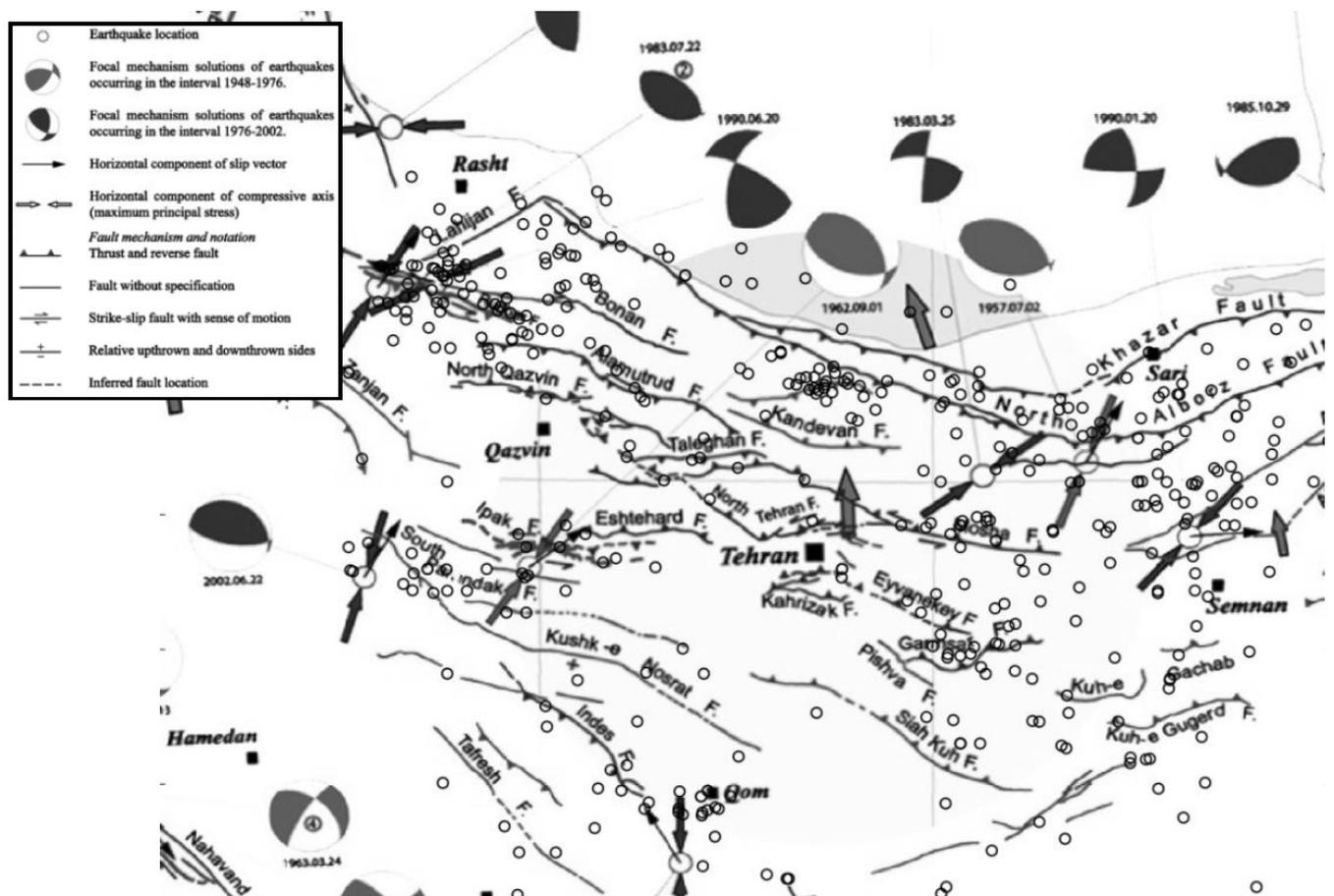


Figure 2. The active faults as well as the distribution of the collected seismic data in Tehran metropolis.

The active faults were identified consisting of four linear and seven areal seismic sources as seen in Figure 3. Different magnitude definitions in the seismic catalogue were transformed into the moment magnitude definition (M_W). To clarify, the Scordilid relationship (Scordilis, 2006) was employed in order to convert surface wave and body wave magnitudes into the moment magnitude. The local magnitude (M_L) was transformed into the moment magnitude by using the Shoja-Taheri relationship (Shoja-Taheri and Naserieh, 2007). Additionally, the distance between the event epicentres to the centre of the given region is calculated by using the Haversine relationship (Sinnott, 1984).

Seismic de-clustering

Three different methodologies were implemented to eliminate foreshocks and aftershocks in the gathered earthquake database. The methodologies consist of (1) Gardner and Knopoff method (Gardner and Knopoff, 1974); (2) Grounthal method (Van Stiphout et al., 2012); and (3) Uhrhammer method (Uhrhammer, 1986). The three mentioned methods use a window scheme in order to recognize foreshocks and aftershocks. Since the results based on the three pre-mentioned methodologies were quite close together, the authors decided to use the Gardner and Knopoff method (Gardner and Knopoff, 1974) for further investigations in this study. In the case of each seismic source, the Gardner and Knopoff method uses a linear regression between time and magnitude as well as a linear regression between magnitude and distance. The events which are below the regression lines are considered as foreshocks and aftershocks and were eliminated from the seismic database. However, the Gardner and Knopoff method is limited to the magnitudes up to 6.4 (Gardner

and Knopoff, 1974). Therefore, the Van Stiphout relationship (Van Stiphout et al., 2012) was implemented in the case of magnitudes higher than 6.4. Finally, after elimination of the foreshocks and aftershocks, 237 data were obtained as shown in Figure 4.

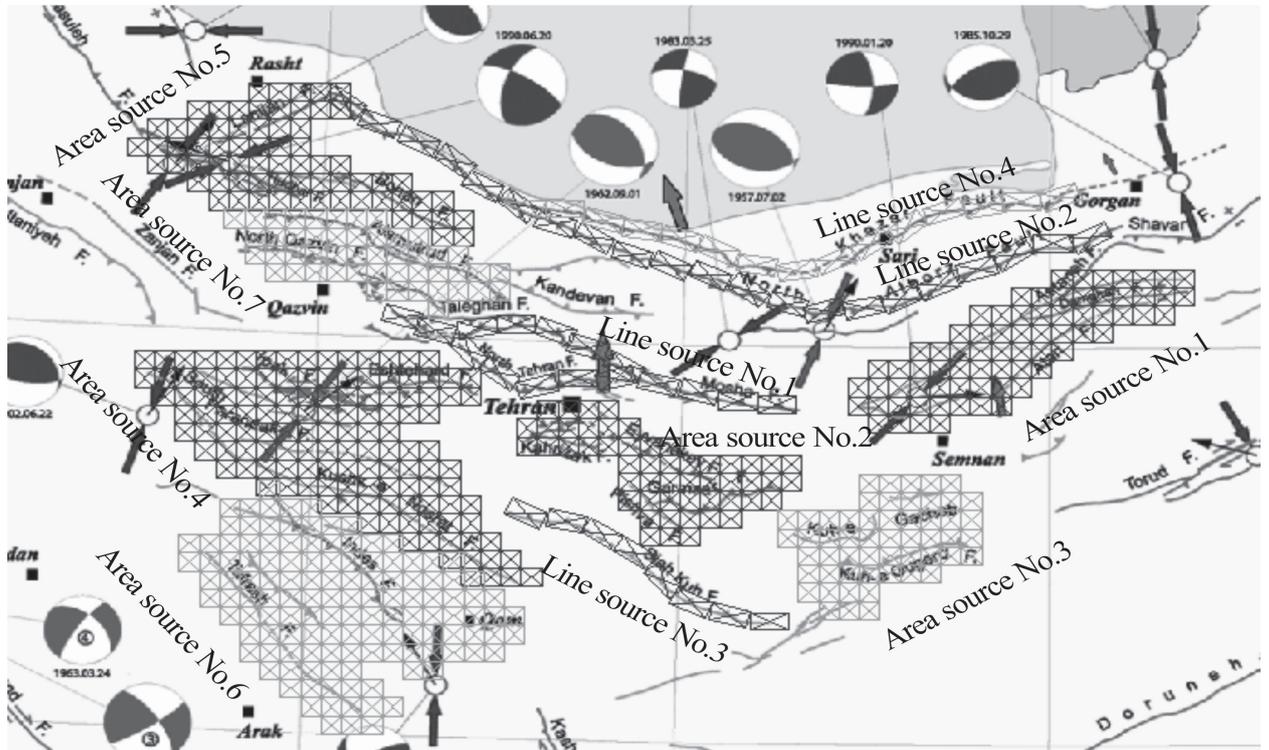


Figure 3. The 11 available seismic sources around Tehran metropolis.

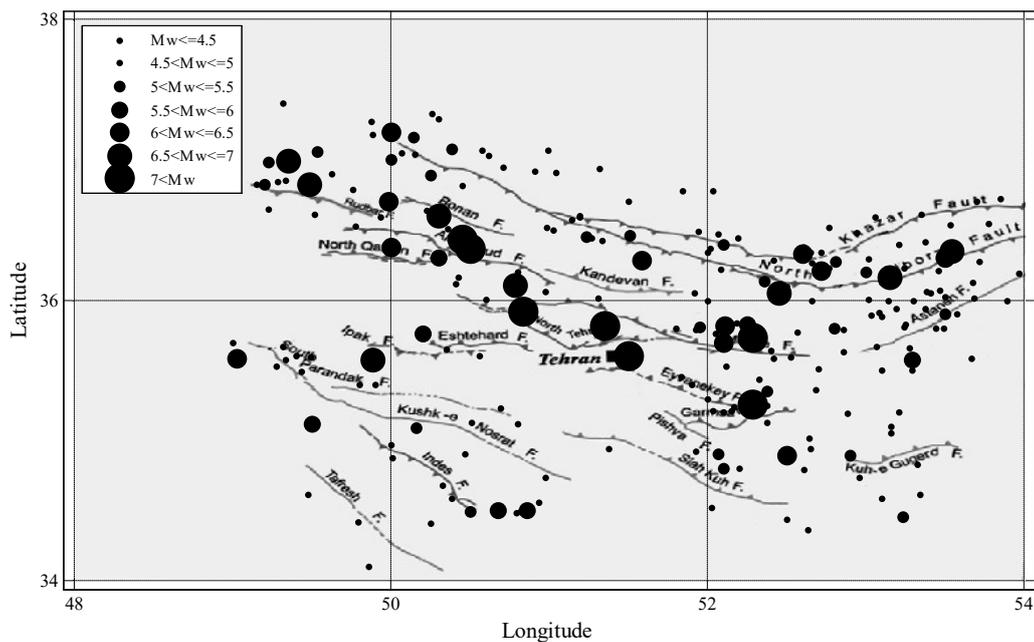


Figure 4. The distribution of the earthquake data in Tehran metropolis after foreshocks and aftershocks elimination.

