Scientia Iranica A (2015) 22(2), 400–409

SCIENTIA IRANICA

Sharif University of Technology
Scientia Iranica
Transactions A: Civil Engineering
www.scientiairanica.com

Research Note

An artificial statistical method to estimate seismicity parameter from incomplete earthquake catalogs A case study in metropolitan Tehran, Iran

A. Yazdani[a,*], A. Nicknam[b], M. Khanzadi[b] and S. Motaghed[b]

a. *Department of Civil Engineering, University of Kurdistan, Sanandaj, P.O. Box 416, Iran.*
b. *School of Civil Engineering, Iran University of Science & Technology, P.O. Box 16765-163, Tehran, Iran.*

Received 7 December 2013; received in revised form 20 May 2014; accepted 26 July 2014

KEYWORDS
Completeness region;
Missing earthquakes;
Stochastic synthetic seismic catalog;
b-value;
Bootstrap sampling.

Abstract. Uncertainties in earthquake catalogs, earthquake recurrence parameters, and in the variation of ground motion parameters are often considered in the evaluation of seismic hazard analysis. The purpose of this study is to develop an artificial statistical procedure based on Bayes' formulation and weighted bootstrap sampling to estimate seismicity parameter (b-value of the Gutenberg-Richter law) from both historical and instrumental data in a given region. The procedure allows for uncertainty in the period of completeness, and assigns different weights to historical seismicity as compared to instrumental seismicity. Variation of seismicity within seismic sources is allowed with this procedure. This variation generalizes the condition of spatially homogeneous seismicity within seismic sources and permits an accurate representation of historical seismicity. As a case study, the earthquake catalog of the greater Tehran, Iran, is considered to estimate seismicity parameters as well as Probabilistic Seismic Hazard Analysis (PSHA) using the proposed procedure, and then the results are compared with those obtained from a conventional PSHA method. This comparison confirms the applicability of the procedure used in this study.

1. Introduction

Probabilistic Seismic Hazard Analysis (PSHA) is the most reasonable and applicable approach of evaluating the design values of earthquake ground motion [1-4]. The PSHA employs source model, occurrence model and attenuation relationships to quantify the probability of exceedance of the threshold intensity measure from all possible magnitude and distances [1,5]. Occurrence model is a representative of background seismicity of an area, but there is no distinct description of the seismicity properties, such as the distribution of earthquakes in space, time and magnitude domain [6]. So, modern earthquake catalogs, unlike older ones, tend to include the seismological information, such as origin times, hypocenter locations, earthquakes magnitudes, faults mechanism, and seismic moment tensors [7]. These catalogs are supposed to provide all the necessary information for further processing [8-11]. The main problem is how to combine this catalog information with the older data to best determine the seismicity parameters. It is quite postulated that statistical approaches are the best tools to deal with earthquake recurrence models due to their randomness [12]. The well-known Gutenberg-Richter (G-R) relation is commonly used for getting together the regional earthquakes information

*. *Corresponding author. Tel.: +98 871 6668457;*
Fax: +98 871 6660073
E-mail addresses: a.yazdani@uok.ac.ir and a.yazdani@qhttc.com (A. Yazdani)

and putting the data into a statistic version of the problem [1].

The results of Seismic Hazard Analysis (SHA) including G-R relation are highly influenced by input data which mathematically appear in the b-value, maximum magnitude, and occurrence rate [13]. The analyst should best interpret the inputs and consider the uncertainty attribute to these factors [4]. Apparently, any defect of the input data used in SHA (such as the accuracy, reliability and particularly the catalog incompleteness) may influence the results. Catalog incompleteness is in turn a function of the detection capability of the network which, in a medium-seismicity area, depends mainly on the density and distribution of seismic stations, as well as technical skill and engineering knowledge to recognize their recording characteristics. The former depends upon the economic power of the country for preparing the upgraded instruments. The above issues are the main causes of probable missing earthquakes data.

PSHA methods consider each earthquake as a point event in space and time. The basic assumption in these temporal point-process models is the time-stationarity of seismic occurrence. The estimation of recurrence parameters and incompleteness is a coupled problem. This paper is focused on seismic catalog incompleteness and associated problems in seismicity parameters estimation and proposes a procedure for catalog completion. For this purpose, the intention is to provide a reliable and general framework for the authentic earthquake recurrence parameter estimation. In this study, a procedure is presented for generating the missing earthquakes, data completion, and b-value estimation based on the principle of stationary stochastic processes. To illustrate the strength of the method, it is implemented in Tehran city as a case study and the proposed technique and the findings are compared with those of a traditional approach.

2. Earthquake catalog: Incompleteness and remedial issues

The incompleteness or shortage of data is a frequent and problematic issue in statistical analysis. The reliability of any statistical analysis is highly influenced by the quantity and quality of its input data [14]. Seismic catalog is an example of probable-incomplete data set.

Incomplete coverage due to limited sensitivity of recording network, variable recording accuracy in time and space, different ways of data interpretation, alteration of temporal-spatial distribution, and the mis-reported earthquake hypocenter are the major factors affecting the catalog incompleteness [9,15]. These factors are the results of political and social status, demographic conditions and construction circumstances for historical era and the coverage and accuracy of network for instrumental period. The earthquake magnitude distribution is usually determined by power law as:

$$\log\left[N(M)\right] = a - bM, \tag{1}$$

where N is the number of events with magnitudes not less than M, and a and b are constant coefficients. The PSHA depends highly on b-value which in turn depends on catalog incompleteness. As a result, many studies have been conducted on reliable b-value estimation and its time variability [16].