Seismogenic zones and magnitude-frequency distribution

The seismicity characteristics, for each seismic source, were calculated by considering historical data. The Kijko method (Kijko and Sellevoll, 1989; Kijko and Sellevoll, 1992), which has been implemented in the Kijko2001 software platform (see Data and Resources Section), was used to calculate the β and λ_{m0} parameters. The maximum credible earthquake is calculated by using Equation (2) (Wells and Coppersmith, 1994). The Surface Rupture Length (SRL) uncertainty is equal to 20 percent of the fault length (Wells and Coppersmith, 1994) and at least 50 km (Biasi and Weldon, 2006). The seismic source mechanism is obtained based on the Iranian active faults map. The results are shown in Table (1). It is worth mentioning that the current results are compared, as seen in Table (1), to the Tavakoli and Ghafory Ashtiany (1999) results as a verification procedure.

$$M = a + b \times \log(SRL) \quad (2)$$

where M is moment magnitude; SRL is the surface rupture length; a and b are constants as shown in Table (2).

Table 1. The seismicity characteristics for each seismic zone in the study region.

No.	Seismic Source name	No. of earthquake events	Rupture length(Km)	Slip type	λ_{m0}	M_{max} (M_w)	β^*	Tavakoli and Ghafori-Ashtiany 1999		
								Zone	$M_{max}(M_w)$	β^*
1	Mosha	28	165	Strike-Slip	0.46	7.3	1.94	15	7.9	1.41
2	North-Alborz	54	428	Reverse	0.61	7.36	1.79	15	7.9	1.41
3	Astaneh/Kandovan	35	180/85	Strike-Slip/	0.55	7.07	2.04	15	7.9	1.41

				Reverse						
4	Siah-kuh	12	145	Reverse	0.39	7.07	1.92	8	7.4	1.99
5	Eyvanakey/ Garmsar/ Pishva	17	95/70/40	Reverse/ Reverse/ Normal	0.29	7.07	1.96	15	7.9	1.41
6	Gugerd-kuh	14	70	Reverse	0.26	7.07	2.2	8	7.4	1.99
7	Khazar	17	295	Reverse	0.32	7.16	2.11	20	7.5	2.32
8	Kushk-e- Nosrat/ Eshtehard/ Ipak/ Parandak- South	25	225/70/70/120	Normal/ Reverse/ Reverse/ Normal	0.38	7.07	1.48	8	7.4	1.99
9	Rudbar/ Bonan / Masuleh	19	75/70/95	Reverse/ Normal/ Reverse	0.37	7.1	1.96	15	7.9	1.41
10	Indes/ Tafresh	18	98/120	Reverse/ Normal	0.39	7.07	2.02	9	7.3	1.94
11	Alamurt-rud/ Qazvin-north	12	135/75	Reverse/ Reverse	0.7	7.07	2.2	15	7.9	1.41

λ_{m0} is the rate of earthquakes with magnitudes higher than m_0
* β is the seismicity rate obtained by the Gutenberg-Richter relationship

Table 2. a and b values employed in Equation (2).

Slip type	A	b
Strike-Slip	5.16	1.12
Reverse	5	1.22
Normal	4.86	1.32

Seismic Catalogue Simulation

Simulation of Catalogue by using the Han and Choi Algorithm

The Han and Choi algorithm (Han and Choi, 2008) is utilised in order to generate a simulated earthquake catalogue. The algorithm procedure is shown in Figure 5. The earthquake catalogue consists of 237 events in 62 years. Therefore, to keep the seismicity rate constant, 3822 events are needed for every 1000 years. This 62 year is a meaningful

restraint in the current study, but it is unavoidable. However, the Kijko method tries to consider this lack of data in the estimation of the seismicity rate. The whole region is divided into a refined rectangular mesh with an increment of 0.1 degrees in longitude and latitude. Each cell in the mesh is defined by two indices, i.e. (i,j) in which i is the row (latitude) number and j is the column (longitude) number. Then, the annual occurrence rate of Mmin is calculated for each cell and denoted by $\lambda_4(i,j)$. The distribution of λ_4 is shown in Figure 6 in which the Frankel method (Frankel et al., 1996) was used to smooth it. The weighting matrix is written in Equation (3) based on (Frankel et al., 1996). Then, Equation (4) is used to smooth the annual occurrence rate matrix by ten iterations (as done in Han and Choi, 2008).

$$W = \begin{bmatrix} 1/16 & 1/8 & 1/16 \\ 1/8 & 1/4 & 1/8 \\ 1/16 & 1/8 & 1/16 \end{bmatrix} \quad (3)$$

$$\lambda_4(i, j) = \sum_{l=i-1}^{i+1} \sum_{m=j-1}^{j+1} \lambda_4(i, j) \times W_{l-(i-2), m-(j-2)} \quad (4)$$

For simplicity, the annual occurrence rate matrix is transformed into a set of vectors by using Equations (5) and (6). Then, the cumulative distribution function (CDF) of this vector is calculated by using Equation (7) as shown in Figure 7.

To simulate the new location of the earthquake, a random number between 0 and 1 is generated as a random CDF number. Then, the corresponding F_λ is obtained based on Figure 7 (Equation 7). By using Equation (5), i and j numbers are derived which are corresponding to the row and column in the simulated annual occurrence rate matrix.

For the magnitude of the simulated catalogue, the same strategy as the location simulation is used. The only difference is that the CDF (Equation 7) of magnitude, as written in

Equation (8), is replaced instead of the location CDF. The simulation process is done for three sets of catalogues consisting of 10^4 , 10^5 and 10^6 seismic events.

$$k_{ij} = (i-1)N_c + j \quad (5)$$

$$\lambda_4(k_{ij}) = \lambda_4((i-1)N_c + j) = \lambda_4(i, j) \quad (6)$$

$$F_\lambda(k_{ij}) = \sum_{i=1}^{k_{ij}} P(k_{ij}) \quad (7)$$

$$F_M(m) = P[M < m \mid M > m_0] = \frac{\lambda_{m_0} - \lambda_m}{\lambda_{m_0}} \quad (8)$$

where $F_M(m)$ is the CDF of magnitude (m); m_0 is the magnitude lower bound; λ_m is the rate of earthquakes with magnitudes higher than m ; λ_{m_0} is the rate of earthquakes with magnitudes higher than m_0 .

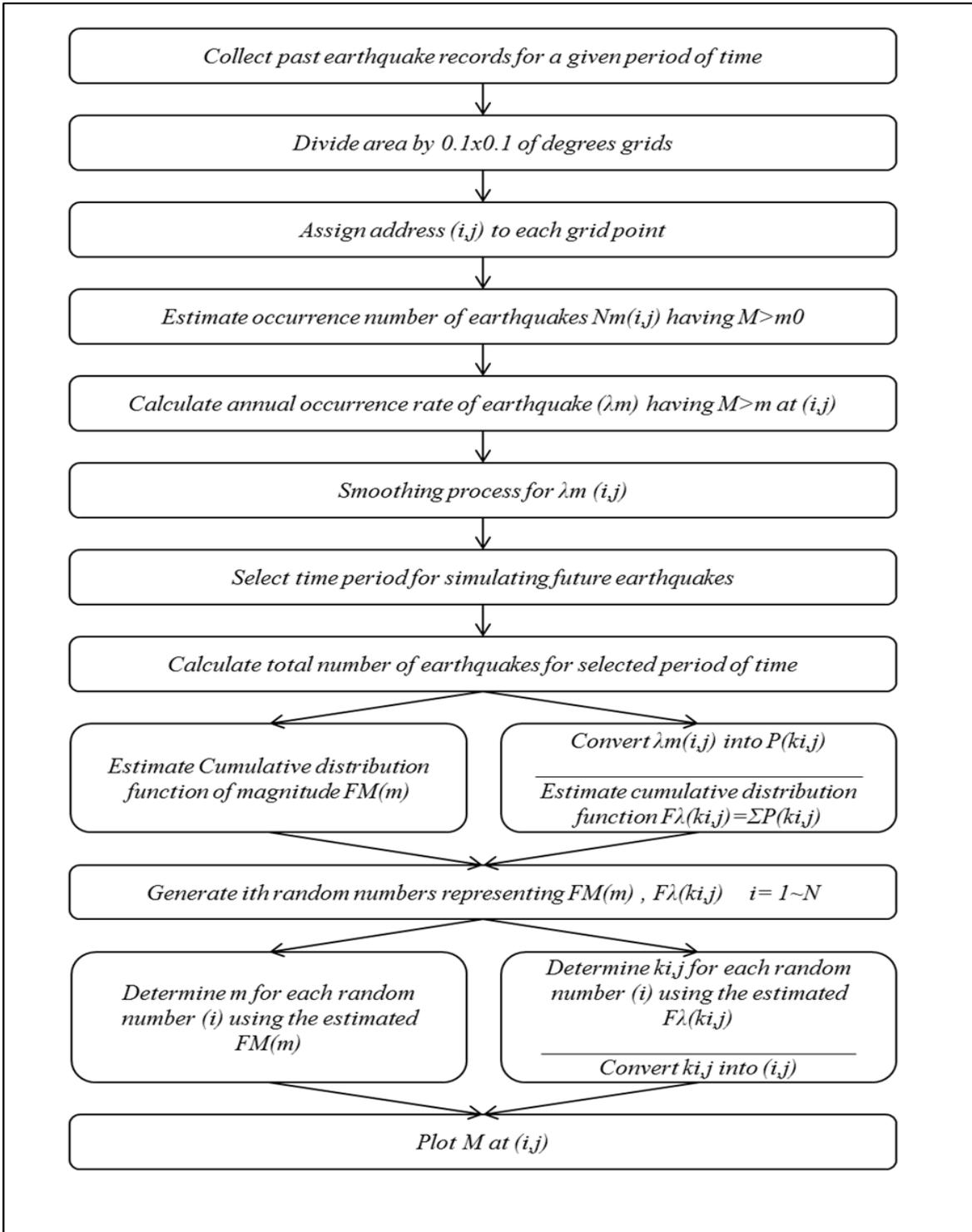


Figure 5. The Han and Choi algorithm (Han and Choi, 2008).

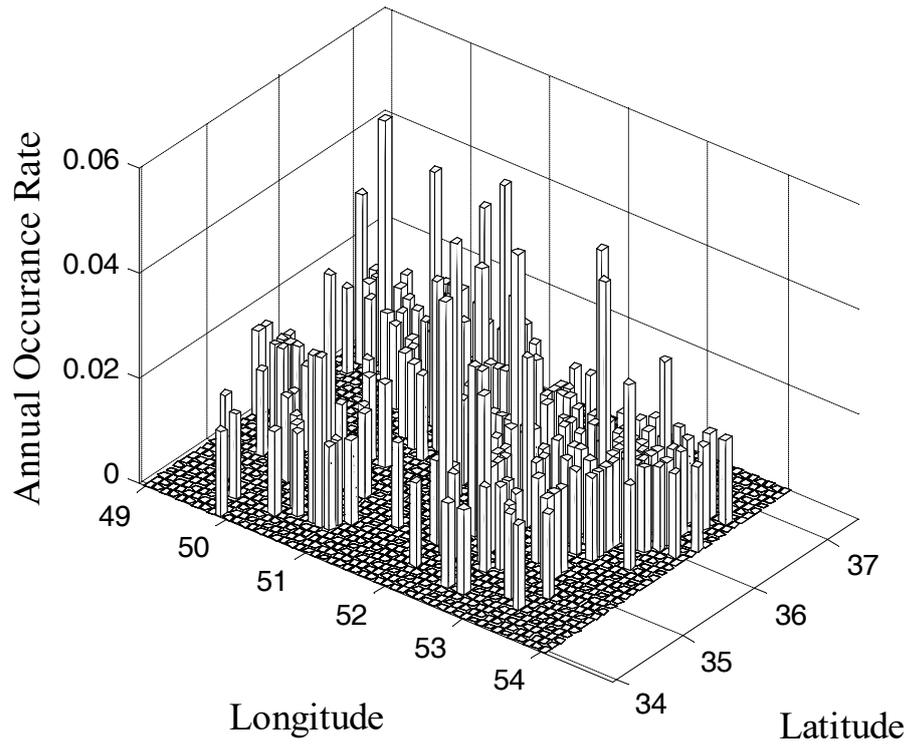


Figure 6. The annual occurrence rate based on the available catalogue in the considered region after smoothing.

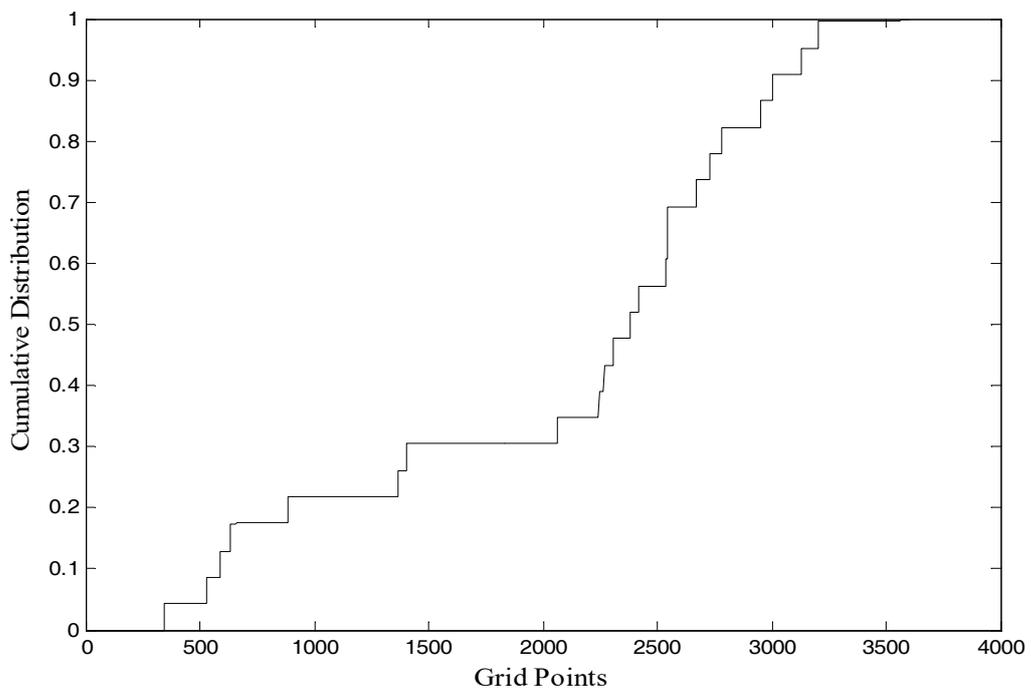


Figure 7. The cumulative distribution function of the yearly occurrence rate vector.

Simulation of Catalogue by using EqHaz software platform

The EqHaz platform (Assatourians and Atkinson, 2013) is an open-source software which was introduced for earthquake catalogue simulation by using Monte-Carlo (Ulam, 1961) method. The input data for the EqHaz consist of (1) seismicity rate of a given fault; (2) the rate of earthquakes with $M \geq m$; and (3) maximum magnitude. The algorithm within the EqHaz is based on the Monte-Carlo simulation method. In other words, the Monte-Carlo simulation method employs the distribution of magnitude, distance and time to re-generate new earthquake catalogues. The data based on Table (1) is used as input to the EqHaz platform, and three bins of catalogues generated including 62, 10^4 and 10^5 years.

Seismic Hazard Analysis

Classical Seismic Hazard Analysis

The classical Probabilistic Seismic Hazard Analysis (PSHA) is performed by using Equation (1) (McGuire, 1995). The magnitude Probability Density Function (PDF) is calculated by using Equation (9) (Cornell, 1968; Green and Hall, 1994). The distance PDF is based on the assumption that the probability of earthquake occurrence is equal in all mesh grids. The Boore and Atkinson GMPE (Boore and Atkinson, 2008) has been used in this study. The fault mechanism is available in Table (1), and the soil shear wave velocity is obtained based on (Zare et al., 1999) and BHRC data (see Data and Resources section). The linear faults are divided into 20 km segments, and the area sources are divided into 10km X 10km cell. Therefore, all the seismic sources are divided into 662 sections in which the distance between the cell centre and the city centre is calculated by using the Haversine relationship (Sinnott, 1984). The obtained hazard curve based on the classical PSHA is shown in Figure 8a.

$$f_M(m) = k\beta e^{-\beta(m-m_{mi})}, m_{\min} \leq m \leq m_{\max} \quad (9)$$

$$k = [1 - e^{-\beta(m_{\max}-m_{\min})}]^{-1}$$

where $f(M)$ is the magnitude PDF; β is the seismicity rate obtained by the Gutenberg-Richter relationship; m_{\min} and m_{\max} are, respectively, the lower and upper bounds for the magnitude; and k is a constant value.

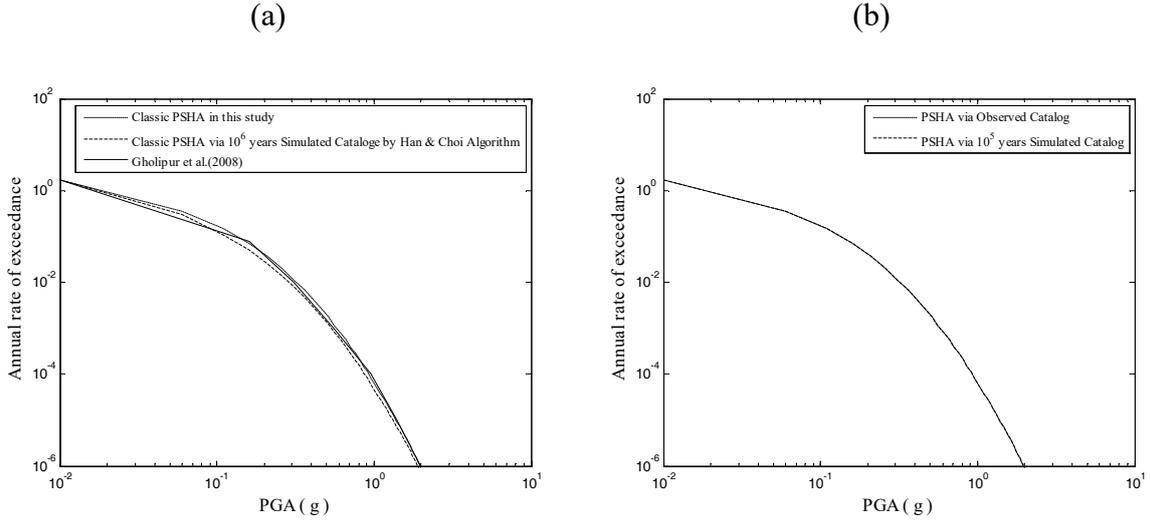


Figure 8. (a) The comparison of hazard curves for the site located in 35.75N and 51.45E based on this study and Gholipour et al.(2008)., (b) The comparison of hazard curves based on this study and EqHaz (Assatourians and G.M. Atkinson, 2013), the two curves are completely overlapped.

Seismic Hazard Analysis by employing the magnitude-distance joint distribution

To use the joint distribution of magnitude-distance, the considered region is divided into a rectangular mesh. Each square mesh cell has a dimension equal to 10 km. The magnitude range is between M_{\min} equal to 4 and M_{\max} with an increment of 0.5 where M_{\max} is different for each seismic source (see Table 1). For each magnitude value, a Geometrical Earthquake Occurrence (GEO) matrix is assigned in which each cell contains the number of events occurred in this location. The GEO matrix dimension is the same as the area mesh as schematically written in Equation (10).

are generated in the case of Equation (11). For example, the cell which contains 2268 number value in Equation (11) corresponds to the 0.005 of the earthquake catalogue.

The M-R joint distribution PDF is obtained based on this fact that the integral of the PDF should be equal to unity. Therefore, Equation (12) is used to calculate the Event Rate Coefficient (ERC). The joint PDF is then obtained by multiplying the GEO matrix by the ERC parameter as written in Equation (13). The PSHA integral is re-written in Equation (14) by implementing the joint distribution of magnitude-distance.

$$\begin{aligned}
F_{NO}(i, j, k) &= [Number\ of\ occurrence] \times \dots \\
\dots\ Coff.NO. & \left| \sum_{z=1}^k \sum_{y=1}^j \sum_{x=1}^i F_{NO}(i, j, k) = 1 \right. \\
\rightarrow Coff.NO. & \times \sum_{z=1}^k \sum_{y=1}^j \sum_{x=1}^i Number\ of\ occurrence(i, j, k) = 1 \\
\rightarrow Coff.NO. & = 1 / \sum_{z=1}^k \sum_{y=1}^j \sum_{x=1}^i Number\ of\ occurrence(i, j, k)
\end{aligned} \tag{12}$$

where Coff.NO. is the occurrence rate coefficient, i is the number of latitude mesh, j is the number of longitude mesh, k is the number of magnitude intervals, and *Number of occurrence* is the occurrence rate matrix.

$$F_{NO}(i, j, k) = [Number\ of\ occurrence] \times Coff.NO. \tag{13}$$

where F_{NO} is the joint M-R distribution matrix.

$$\lambda(IM > x) = \sum_{i=1}^{n_{sources}} \lambda(M_i > m_{min}) \int_{m_{min}}^{m_{max}} \int_0^{r_{max}} P(IM > x | m, r) f_{m,r}(M, R) dm dr \tag{14}$$

where all the parameters definition are identical to Equation (1) except $f_{m,r}(M, R)$ which is joint M-R PDF identical to F_{NO} in Equation (13).