In general, regression methods are employed to estimate seismicity parameters. Reliability of estimation depends upon plentifulness and homogeneity of employed data. In regions with moderate seismicity, data augmentation is possible through enlarging the time interval, spatial interval, or both. The enlargement could lead to heterogeneity. For example, if all the historical data are used, the incompleteness of the earthquake catalog could commit bias in recurrence model parameters. In addition, the incompleteness degree depends on the time. Thus, temporal distribution of the earthquake is estimated incorrectly.

There are two different strategies to remedy the catalog incompleteness. The first strategy is based on estimation of b-value by employing a fairly complete and recent catalog data (e.g., recent 50 years) with data homogeneity assumption, and parameters estimation using Least Squares (LS) regression. Despite the widespread adoption of this strategy, using LS technique for b-value estimation does not have any statistical motivation [17]. Lamarre et al. [18] improved this method by introducing a max likelihood approach for G-R law estimation in ranges associated with uneven observations and different magnitudes. The time period of instrumental data is usually shorter than the recurrence rate of large earthquakes and large earthquake can influence the estimated parameters strongly, so these procedures are often associated with complications [9,18]. The other strategy is estimation of b-value based on entire catalog data, by utilizing statistical concepts for removing data heterogeneity. Dong et al. [19,20] introduced a Bayesian framework for developing recurrence relationship which is compatible with the geological, historical, and instrumental data. Kijko and Sellevoll elicited earthquake hazard parameters from incomplete data sets via maximum likelihood estimation [21]. Lamarre et al. [18] described a procedure based on the bootstrap statistical method to combine the uncertainties in earthquake catalog, recurrence and attenuation models.

In this study, we present an alternative method for catalog completion and b-value estimation based on Stochastic Synthetic Seismic Catalog (SSSC) generation under the assumption of stationarity. The key feature of this method is correction of detection probability through the use of Bayes' formulation, so it is possible to utilize the entire earthquake catalog in a region and estimate the uncertainties in b-value using the bootstrap sampling method. The proposed procedure describes seismicity distribution by the G-R relation under the assumption of stationarity. It also provides uncertainties resulting from the estimation method.

3. Methodology

In this section, a new approach is described for the completion of existing incomplete catalog and generating SSSC based on stationarity assumption of the process. This approach employs some statistical tools namely stochastic processes, missing data theory, Bayes' theorem and bootstrap sampling method.

The SSSC covers the entire recorded and unrecorded (but likely to occur) earthquakes in a region, so that the reported real data remains unchanged in the catalog (weighted by one) while those likely to happen (missing data) are weighted by their occurrence probability value. In other words, we append new earthquakes which probably happened but were not recorded (i.e., missing earthquakes). Therefore, the new catalog, or SSSC, will contain all the recorded and un-recorded earthquakes along with a coefficient indicating the recording characteristics of the events.

3.1. Catalog incompleteness

The catalog incompleteness is clarified by completeness region definition using detection probability concept. Completeness region refers to a certain geographical region, magnitude range, and time period where the detection probability is homogeneous [18]. Statistically speaking, in the given ranges, the mean of detection probability equals detection probability values, i.e. distribution is uniform.

In stationary processes, the future seismicity rate is estimated directly from the past information [10]. Magnitude range is the smallest range where stationarity is met which should be small enough to best provide homogeneity. It is not necessary for magnitude ranges to be equal to each other. In regions with moderate seismicity, where the earthquake of upper magnitude is rare, sometimes it is useful to enlarge ranges with magnitude. Geographical regions should be considered as the largest area where recording homogeneity can be assumed. Geographical region enlargement can provide the possibility to satisfy stationarity in smaller magnitude ranges.

The time period is determined based on the homogeneity of the recording in each range. Therefore, all the factors that affect the ability of recording, such as demographic changes, social and political events, the seismogram network, and development and completion play a role in determination of homogenous time periods. As a result, time intervals can be determined as the time between important changes of the aforementioned factors.

Boyd [22] incorporated the foreshocks and aftershocks into time-independent probabilistic seismic-hazard analyses by mathematically considering the cluster of all shocks as a union of events in which each event in the cluster has some probability of exceeding a given ground motion. But in this study, after depicting earthquakes from the catalog, and before any processing, aftershocks and foreshocks were removed. It should be noted that the catalog completeness is not concerned with aftershocks and foreshocks but depends on characteristics, accuracy and distribution of recording equipment and earthquake magnitude. For instance, where human witnesses were responsible for recording earthquakes, presence of human in the region (i.e. population distribution and density), event time, minimum sensible magnitude, condition and types of construction materials, which are effective on human's understanding of the magnitude and society development, are of great importance.

Shcherbakov et al. [23] indicated that most statistical parameters of aftershocks do not depend on the magnitude of the mainshock. But, in general, the reason for removing aftershocks and foreshocks is that the current methods of catalog completeness test are based on G-R law, which in turn, is based on the presumption of linear behavior of earthquakes in magnitude domain.

3.2. Detection probability

In PSHA, the earthquake is assumed as a stationary process and thus seismicity rate can be estimated from the past information [10]. A stationary process is defined as a stochastic process whose distribution is invariant over time and situations. The record ratio in stationary processed is defined as:

$$\mathrm{RR}_{\mathrm{CR}(t,g,m)}=\frac{n_{\mathrm{CR}}(t,g,m)}{n_{\mathrm{CCR}}(t_c,g,m)}, \tag{2}$$

where $n_{\mathrm{CR}}(t,g,m)$ is the earthquake occurrence rate in a given completeness region as a function of geographical region (g), magnitude range (m), and time period (t). $n_{\mathrm{CCR}}(t_c,g,m)$ is the earthquake occurrence rate in the cause of Complete Catalog Completeness Region (CCCR), i.e. for the same geographical region (g) and magnitude range (m), but the time period of complete catalog (t_c). The $n_{\mathrm{CR}}(t,g,m)$ is equal to $N_{\mathrm{CR}}(t,g,m)/t$, where $N_{\mathrm{CR}}(t_c,g,m)$ is the number of earthquakes in a given completeness region and $n_{\mathrm{CCR}}(t_c,g,m)$ are achieved form the same approach for CCCSs ($n_{\mathrm{CCR}}(t_c,g,m)=N_{\mathrm{CCR}}(t_c,g,m)/t_c$).

It should be noted that the number of earthquakes in a completeness region also depend on the uncertainties in magnitude, time and epicenter position. So the number of earthquakes in a completeness region has a probabilistic nature.

The recorded probability is defined as the chance of recording and/or reporting the occurred earthquake in a catalog and is determined from the record ratio using the negative binomial distribution [18]:

$$\mathrm{RP}_{\mathrm{CR}(t,g,m)} = (\mathrm{RR}_{\mathrm{CR}(t,g,m)})^{N_{\mathrm{CR}}(t,g,m)}. \tag{3}$$

Furthermore, annual recorded probability is calculated as:

$$\begin{aligned}\mathrm{ARP}_{t,g,m} &= (\mathrm{RP}_{\mathrm{CR}(t,g,m)})^{1/t} = (\mathrm{RR}_{\mathrm{CR}(t,g,m)})^{\frac{N_{\mathrm{CR}}(t,g,m)}{t}}\\ &= (\mathrm{RR}_{\mathrm{CR}(t,g,m)})^{n_{\mathrm{CR}}(t,g,m)}.\end{aligned} \tag{4}$$

The time interval of recorded probability in the time period between the two predefined event times (T) is calculated as:

$$\mathrm{TIRP}_{Tgm} = (\mathrm{ARP}_{t,g,m})^{T} = (\mathrm{RP}_{\mathrm{CR}(t,g,m)})^{T/t}. \tag{5}$$

Since the definition of the time interval of recorded probability is based on the occurrence ratio, it can be regarded as a conditional probability:

$$\mathrm{TIRP}_{Tgm} = p_{T,g,m}(A/B), \tag{6}$$

where, A and B stand for earthquake recording and occurrence probabilities, respectively. Finally, detection probability (or unrecorded probability) is defined as the probability of unrecorded events in completeness region between two successive main events and calculated as a complementary probability. The two time periods, t and t_c, should be sufficiently large so that the stationarity assumption of the earthquake occurrence can be satisfied.

3.3. Catalog completion

According to the above definitions and relations, missing earthquakes can be appended to the catalog based on earthquake scenario definition. A scenario is defined as a set of probable earthquakes between two apparently consecutive earthquakes in catalog after removing clusters, duplicates and magnitude scale conversion. These two consecutive earthquakes can be in the same completeness region in the case of incomplete catalog. Regarding the magnitude uncertainty, these two events also can be in different completeness regions. The upper bound of the magnitude range of the previous earthquakes should be greater than a certain threshold magnitude, m_t [18]. By the above definition, the SSSC can be generated.

3.4. Earthquake occurrence probability

According to the uneven occurrence probability of the earthquakes in the SSSC, two different conditions for assigning probabilities to earthquakes are considered. In the first condition, for the recorded earthquake, event weight is equal to conditional probability of occurrence in relationship to record that is equal to one. By considering the uncertainties of the reported events, the probability of event can be considered to be smaller than one. In this case, the record probability of the completeness region may also need to be modified. In another condition, for added earthquakes, event weight is equal to conditional probability of occurrence in relationship to detection probability that should be calculated. These calculated probabilities will be considered as bootstrap sampling weights.

The conditional probability of event occurrence given its detection, $P(B/A')$, and conditional probability of event non-occurrence given its detection, $P(B'/A')$, are calculated. B' and A' denote non-occurrence and non-reported probabilities, respectively. According to Bayes' theorem, posterior probability of B can be obtained as:

$$\begin{aligned}P(B/A') &= \frac{P(A'\cap B)}{P(A')}\\ &= \left(1-\frac{1}{P(A'/B)}+\frac{1}{P(B)P(A'/B)}\right)^{-1},\end{aligned} \tag{7}$$

where, $P(A'/B)$ and $P(B)$ are detection probability and occurrence probability, respectively.

The earthquake annual rate in the completed part of the catalog is divided by the annual rate of earthquake in the recent time period of that completeness region (e.g. completeness region with zero detection probability). It is implied that the annual rate of earthquake in the complete part of the catalog is known and is equal to the annual earthquake rate. Assuming the Poisson distribution, the occurrence probability between two consecutive earthquakes can be determined as:

$$P(B) = P(x>0) = 1-P(x=0) = 1-e^{-v}, \tag{8}$$

where v is the earthquake rate within the time period between the two respective earthquakes, which is equal to annual earthquake rate multiplied by the number of years between the two consecutive earthquakes.

Furthermore, Magnitude uncertainty should also be taken into account. For added earthquakes, magnitude uncertainty is accounted for using uniform distribution. For the recorded earthquakes this uncertainty can be applied using statistical distributions or fuzzy numbers. Distribution parameters bins are determined based on catalog information precision.

Augmenting the calculated weights to the completed catalog and allocating the weights to the events, the SSSC and quantify of heterogeneity of the initial catalog in incomplete periods can be generated. However the completed catalog is also non-homogeneous, so it cannot be used for seismicity parameters estimation directly, because the events have unequal weights and they cannot have equal roles in regression. For this reason, the bootstrap weighted sampling is used to compute the recurrence model parameters.