Discussion of the results

Classical Seismic Hazard Analysis

The classical PSHA is obtained based on Equation (1) and the dataset in Table (2). The result is shown in Figure 8a. To have a point of comparison, the obtained hazard curve in the site located in 35.75N and 51.45E based on this study is compared with the result in (Gholipour et al., 2008) which was a national project. The assumptions and considered catalogue are not the same between the current study and (Gholipour et al., 2008); however, many common points are present between them. The two curves show good agreement as seen in Figure 8a that confirm the accuracy of the current PSHA results in this study.

Classical Seismic Hazard Analysis by using the simulated data

The classical PSHA (Equation (1)) is calculated in this section by implementing the simulated data, i.e. using the Han and Choi algorithm as well as the EqHaz platform. The standard hazard curves, based on the observed data (Table 1) as well as the simulated data, 10^6 years based on the Han and Choi method, are compared in Figure 8a in which all show good agreement together. However, the hazard curve based on the simulated data is always lower than the hazard curve based on the observed data in the range of 0.05g-0.5g regarding PGA. The difference comes from the value used in Equation (9). As the simulated catalogue has significantly more data than the observed catalogue, the regression between magnitude and rate of occurrences has been changed. In other words, the seismicity is less in the case of the simulated catalogue when compared to the observed catalogue. This may be an effect of generating several large magnitude events within the simulated catalogue.

The standard hazard curves, based on the observed data (Table 1) and the simulated data (10^5 years based on EqHaz platform (Assatourians and G.M. Atkinson, 2013)), are

compared in Figure 8b which are almost identical. The reason is that the EqHaz uses precisely the same seismic characteristics as the results obtained based on the observed catalogue. Therefore, it is evident that the two hazard curves are identical in Figure 8b.

Classical Seismic Hazard Analysis by using the joint distribution of magnitude-distance

In this section, the PSHA is calculated based on Equation (14) in which employs the M-R joint distribution. The hazard curves based on the Han and Choi algorithm (Han and Choi, 2008) in the case of 62, 10^4 , 10^5 and 10^6 years catalogues are shown in Figure 9a. As seen in Figure 9a, the 62 years of simulated catalogue is not reliable since the duration is too short. On the other hand, the 10^4 , 10^5 and 10^6 years catalogues result in nearly the same hazard curves. Therefore, the hazard curve based on the Han and Choi 10^6 years catalogue is used hereafter for further investigations in this study. The same data are provided in Figure 9b in the case of the EqHaz platform in which the hazard curve based on the 10^5 years EqHaz simulated catalogue will be used hereafter.

In order to distinguish the difference between Equation (1) and Equation (14), three different hazard curves are shown in Figure 10a consisting of (a) the classical PSHA identical to Figure 8a which employs Equation (1); (b) the hazard curve using Equation (14) and the 10^6 years simulated data by Han and Choi algorithm (Han and Choi, 2008); (c) the hazard curve using Equation (14) and based on the observed data. The same data are provided in Figure 10b in the case of using the EqHaz platform (Assatourians and G.M. Atkinson, 2013).

As seen in Figure 10a, the classical hazard curve is upper than the joint distribution hazard curves. The reason is that the magnitude, as well as the distance distributions, are based on the observed past events, however, in the case of the joint distribution function (Equation (14)); a standard deviation is taken into consideration in each cell. This standard deviation

makes it possible to consider some events in the locations that there is not any past event as well as some magnitudes that did not happen in the observed catalogue.

This justification in Figure 10b is different. As seen in Figure 10b, the classical PSHA still upper than the other two curves. However, the hazard curve based on the EqHaz platform show higher seismic hazards in high return periods when compared to the hazard curve based on the original observed catalogue. This phenomenon is vice versa in the low return period region as seen in Figure 10b. The reason is apparently coming from the difference between Equation (1) and Equation (14) which is the magnitude and distance distributions. For more elaborate with this issue, the magnitude probabilities are shown in Figure 11(a) to Figure 11(d) in four cases, i.e. the Gutenberg-Richter (1958), the observed catalogue, the Han and Choi catalogue, and the EqHaz catalogue. Figure 11(a) to 11(d) illustrate the magnitude distribution in the case of four seismic sources, i.e. line sources No. 1 and No.2 as well as area sources No.1 and No.3. The Han and Choi magnitude distribution is also nearly exponential in Figure 11(c). In general, the Han and Choi distribution is more compatible with the observed distribution when compared to the EqHaz distribution.

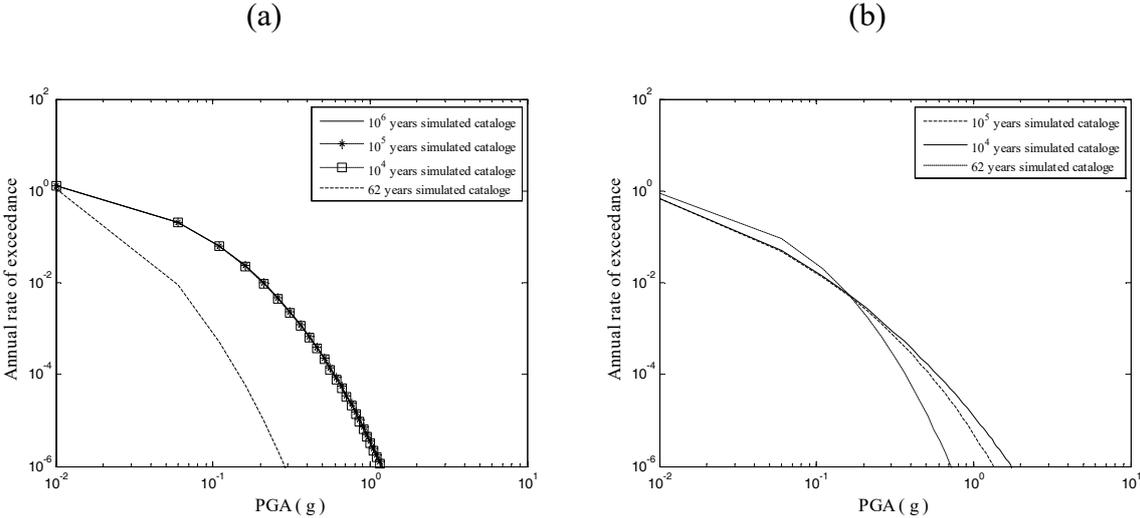


Figure 9. (a) The comparison of hazard curves in the case of Han and Choi algorithm (Han and Choi, 2008) for different catalogue durations, (b) The comparison of

hazard curves in the case of EqHaz platform (Assatourians and G.M. Atkinson, 2013) for different catalogue durations.

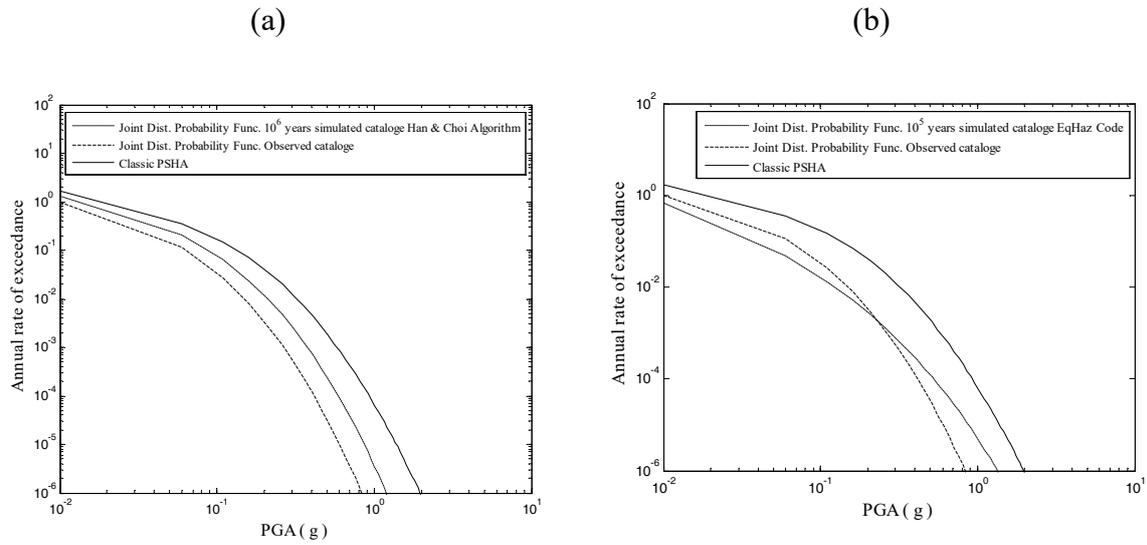
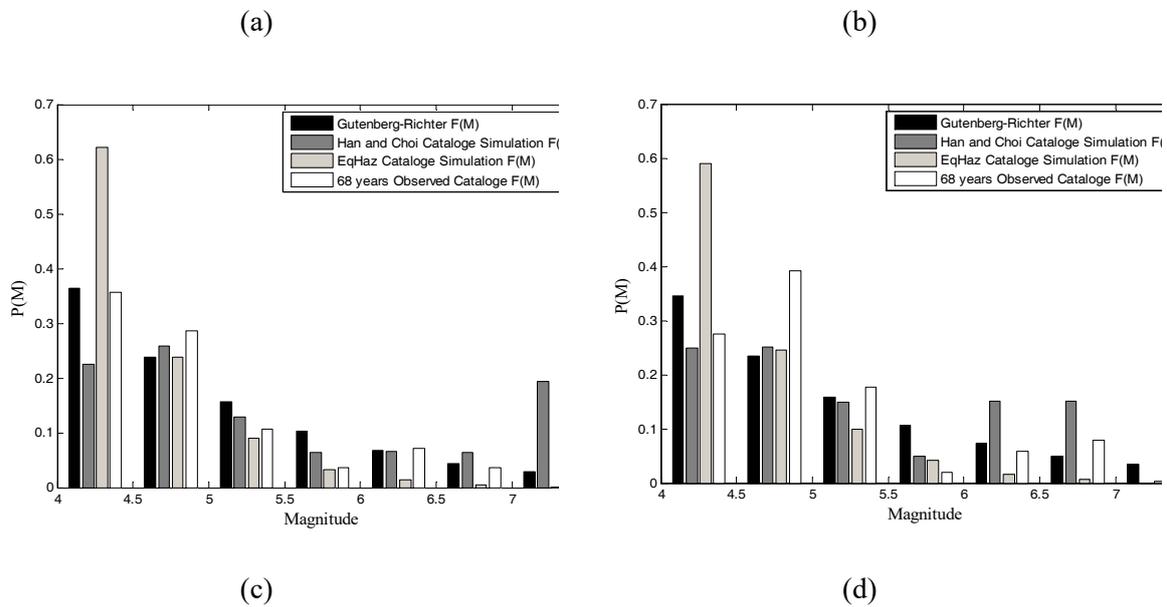


Figure 10. (a) The comparison of hazard curves in the case of using Equations (1) and (14) based on the Han and Choi algorithm (Han and Choi, 2008), (b) The comparison of hazard curves in the case of using Equations (1) and (14) based on the EqHaz platform (Assatourians and G.M. Atkinson, 2013).