3.5. Weighted sampling of completed catalog

The bootstrap statistical method is used to deal with the incompleteness catalog and weighted sampling generation. In this method, sampling of the SSSC is done with the replacement procedure [24]. Bootstrap sample includes events that might be occurred, but have not actually occurred. Here, the initial set of n members is used to produce D bootstrap sets each with n members. D represents a big number like 10000 or more. The probability of the presence of each member in the bootstrap sample is proportional to each member's weight. Due to the stochastic nature of bootstrap sampling, there is no clear-cut rule for producing samples. The problem can be solved by increasing the number of samples.

The bootstrap sampling generates D sets of completed catalog for a region. The b-value for each of these sets can be obtained by Maximum likelihood regression analysis. So, a D-member set composed of the b-values and their mean and standard deviation can be obtained.

After the recurrence parameter is characterized, a probability density function is created that uses the distance from the rupture to the site to quantify the earthquake's location, and conditions for the source property are defined. Evaluation of seismic hazard requires an estimate of the expected ground motion and an attenuation relation at the site of interest. Analytically, the effects of all earthquakes of different sizes, occurring at different locations within different earthquake sources and having various frequencies of occurrence, are integrated into a single seismic-hazard curve that shows the frequencies of different levels of ground shaking being exceeded at a site during a specified period of time. The bootstrap sampling uses randomly generated points in the simulation of stochastic processes to cover the range of values that enter into calculation. The technique has the advantage of being relatively easy to be implemented on a computer and allowing uncertainty in the input parameters to be dealt with in a very powerful way by the generation of random numbers.

4. Case study: Metropolitan Tehran

Tehran, the capital, political and economic center of Iran, is located among the southern foothills of the Alborz mountain range, bounded by active faults [25]. Most active faults affecting the Central Alborz are parallel to the range and accommodate oblique convergence across the mountain belt. In the south of the Alborz range, the main active faults are the North Tehran, Mosha and Niavaran faults.

Previous conventional PSHA studies by Tavakoli and Ghafory-Ashtiany [26] and Ghodrati-Amiri et al. [27] and Monte-Carlo procedure PSHA by Yazdani and Abdi [28] calculated values for design-basis acceleration at greater levels than the value suggested by the seismic code. The seismic assessment at the site of interest depends mainly on the catalog of earthquakes and potential seismic sources that were compiled from available references containing historical and instrumental events in a radius of 200 km. Historical earthquakes in Iran (pre1900) were reviewed by Ambraseys and Melville [29] and Berberian [30]. Early (pre-1964) and recent (after 1964) instrumentally recorded events are collected from Shahvar et al. [31]. All other required parameters are taken from Yazdani and Abdi [28]. An overview of the earthquake catalog of Tehran shows its sparseness and heterogeneity. In other words, the historical data is very rare. In regions where the seismogenic progress is relatively unknown, the earthquake data is sparse. In this condition, instrumental data is often integrated with paleoseismological data to cover the longer duration, thus the data is inconsistent. Due to incomplete recording, the sequence of earthquake events displays a high degree of non-stationarity. For two different magnitude ranges, the events in Tehran region are shown in Figure 1. It is obvious that the part of the catalogue spanning from 1900 to 1960 and 1900 to 1990 is poorly reported for magnitude bins of 4.5-5.5 and 3.5-4.5, respectively, which are due to lack of observations. However, a good recording was observed for later years. Especially the number of earthquakes of magnitude 3.5-4.5 has enormous increase in recent years. It introduces a strong bias in the estimation of recurrence model parameters.

The data are controlled in different magnitude range to check the stationarity in time. The attained stationary magnitude ranges are 5-6, 6-7, and upper 7 ranges. The degree of incompleteness of the earthquake data needs to be assessed in order to use historical and large earthquake to refine the modeling of these large earthquakes in the future. The selected area, which is spatially homogeneous [27,32], includes residential and non-residential areas. The epicenter location of historical and instrumental recorded events in greater Tehran is presented in Figure 2. Figure 2(b) shows the

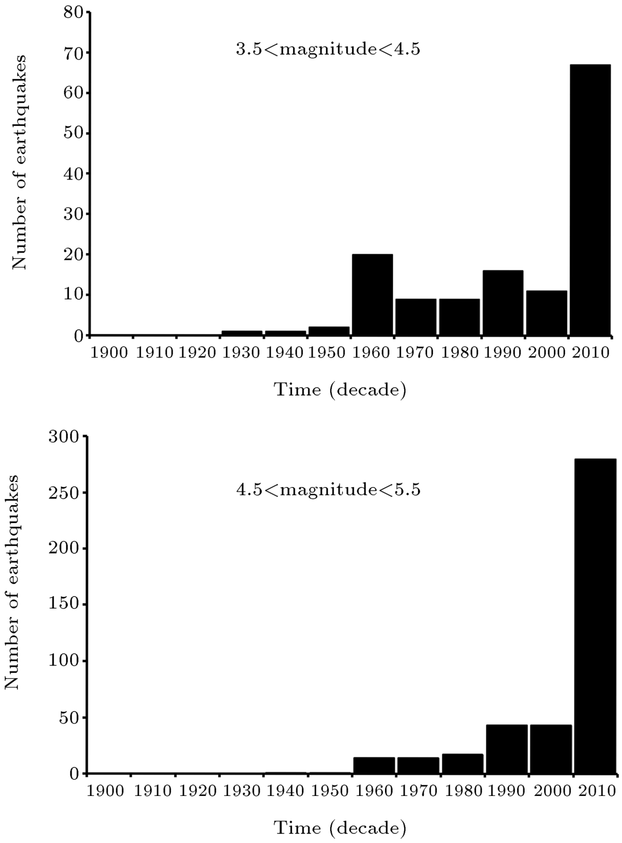

Figure 1. Histogram of the earthquake number in the time period 1900-2010 for Tehran.

approximate population distribution in Tehran area at three different years. The detection probability for each zone is calculated separately. There is no significant difference between detection probability in residential and non-residential areas, which can be due to the lack of accurate information about historical demography. Therefore, it seems illogical to discriminate between aforementioned areas.

Based on the recorded data in greater Tehran, Table 1 indicates the completeness time periods for different magnitudes [32]. Regarding the presumption of constant event rate in magnitude bins through time, the authors proposed the starting year of complete recording for each magnitude range, via the completeness analysis. The number of events in different

Table 1. Completeness time periods for Tehran [32].

Time period	Magnitude
0-855	-
855-1601	≥ 7
1601-1930	≥ 6
1930-1965	≥ 5
1965-1990	≥ 4.5
1990-2012	≥ 4

Table 2. Number of earthquakes for Tehran in completeness regions.

<table>
<tr><th rowspan="2">Time period</th><th colspan="3">Magnitude range</th></tr>
<tr><th>5.0-6.0</th><th>6.0-7.0</th><th>≥ 7</th></tr>
<tr><td>0-855</td><td>0</td><td>2</td><td>4</td></tr>
<tr><td>855-1601</td><td>1</td><td>4</td><td>4</td></tr>
<tr><td>1601-1930</td><td>9</td><td>4</td><td rowspan="4">4</td></tr>
<tr><td>1930-1965</td><td rowspan="3">25</td><td rowspan="3">2</td></tr>
<tr><td>1965-1990</td></tr>
<tr><td>1990-2012</td></tr>
</table>

Table 3. Earthquake rates for Tehran completeness regions.

<table>
<tr><th rowspan="2">Time period</th><th colspan="3">Magnitude range</th></tr>
<tr><th>5.0-6.0</th><th>6.0-7.0</th><th>≥ 7</th></tr>
<tr><td>0-855</td><td>0</td><td>0.0023</td><td>0.0046</td></tr>
<tr><td>855-1601</td><td>0.0013</td><td>0.0053</td><td>0.0053</td></tr>
<tr><td>1601-1930</td><td>0.0272</td><td>0.0121</td><td rowspan="4">0.0097</td></tr>
<tr><td>1930-1965</td><td rowspan="3">0.3048</td><td rowspan="3">0.0243</td></tr>
<tr><td>1965-1990</td></tr>
<tr><td>1990-2012</td></tr>
</table>

Table 4. Record ratio for Tehran completeness regions.

<table>
<tr><th rowspan="2">Time period</th><th colspan="3">Magnitude range</th></tr>
<tr><th>5.0-6.0</th><th>6.0-7.0</th><th>≥ 7</th></tr>
<tr><td>0-855</td><td>0</td><td>0.0959</td><td>0.4818</td></tr>
<tr><td>855-1601</td><td>0.0044</td><td>0.2201</td><td>0.5530</td></tr>
<tr><td>1601-1930</td><td>0.0894</td><td>0.4969</td><td rowspan="4">1.0</td></tr>
<tr><td>1930-1965</td><td rowspan="3">1.0</td><td rowspan="3">1.0</td></tr>
<tr><td>1965-1990</td></tr>
<tr><td>1990-2012</td></tr>
</table>

magnitude ranges based on completeness region in greater Tehran by assuming the temporal and spatial homogeneity are presented in Table 2. In this case study, magnitudes and epicenters error of the recorded events were not considered for simplification.

Regarding the earthquake number in different completeness region, the earthquake rate is calculated as mentioned in Table 3. Table 4 indicates the record ratio using Eq. (1) in completeness regions. This ratio is equal to one for complete completeness region. The small values show the unsuitable recording condition and incomplete reporting of the earthquakes. The record probabilities for each completeness region using Eq. (2) are shown in Table 5. These values are the record probabilities in the completeness region time period. The annual record ratios according to