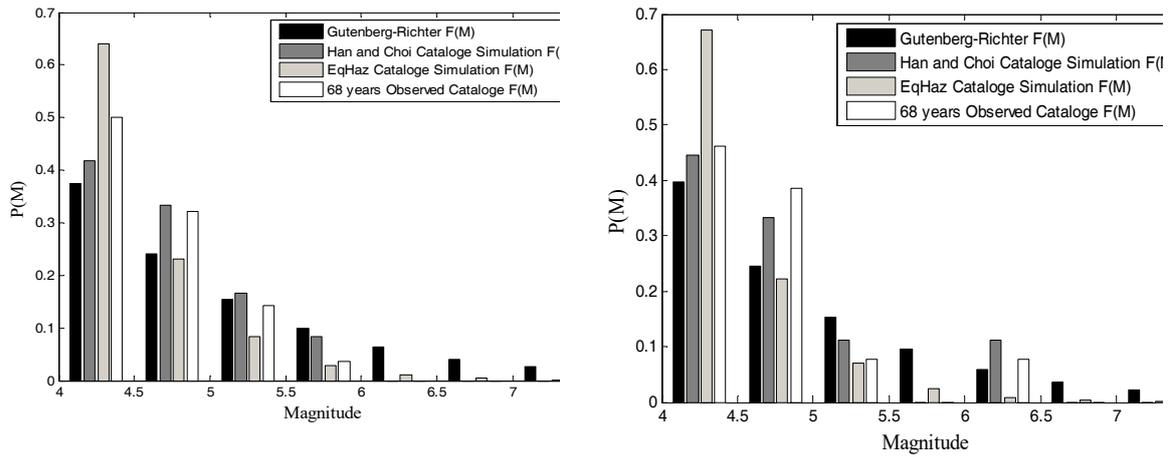


Figure 11. The comparison of magnitude probability in the cases of Gutenberg-Richter function, the Han and Choi method, EqHaz platform and 68 years observed data; (a) Line source No. 1, (b) Line source No. 2, (c) Area source No. 1, and (d) Area source No. 3.

Conclusion

In this study, the multivariate statistical distribution of magnitude-distance is altered into the classical hazard analysis, and the obtained results are discussed. The classical SHA always depends on the seismicity rate parameters. For example, the classical SHA for two regions with the same seismicity rate parameters, but different earthquake catalogues will result in the same hazard curves which are seriously challengeable.

The Han and Choi, as well as the EqHaz software platform, were taken into consideration to obtain simulated earthquake catalogues. The results show that the Han and Choi method is more compatible with the observed data distribution when compared with the EqHaz results. Also, the hazard curve based on the Han and Choi algorithm is always lower than the hazard curve based on the classical SHA. This phenomenon shows that the classical SHA is still on the safe side at least within the assumptions and the case study of this paper. The main reason for this phenomenon is that, in the classical SHA, only the occurrence probability is taken into consideration in the locations in which past recorded data are available in the catalogue. However, in the case of the Han and Choi algorithm, a

variation is considered around each location and magnitude. In other words, events are taken into consideration, within the Han and Choi algorithm, in the locations that there were no previous events in the observed catalogue.

It is worth mentioning that consideration of the M-R joint distribution results in meaningful different hazard curves when compared to the classical SHA hazard curves, at least within this study's limitations. More studies are obviously needed to prove this hypothesis.

Data and Resources

The ground-motion records were provided by the Building and Housing Research Centre (BHRC), Iran (<http://www.bhrc.ac.ir/portal/>, last accessed February 2014). Tehran location is 35°41'46"N 51°25'23"E based on the Wikipedia Internet encyclopaedia (Available at <http://www.wikipedia.org>, last accessed September 2013).

Iran faults map was provided by International Institute of earthquake engineering and seismology, Iran (<http://www.iiees.ac.ir/>, last accessed February 2014).

Kijko2001 software platform was provided by Council for Geoscience (CGS), South Africa (<http://www.geoscience.org.za>, last accessed September 2013).

References

- Assatourians, K., and G. M. Atkinson (2013). EqHaz: An open source probabilistic seismic hazard code based on the Monte Carlo simulation approach, *Seismological Research Letters*, **84**(3), 516-524.
- Baker, J. W. (2008). An introduction to probabilistic seismic hazard analysis (PSHA). *White paper, version, 1*, 72.
- Berberian, M. (2005). The 2003 Bam urban earthquake: a predictable seismotectonic pattern along the western margin of the rigid Lut block, southeast Iran, *Earthquake Spectra*, **21**(S1), 35-99.

- Berberian, M. (1994). Natural hazards and the first earthquake catalogue of Iran. International Institute of Earthquake Engineers and Seismology.
- Biasi, G. P., and R. J. Weldon (2006). Estimating surface rupture length and magnitude of paleoearthquakes from point measurements of rupture displacement. *Bulletin of the Seismological Society of America*, **96**(5), 1612-1623.
- Boore, D. M., and G. M. Atkinson (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, *Earthquake Spectra*, **24**(1), 99-138.
- Bazzurro, P., and C. A. Cornell (1999). Disaggregation of seismic hazard. *Bulletin of the Seismological Society of America*, **89**(2), 501-520.
- Cornell, C. A. (1968). Engineering seismic risk analysis, *Bulletin of the Seismological Society of America*, **58**(5), 1583-1606.
- Frankel, A. D., C. Mueller, T. Barnhard, D. Perkins, E. Leyendecker, N. Dickman, N., S. Hanson, and Hopper, M. (1996). *National seismic-hazard maps: documentation June 1996* (pp. 96-532). Reston, VA: US Geological Survey.
- Gardner, J. K., and L. Knopoff (1974). Is the sequence of earthquakes in southern California, with aftershocks removed, Poissonian. *Bulletin of the Seismological Society of America*, **64**(5), 1363-1367.
- Gholipour, Y., Y. Bozorgnia, M. Rahnema, M. Berberian, and J. Shojataheri (2008). Probabilistic seismic hazard analysis, phase I—greater Tehran regions. *Final report. Faculty of Engineering, University of Tehran, Tehran*.
- Green, R. A., and W. J. Hall (1994). *An overview of selected seismic hazard analysis methodologies*. University of Illinois Engineering Experiment Station. College of Engineering. University of Illinois at Urbana-Champaign.
- Han, S. W., and Y. S. Choi (2008). Seismic hazard analysis in low and moderate seismic region-Korean peninsula. *Structural Safety*, **30**(6), 543-558.

- Isacks, B., and J. Oliver, (1964). Seismic waves with frequencies from 1 to 100 cycles per second recorded in a deep mine in northern New Jersey. *Bulletin of the Seismological Society of America*, **54**(6A), 1941-1979.
- Kagan, V. Y. (2002). Modern California Earthquake Catalogs and their Comparison. *Seismological Research Letters*, **73**(6), 921.
- Kijko, A., and M. A. Sellevoll (1989). Estimation of earthquake hazard parameters from incomplete data files. Part I. Utilization of extreme and complete catalogs with different threshold magnitudes. *Bulletin of the Seismological Society of America*, **79**(3), 645-654.
- Kijko, A., and M. A. Sellevoll (1992). Estimation of earthquake hazard parameters from incomplete data files. Part II. Incorporation of magnitude heterogeneity. *Bulletin of the Seismological Society of America*, **82**(1), 120-134.
- McGuire, R. K. (1995). Probabilistic seismic hazard analysis and design earthquakes: closing the loop. *Bulletin of the Seismological Society of America*, **85**(5), 1275-1284.
- McGuire, R. K. (2004). *Seismic hazard and risk analysis*. Earthquake Engineering Research Institute.
- Richter, C. F. (1958). *Elementary seismology*.
- Rosenblueth, E., and L. Esteva (1966). On seismicity. In *Seminar in the application of statistics to structural mechanics*. Dept. of Civil Engineering, Univ. of Pennsylvania, Philadelphia.
- Scordilis, E. M. (2006). Empirical global relations converting M_S and m_b to moment magnitude, *Journal of seismology*, **10**(2), 225-236.
- Shahvar, M. P., M. Zare, and S. Castellaro (2013). A unified seismic catalog for the Iranian plateau (1900–2011), *Seismological Research Letters*, **84**(2), 233-249.
- Shoja-Taheri, J., S. Naserieh, and H. Ghofrani (2007). M_L and M_W scales in the Iranian Plateau based on the strong-motion records. *Bulletin of the Seismological Society of America*, **97**(2), 661-669.

- Sinnott, R. W. (1984). Sky and telescope, *Virtues of the Haversine*, **68**(2), 159.
- Tavakoli, B., and M. Ghafory-Ashtiany (1999). Seismic hazard assessment of Iran. *Annals of Geophysics*, **42**(6).
- Ulam, S. M. (1961). Monte Carlo calculations in problems of mathematical physics. *Modern Mathematics for the Engineers*, 261-281.
- Van Stiphout, T., J. Zhuang, and D. Marsan (2012). Seismicity declustering, *Community Online Resource for Statistical Seismicity Analysis*, 10.
- Wells, D. L., and K. J. Coppersmith (1994). New empirical relationships among magnitude, rupture length, rupture width, rupture area, and surface displacement, *Bulletin of the seismological Society of America*, **84**(4), 974-1002.
- Zare, M., P. Y. Bard, and M. Ghafory-Ashtiany (1999). Site characterizations for the Iranian strong motion network, *Soil Dynamics and Earthquake Engineering*, **18**(2), 101-123.