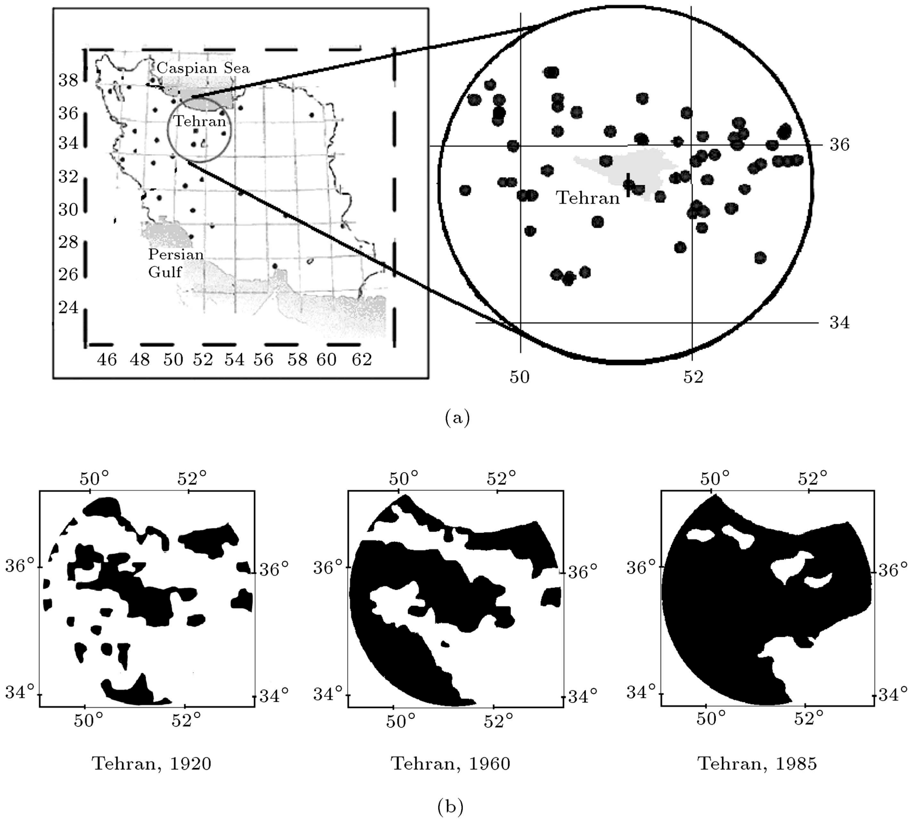

Figure 2. (a) The epicenter location of historical and instrumental recorded events in greater Tehran. (b) Samples of demographic map in greater Tehran, black color represents the residential area.

Table 5. Record probability for Tehran completeness region.

<table>
<tr><th rowspan="2">Time period</th><th colspan="3">Magnitude range</th></tr>
<tr><th>5.0-6.0</th><th>6.0-7.0</th><th>≥ 7</th></tr>
<tr><td>0-855</td><td>0</td><td>0.0258</td><td>0.0539</td></tr>
<tr><td>855-1601</td><td>0.0044</td><td>0.0184</td><td>0.0935</td></tr>
<tr><td>1601-1930</td><td>3.67E-10</td><td>0.0609</td><td rowspan="4">1.0</td></tr>
<tr><td>1930-1965</td><td rowspan="3">1.0</td><td rowspan="3">1.0</td></tr>
<tr><td>1965-1990</td></tr>
<tr><td>1990-2012</td></tr>
</table>

Table 6. Annual record probability for Tehran completeness region.

<table>
<tr><th rowspan="2">Time period</th><th colspan="3">Magnitude range</th></tr>
<tr><th>5.0-6.0</th><th>6.0-7.0</th><th>≥ 7</th></tr>
<tr><td>0-855</td><td>0</td><td>0.9957</td><td>0.9965</td></tr>
<tr><td>855-1601</td><td>0.9927</td><td>0.9946</td><td>0.9968</td></tr>
<tr><td>1601-1930</td><td>0.9362</td><td>0.9915</td><td rowspan="4">1.0</td></tr>
<tr><td>1930-1965</td><td rowspan="3">1.0</td><td rowspan="3">1.0</td></tr>
<tr><td>1965-1990</td></tr>
<tr><td>1990-2012</td></tr>
</table>

Eq. (3) are given in Table 6. Given these values, detection probability in the time interval between two recorded earthquakes can be calculated based on Eq. (6).

By the presented procedure, the catalog can be completed in different time periods. Table 7 presents the complete catalog in greater Tehran. The values of the weights in this table indicate the heterogeneity of the catalog. The mean value and standard deviation of b-value through weighted bootstrap sampling procedure (10,000 simulations) are equal to 1.093 and 0.106, respectively.

The mean seismic hazard curve in the centre of greater Tehran is shown in Figure 3. The two different attenuation relationships of Ambraseys and Bommer [33] and Sarma and Srbulov [34] were employed using logic-tree method in seismic hazard calculation. In Figure 3, the mean curve using the seismic parameters computed by present procedure is compared with the mean curve using the conventional PSHA. As expected, variation of the b-value has a significant impact on the hazard curve. This issue is critical in the design of special and important structures with long design life time [35,36]. As shown in Figure 3, in Tehran, considering the uncertainty in seismic parameter causes an increase in the mean value of seismic hazard curve.

Table 7. Complete catalog in greater Tehran (magnitude scale M_w).

Time	Magnitude	Weight	Time	Magnitude	Weight	Time	Magnitude	Weight	Time	Magnitude	Weight
743	7.2	1	1127	6-7	0.201	1809	6.5	1	1961	5.1	1
	5-6	0.999		≥ 7	0.084		5-6	0.757	1962	7	1
	6-7	0.610	1301	6.7	1	1825	6.7	1	1964	5.3	1
	≥ 7	0.384		5-6	0.999		5-6	0.170	1966	5	1
855	7.1	1		6-7	0.894	1830	5.4	1	1968	5.5	1
	5-6	8.0E-05		≥ 7	0.687		5-6	0.998	1968	5	1
	6-7	7.8E-05	1485	7.2	1	1876	5.8	1	1971	5.6	1
	≥ 7	3.1E-05		5-6	0.123		5-6	0.687	1974	5.1	1
856	7.4	1		6-7	0.008	1890	5.5	1	1977	5.3	1
	5-6	0.073		≥ 7	0.003		5-6	0.170	1980	5.3	1
	6-7	0.005	1495	5.9	1	1895	5.4	1	1980	6	1
	≥ 7	0.002		5-6	1		5-6	0.233	1982	5.6	1
864	5.5	1		6-7	0.073	1901	5.6	1	1983	5.4	1
	5-6	1		≥ 7	0.005		5-6	0.951	1988	5.3	1
	6-7	0.998	1600		1	1930	5.5	1	1988	5.1	1
	≥ 7	0.980		5-6	0.999	1935	5.6	1	1988	5	1
958	7.4	1		6-7	0.608	1937	5.7	1	1990	5.9	1
	5-6	0.999		≥ 7	0.334	1940	5.4	1	1991	5.1	1
	6-7	0.536	1608	7.4	1	1945	5.2	1	1993	5	1
	≥ 7	0.278		5-6	0.999	1951	5.4	1	1998	5.1	1
1052	6.8	1	1678	6.5	1	1951	5.1	1	2001	5	1
	5-6	0.998		5-6	0.990	1952	5.2	1	2002	5.1	1
	6-7	0.332	1721	7.1	1	1954	5.1	1	2002	5.2	1
	≥ 7	0.149		5-6	0.999	1957	5.7	1	2002	6.4	1
1119	6.5	1	1803	5.5	1	1957	5.2	1	2002	5.7	1
	5-6	0.073		5-6	0.170	1957	6.7	1	2004	6.3	1
	6-7	0.005	1808	6.6	1	1958	5.4	1	2007	5.9	1
	≥ 7	0.002		5-6	0.007	1958	5.1	1			
1127	6.8	1	1808	5.9	1	1959	5.2	1			
	5-6	0.986		5-6	0.001	1960	5.4	1			