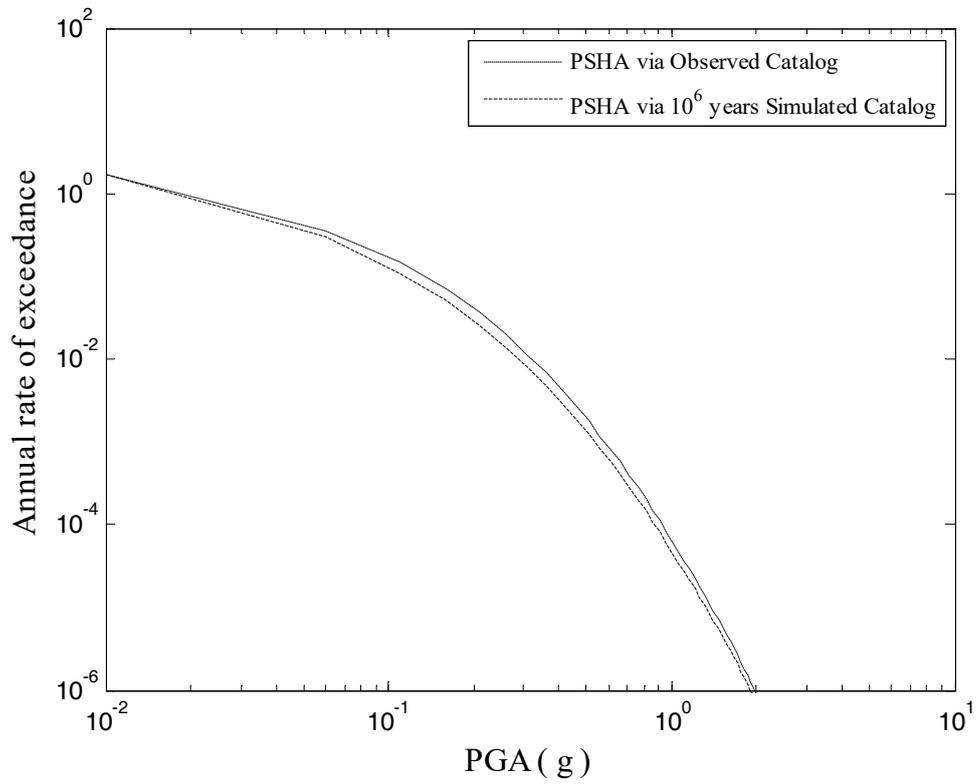


Figure 1. The hazard curves in Tehran metropolis by using the observed conventional catalogue and Han and Choi simulated catalogue.

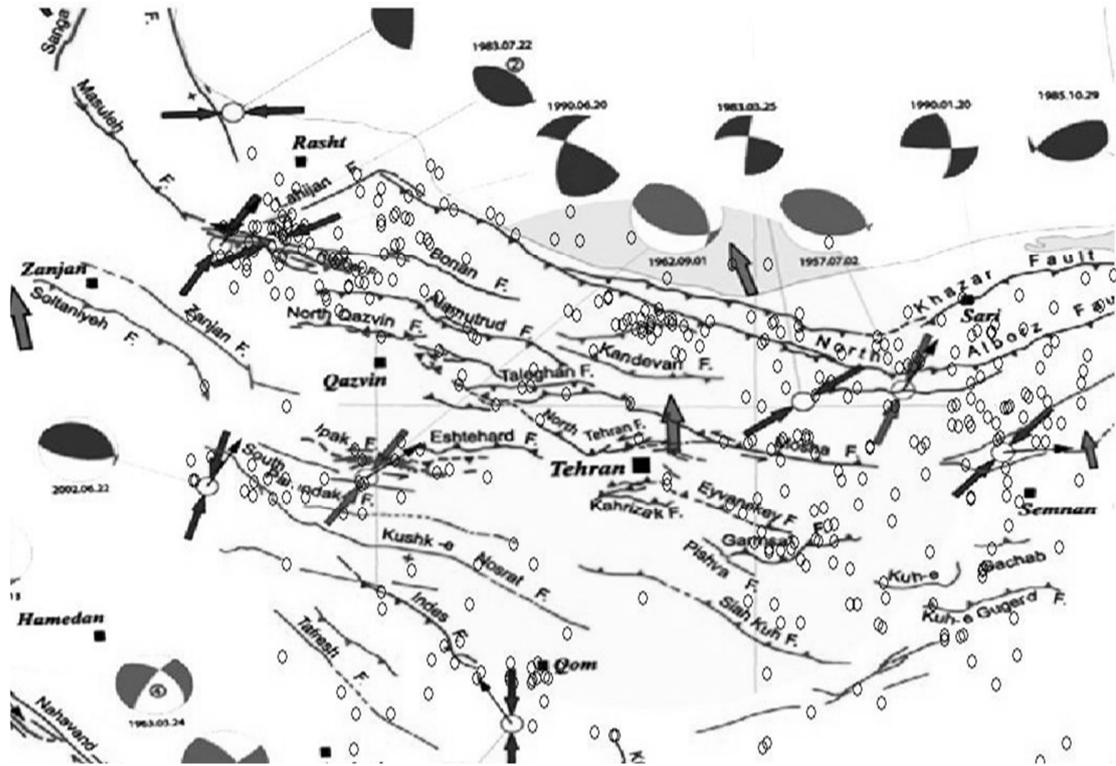


Figure 2. The available active faults as well as the distribution of the collected earthquake data in Tehran metropolis.

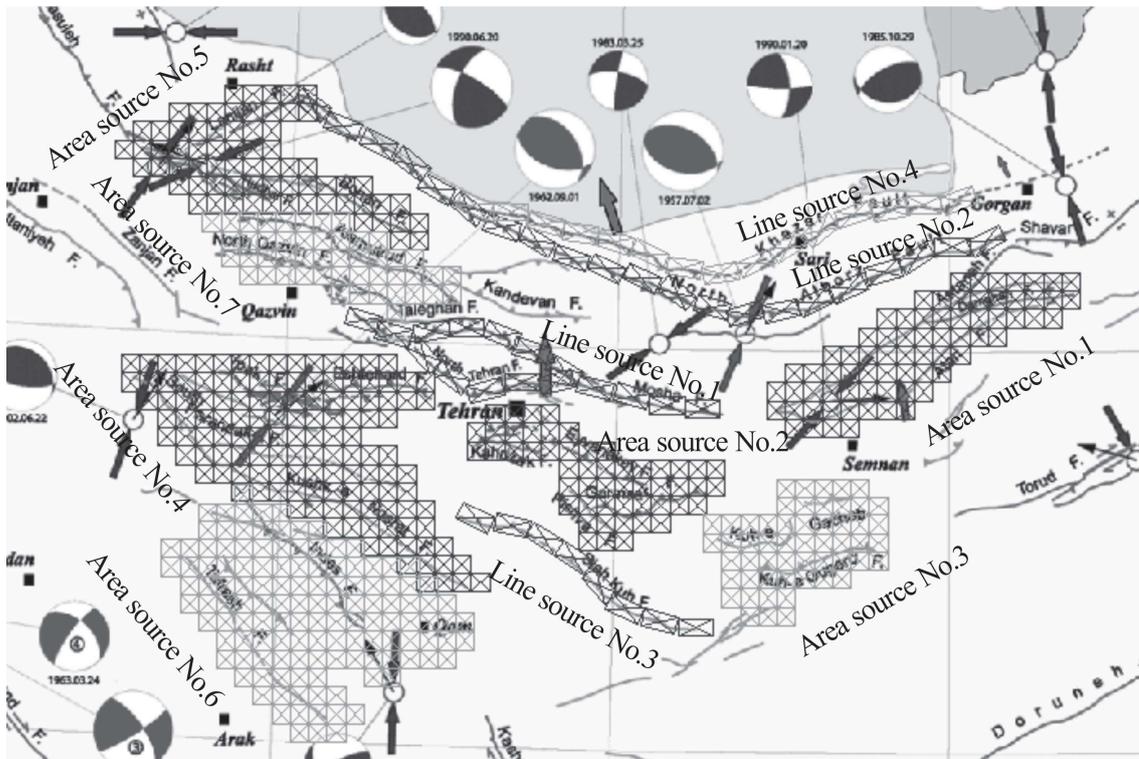


Figure 3. The 11 available seismic sources around Tehran metropolis.

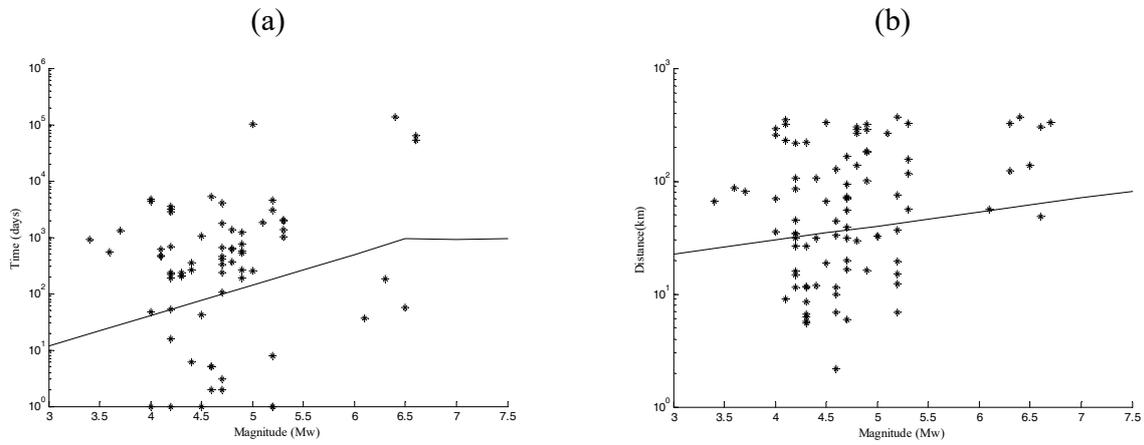


Figure 4. (a) Linear regression between magnitude and time in the case of seismic source No. 4 by using Equations (2) and (4)., (b) Linear regression between magnitude and distance in the case of seismic source No. 4 by using Equation (3).

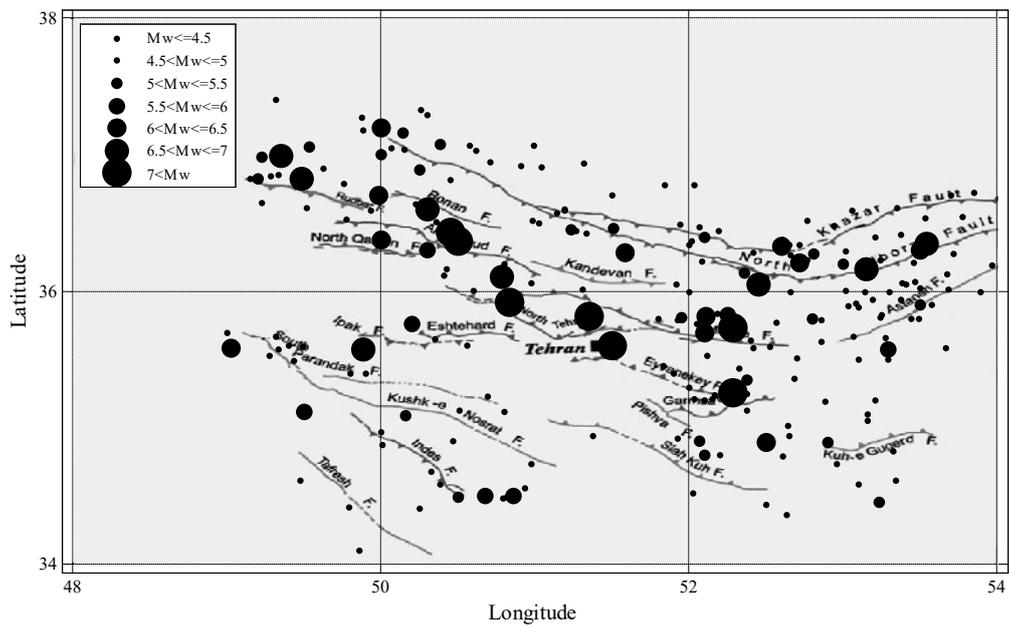


Figure 5. The distribution of the earthquake data in Tehran metropolis after pre and post-earthquake elimination.