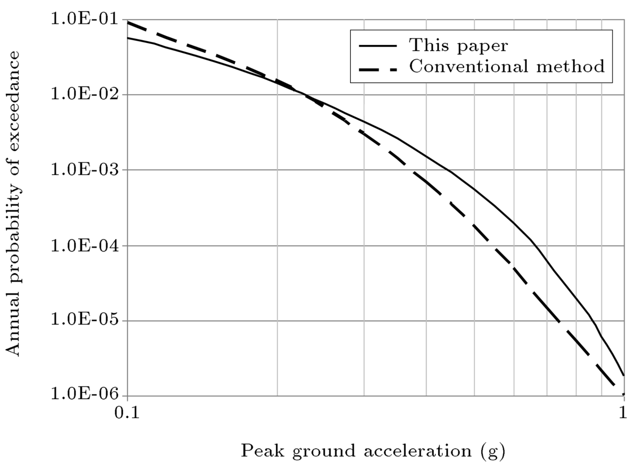

Figure 3. Hazard curve for Tehran.

5. Conclusion

A procedure for improving the accuracy of the PSHA, by entering the seismic data in the complete time period was presented. In order to use the earthquake catalog data through the whole period, the record probability definition is offered by definition of the detection probability. The SSSC was generated based on the proposition of a generic method for catalog completion. The missing earthquake occurrence probabilities were calculated using Bayes' formula.

Assigning weights to events based on occurrence probabilities, the mean values and standard deviations of b-value and consequently the PSHA curve were calculated based on the Bootstrap sampling method. The applicability, simplicity, and efficiency as shown in the case study, are the main advantages of the presented procedure. The standard deviation of b-value can be calculated, thus confidence intervals of SHA curves can be calculated.

Acknowledgment

The authors would like to thank anonymous reviewers for their comments which helped to improve the manuscript.

References

1. Cornell, C.A. "Engineering seismic risk analysis", *Bull. Seism. Soc. Am.*, **58**(5), pp. 1583-1606 (1968).
2. McGuire, R.K. "Effects of uncertainty in seismicity on estimates of seismic hazard for the East Coast of the United States", *Bull. Seism. Soc. Am.*, **66**(3), pp. 827-848 (1977).
3. Jalali, S.A., Banazadeh, M., Abolmaali, A. and Tafakori, E. "Probabilistic seismic demand assessment of steel moment frames with side-plate connections", *Scientia Iranica*, **19**(1), pp. 27-40 (2012).
4. Mayer-Rosa, D. "The assessment of seismic hazard and GSHAP aspects". *Scientia Iranica*, **2**(2), pp. 137-143 (1995).
5. Gupta, I.D. "The state of the art in seismic hazard analysis", *ISET Journal of Earthquake Technology*, **39**(4), pp. 311-346 (2002).
6. Mulargia, F. and Geller, R.J. (Eds.), *Earthquake Science and Seismic Risk Reduction*, Springer (2003).
7. Husen, S. and Hardebeck, J.L. "Earthquake location accuracy", *Community Online Resource for Statistical Seismicity Analysis*, Doi: 10.5078/corssa-55815573. Available at http://www.corssa.org (2010).
8. Dziewonski, A.M. and Woodhouse, J.H. "An experiment in systematic study of global seismicity: Centroid-moment tensor solutions for 201 moderate and large earthquakes of 1981", *J. Geophys Res-Sol. Ea.*, **88**(B4), pp. 3247-3271 (1983).
9. Kagan, Y.Y. "Accuracy of modern global earthquake catalogs", *Phys. Earth Planet. Inter.*, **135**(2-3), pp. 173-209 (2003).
10. Kagan, Y.Y. "Earthquake spatial distribution: The correlation dimension", *Geophys J. Int.*, **168**(3), pp. 1175-1194 (2007).
11. Yazdani, A. and Kowsari, M. "Bayesian estimation of seismic hazards in Iran", *Scientia Iranica*, **20**(3), pp. 422-430 (2013).
12. Corral, A. "Dependence of earthquake recurrence times and independence of magnitudes on seismicity history", *Tectonophy*, **424**(3-4), pp. 177-193 (2006).
13. Matsumura, K. "Stationary seismic hazard in Japan based on the seismo-tectonics", *Proc.11th World Conf. on Earthq. Engin.*, Paper No 715, Elsevier, Amesterdam (1996).
14. Haj Najafi, L., Tehranizadeh, M. and Banazadeh, M. "Reliability-based evaluation of load factors using beta unzipping and Bayesian methods", *Scientia Iranica*, **20**(4), pp. 1175-1192 (2013).
15. Rotondi, R. and Garavaglia, E. "Statistical analysis of the completeness of a seismic catalogue", Nat. Hazards, **25**(3), pp. 245-258 (2002).
16. Mignan, A. and Woessner, J. "Estimating the magnitude of completeness for earthquake catalogs", *Community Online Resource for Statistical Seismicity Analysis*, Ver. 1 (2012). Available at http://www.corssa.org.
17. Sandri, L. and Marzocchi, W. "A technical note on the bias in the estimation of the b-value and its uncertainty through the least squares technique", *Ann. Geophys.*, **50**, pp. 329-339 (2007).
18. Lamarre, M., Townsend, B. and Shah, H.C. "Application of the bootstrap method to quantify uncertainty in seismic hazard estimates", *Bull. Seism. Soc. Am.*, **82**(1), pp. 104-119 (1992).
19. Dong, W., Shah, H.C., Bao, A. and Mortgat, C.P. "Utilization of geophysical information in Bayesian seismic hazard model", *Int. J. Soil Dyn. Earthquake Eng.*, **3**(2), pp. 103-111 (1984).
20. Dong, W.M., Bao, A.B. and Shah, H.C. "Use of maximum entropy principle in earthquake recurrence relationships", *Bull. Seism. Soc. Am.*, **74**(2), pp. 725-737 (1984).
21. Kijko, A. and Sellevoll, M.A. "Estimation of earthquake hazard parameters for incomplete and uncertain data files", *NAT. HAZARDS*, **3**(1), pp. 1-13 (1990).
22. Boyd, O.S. "Including foreshocks and aftershocks in time-independent probabilistic seismic-hazard analyses", *Bull. Seism. Soc. Am.*, **102**(3), pp. 909-917 (2012).
23. Shcherbakov, R., Goda, K., Ivanian, A. and Atkinson, G.M. "Aftershock statistics of major subduction earthquakes", *Bull. Seism. Soc. Am.*, **103**(6), pp. 3222-3234 (2013).
24. Efron, B. "Bootstrap methods: Another look at the jackknife", *Ann. Stat.*, pp. 1-26 (1979).
25. Tchalenko, J.S., Berberian, M., Iranmanesh, H., Bailly, M. and Arsovsky, M. "Tectonic framework of the Tehran region", *Materials for the Study of Seismotectonics of Iran, North-Central Iran. Geol. Surv.*, Iran, Rep, 29, pp. 7-46 (1974).
26. Tavakoli, B. and Ghafory-Ashtiany, M. "Seismic hazard assessment of Iran", *Annali Di Geofisica 42, The Global Seismic Hazard Assessment Program (GSHAP)* 1992-1999, pp. 1013-1021 (1999).
27. Ghodrati Amiri, G., Motamed, R. and Rabet Ehhaghi, H. "Seismic hazard assessment of metropolitan Tehran, Iran", *J. Earthquake Eng.*, **7**, pp. 347-372 (2003).
28. Yazdani, A. and Abdi, M.S. "Stochastic modeling of