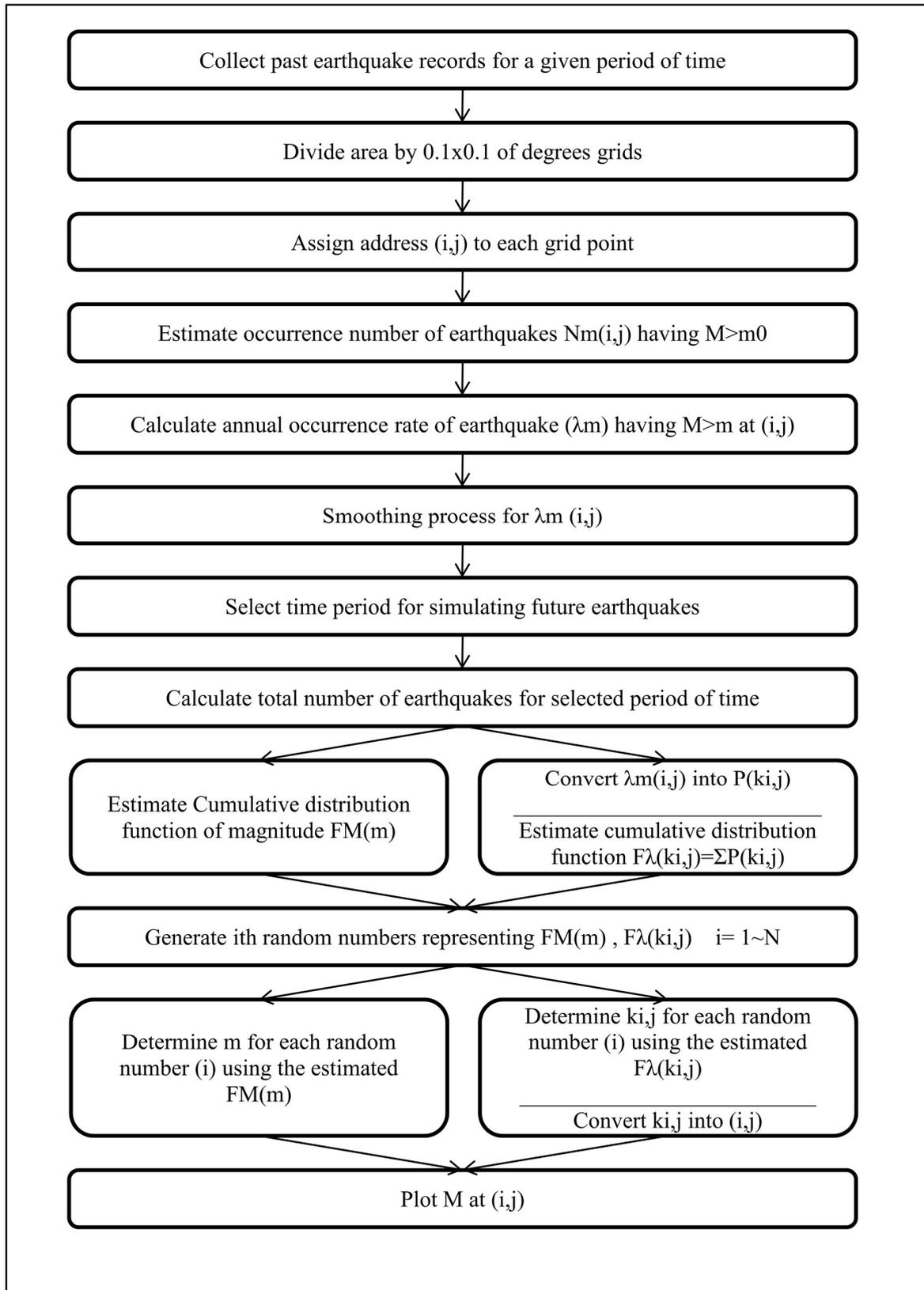


Figure 6. The Han and Choi algorithm (Han and Choi, 2008).

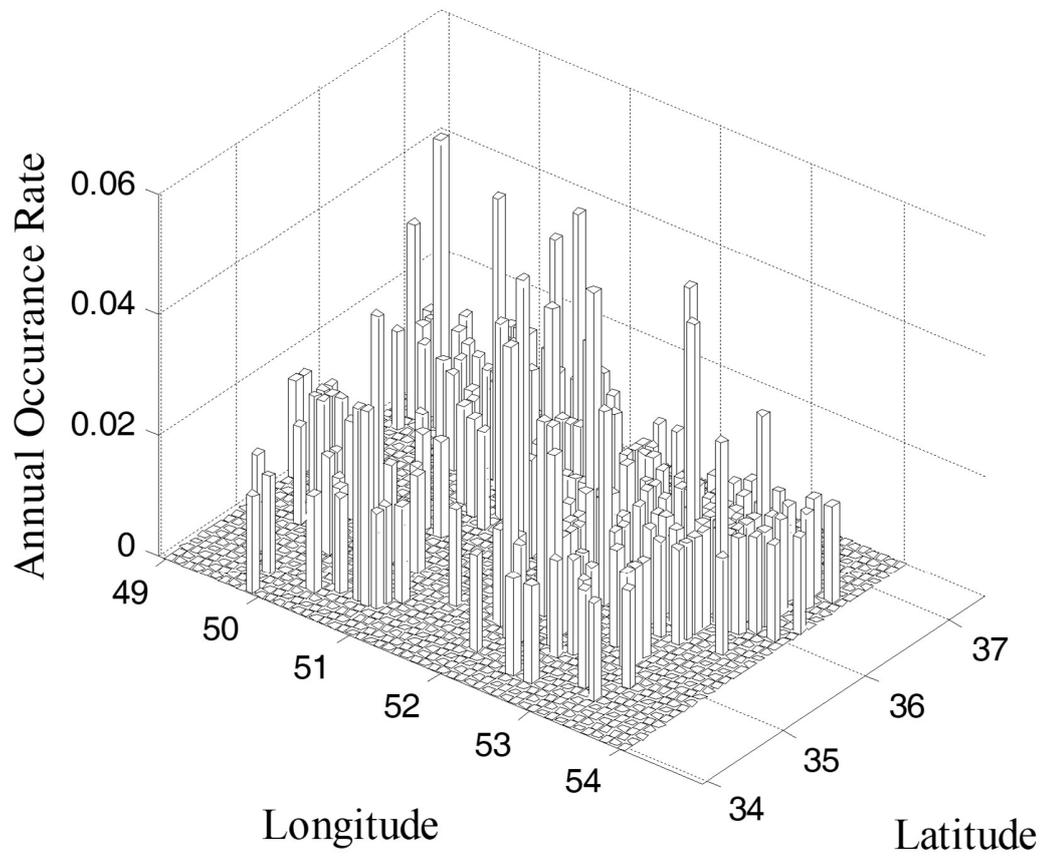


Figure 7. The annual occurrence rate based on the available catalogue in the considered region after smoothing.

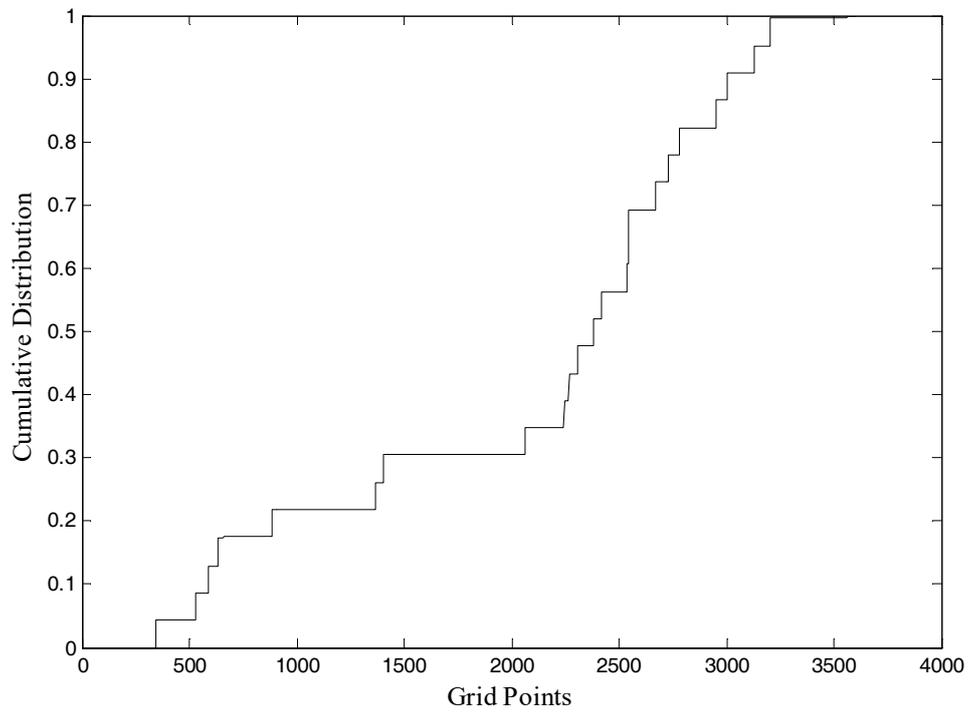


Figure 8. The cumulative distribution function of the annual occurrence rate vector.

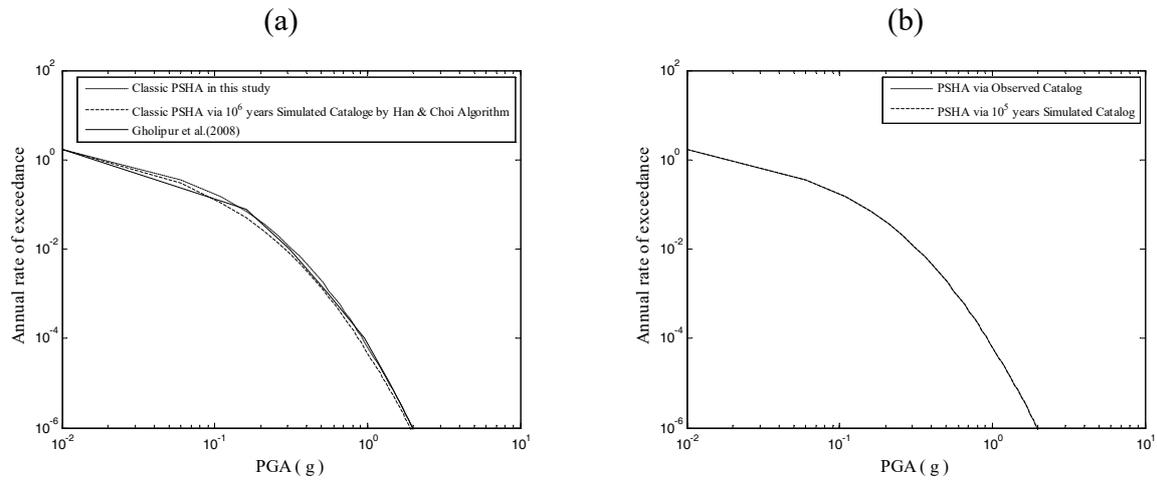


Figure 9. (a) The comparison of hazard curves for the site located in 35.75N and 51.45E based on this study and Gholipour et al.(2008)., (b) The comparison of hazard curves based on this study and EqHaz (Assatourians and G.M. Atkinson, 2013).

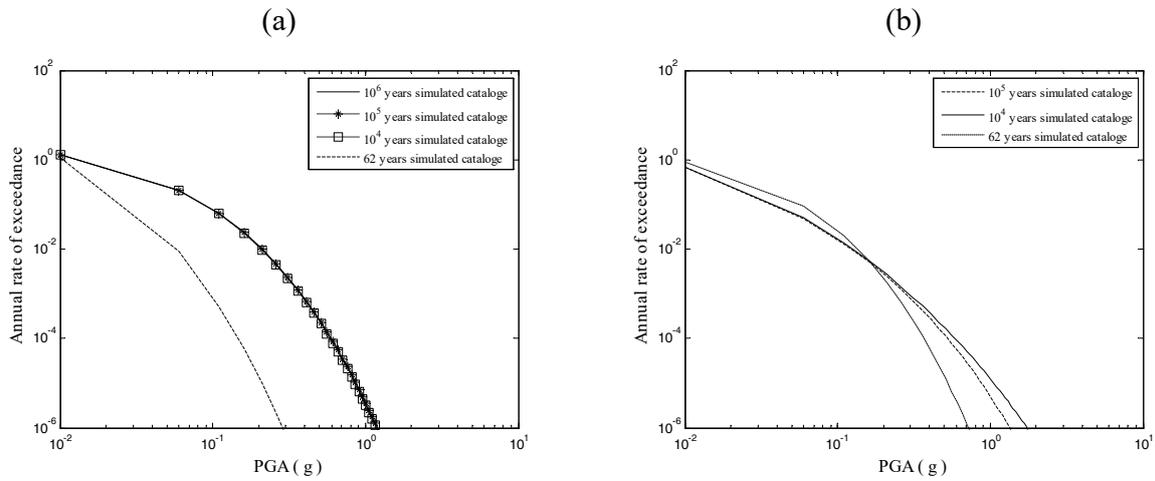


Figure 10. (a) The comparison of hazard curves in the case of Han and Choi algorithm (Han and Choi, 2008) for different catalogue durations, (b) The comparison of hazard curves in the case of EqHaz platform (Assatourians and G.M. Atkinson, 2013) for different catalogue durations.

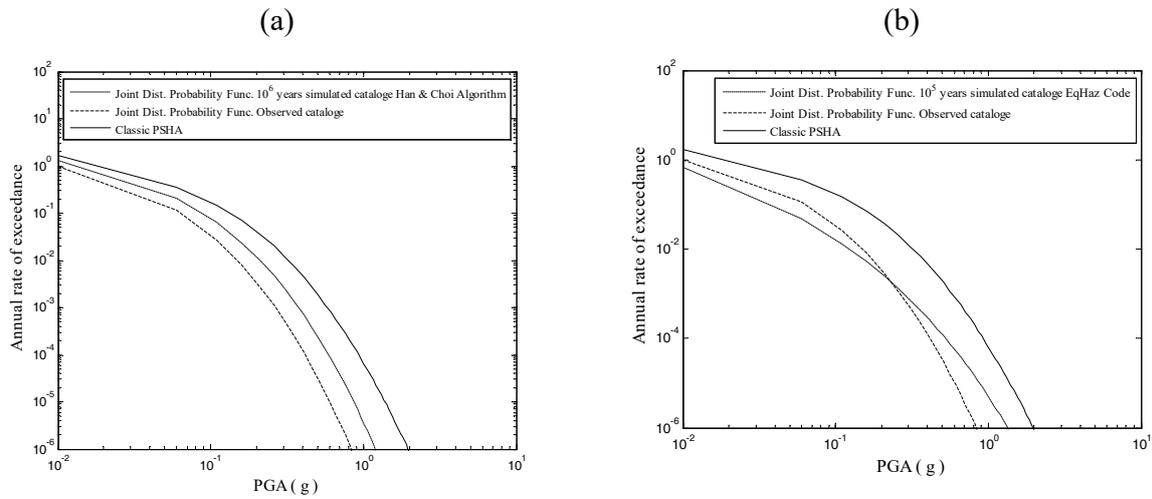


Figure 11. (a) The comparison of hazard curves in the case of using Equations (1) and (17) based on the Han and Choi algorithm (Han and Choi, 2008), (b) The comparison of hazard curves in the case of using Equations (1) and (17) based on the EqHaz platform (Assatourians and G.M. Atkinson, 2013).

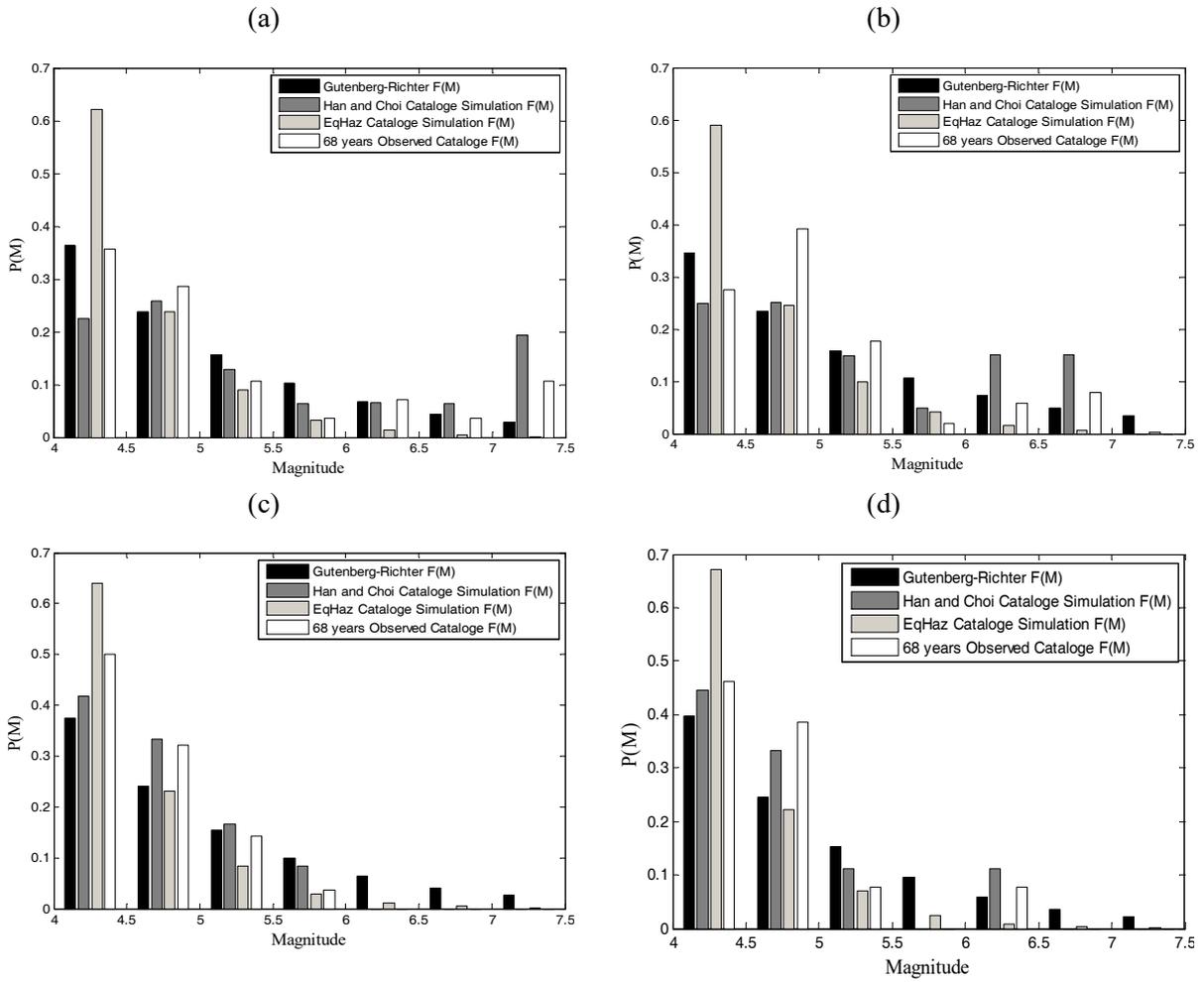


Figure 12. The comparison of magnitude probability in the cases of Gutenberg-Richter function, the Han and Choi method, EqHaz platform and 68 years observed data; (a) Line source No. 1, (b) Line source No. 2, (c) Area source No. 1, and (d) Area source No. 3.

Table 1. The seismicity characteristics for each fault in the region.

No.	Seismic Source name	No. of earthquake events	Rupture length(Km)	Slip type	M_{max} (M_w)	β	Tavakoli and Ghafori-Ashtiany 1999		
							Zone	$M_{max}(M_w)$	β
1	Mosha	28	165	Strike-Slip	7.3	1.94	15	7.9	1.41
2	North-Alborz	54	428	Reverse	7.36	1.79	15	7.9	1.41
3	Astaneh	35	180	Strike-Slip	7.07	2.04	15	7.9	1.41
4	Kandovan		85	Reverse					
5	Siah-kuh	12	145	Reverse	7.07	1.92	8	7.4	1.99
6	Eyvanakey		95	Reverse					
7	Garmsar	17	70	Reverse	7.07	1.96	15	7.9	1.41
8	Pishva		40	Normal					
9	Gugerd-kuh	14	70	Reverse	7.07	2.2	8	7.4	1.99
10	Khazar	17	295	Reverse	7.16	2.11	20	7.5	2.32
11	Kushk-e-Nosrat		225	Normal					
12	Eshtehard	25	70	Reverse	7.07	1.48	8	7.4	1.99
13	Ipak		70	Reverse					
14	Parandak-South		120	Normal					
15	Rudbar		75	Reverse					
16	Bonan	19	70	Normal	7.1	1.96	15	7.9	1.41
17	Masuleh		95	Reverse					
18	Indes	18	98	Reverse	7.07	2.02	9	7.3	1.94
19	Tafresh		120	Normal					
20	Alamurt-rud	12	135	Reverse	7.07	2.2	15	7.9	1.41
21	Qazvin-north		75	Reverse					

Table 2. a and b values for using in Equation (5).

Slip type	a	b
Strike-Slip	5.16	1.12
Reverse	5	1.22
Normal	4.86	1.32