earthquake scenarios in greater Tehran", *J. Earthquake Eng.*, **15**(2), pp. 321-337 (2011).

29. Ambraseys, N.N. and Melville, C.P., *A History of Persian Earthquakes*, Cambridge Univ. Press, New York (1982).
30. Berberian, M., *Natural Hazards and the First Earthquake Catalogue of Iran: Historical Hazards in Iran Prior to 1900*, International Institute of Earthquake Engineering and Seismology, **1**, pp. 603-620 (1994).
31. Shahvar, M.P., Zare, M. and Castellaro, S. "A unified seismic catalog for the Iranian plateau (1900-2011)", *Seismological Research Letters*, **84**(2), pp. 233-249(2013).
32. Gholipour, Y., Bozorgnia, Y., Rahnama, M., Berberian, M. and Shojataheri, J. "Probabilistic seismic hazard analysis, Phase I - Greater Tehran regions", Final Report, Faculty of Engineering, University of Tehran, Tehran (2008). Availabe at www.iranhazard.mporg.ir
33. Ambraseys, N.N. and Bommer, J.J. "The attenuation of ground accelerations in Europe", *Earthq. Eng. Struct. Dyn.*, **20**(12), pp. 1179-1202 (1991).
34. Sarma, S.K. and Srbulov, M. "A simplified method for prediction of kinematic soil-foundation interaction effects on peak horizontal acceleration of a rigid foundation", *Earthq. Eng. Struct. Dyn.*, **25**(8), pp. 815-836 (1996).
35. Bakhshi, A. and Asadi, P. "Probabilistic evaluation of seismic design parameters of RC frames based on fragility curves", *Scientia Iranica*, **20**(2), pp. 231-241 (2013).
36. Mirzaei, R. and Bahar, O. "A time varying optimal algorithm for active structural response control", *Sceintia Iranica*, **21**(6), pp. 1833-1844 (2014).

Biographies

Azad Yazdani, PhD, is Associate Professor in the Department of Civil Engineering, University of Kurdistan, Sanandaj, Iran. His research interests lie in the intersection of the general fields of seismology and structural engineering. He has published around 45 technical articles.

Ahmad Nicknam, PhD, is Assistant Professor in the Department of Civil Engineering at Iran University of Science and Technology, Tehran, Iran. He has published numerous papers in international journals and conferences as authored/co-authored in the fields of earthquake simulation at near and far source sites including directivity effects, nonlinear dynamic analysis and performance evaluation of structures, retrofitting structures, deterministic and probabilistic USHA.

Mostafa Khanzadi is an Assistant Professor of Civil Engineering with over 25 years of professional experience in civil and structural engineering. His main expertise and experience are in the fields of structural analysis, concrete technology and construction management.

Sasan Motaghed is a PhD student in the Department of Civil Engineering at Iran University of Science and Technology, Tehran, Iran, where he also received his MSc and BSc degrees. His research interests include: seismic risk analysis, statistical seismology, evaluating the seismic behavior of reinforced concrete structures. He has been working as a lecturer since 2006 in Behbahan Khatam Alanbia University of Technology.