Quantifying Seismic Parameter Uncertainties and Their Effects on Probabilistic Seismic Hazard Analysis: A Case Study of Iran

MILAD KOWSARI, NASROLLAH EFTEKHARI, ANDRZEJ KIJKO, EHSAN YOUSEFI DADRAS, HAMED GHAZI, and ELHAM SHABANI

Abstract—Probabilistic seismic hazard analysis (PSHA) can accommodate various sources of uncertainties and it provides a rational framework for the precise portrayal of the hazard of a given region. Often, the information used in the PSHA is incomplete and uncertain; therefore, the question arises how the uncertainty of the input data affects the estimated hazard characteristics. In this study, sensitivity analysis (SA) was conducted to identify the most dominant inputs affecting the assessment of the key seismicity parameters (SPs), including the mean seismic activity rate $k$, $b$ value of Gutenberg–Richter, and the maximum possible earthquake magnitude $m_{\text{max}}$. The study was applied in five areas of Iran, for which such analyses have not been conducted in previous studies. Subsequently, Monte Carlo simulation was employed to determine the effects of the uncertain input parameters on PSHA relevant to spectral accelerations corresponding to 10% and 2% probability of exceedance at least once in 50 years. For this purpose, a unified and declustered earthquake catalogue was used for the five major seismotectonic provinces of Iran (Alborz-Azarbayejan, Zagros, Central-East Iran, Koppeh Dagh, and Makran). The results showed that the last (complete) part of the catalogue has a significant effect on the estimated value of seismic activity and the $b$ value. In contrast, its influence is insignificant on the area-characteristic maximum possible earthquake magnitude, for which the most influential inputs are the maximum observed earthquake and its uncertainty. Furthermore, the uncertainties of the input SPs affected the seismic hazard estimates substantially and led to significant variability in the estimated ground motion characteristics.

Key words: PSHA, sensitivity analysis, uncertainty analysis, seismicity parameters.

1. Introduction

Probabilistic seismic hazard analysis (PSHA) provides estimates of the probability of exceedance (PE) of a ground motion intensity measure at a given site during a given period of exposure. The input parameters required to assess PSHA are determined from historical, instrumental, and geological observations (Cornell 1968; Sokolov and Wenzel 2015). The seismicity parameters (SPs), including the slope of the Gutenberg–Richter (G–R) relationship ($b$ value = $b$/ln10), earthquake activity rate ($k$), and the maximum magnitude ($m_{\text{max}}$), are the principal inputs of any PSHA. The estimates of the SPs can be established by studying the spatial and temporal distribution of the earthquakes, as well as the frequency of occurrence of the earthquake magnitudes (i.e., Gutenberg–Richter law). In most PSHA studies, these parameters are assumed to remain constant over time and space; however, their significant spatial and temporal variations are evident (Zolfaghari 2015; Khodaverdian et al. 2016; Yadav 2016; Kazemian and Hatami 2017; Kazemi-Beydokhti et al. 2017). Such spatial and temporal differences combined with the uncertainties in occurrence time, size, and location of the recorded earthquakes can affect the reliability of the seismic hazard assessment significantly (Kagan 2003; Woessner and Wiemer 2005).
The uncertainties in PSHA are divided commonly into two categories, namely, aleatory and epistemic. Aleatory variability derives from an apparent randomness in nature that is irreducible. On the other hand, epistemic uncertainty derives from a lack of knowledge of the appropriate model and its parameterisation. The epistemic uncertainties can be divided further into parameter and model uncertainties. Parameter uncertainty stems from difficulties in estimating the parameters due to the limited number, poor representativeness, and imprecision of data, whereas model uncertainty is subject to the choice of variables, dependencies, assumptions, and the processes (Rohmer et al. 2014). Generally, the uncertainties related to the SPs are categorised as the epistemic uncertainty (McGuire 2004). Therefore, the uncertainty of the input parameters should be analysed to evaluate their effects on the variability of the PSHA results. Traditionally, such analysis consists of choosing the distribution that represents the extent of uncertainty in the inputs (i.e., uncertainty modelling), then the extent of uncertainty in the outputs is computed (i.e., uncertainty propagation) and finally the contribution of each model inputs to the output variability is determined (De Rocquigny et al. 2008). The latter technique is known as sensitivity analysis (SA). In other words, the SA can be applied to study (qualitatively or quantitatively) how changes in the model input affect the resulting response (Saltelli et al. 2000).

Iran is a highly seismic-prone country that frequently suffers destructive and catastrophic earthquakes, causing significant loss of life and widespread damage (Yazdani and Kowsari 2013). To boost confidence in the assessment of seismic hazard in the country, it is vital to understand the effects of SP uncertainties. Several such SP studies have been conducted in Iran (Mirzaei et al. 1997; Tavakoli and Ghafory-Ashtiany 1999; Bastami and Kowsari 2014; Kalaneh and Agh-Atabei 2016; Khodaverdian et al. 2016; Madahizadeh et al. 2016; Mohammadi et al. 2016; Salamat et al. 2017; Raeesi et al. 2017); however, the effect of SP uncertainties on the PSHA estimates has not always been quantified explicitly. Therefore, we performed uncertainty analysis (UA) and SA to determine the contribution of the data and to establish the most significant parameters affecting the estimated SPs. Furthermore, UA was conducted to show how the uncertainties in the SPs could affect the PSHA results. Our study results revealed the primary sources of hazard uncertainties and could help to prioritize efforts to reduce hazard variability.

2. The Study Region and Data Used

The Iranian Plateau lies in the Alpine seismic and orogenic belt that frequently experiences destructive earthquakes. Over the past four decades, such catastrophic events have caused more than 100,000 casualties. The seismic activity stems primarily from the location of the area, i.e., a 1000-km-wide zone that accommodates the 35-mm/yr convergence rate between the Eurasian and Arabian plates by strike-slip and reverse faults (Berberian and Yeats 1999; Engdahl et al. 2006). As the seismic characteristics of all the areas of this plateau are not similar (Ansari et al. 2009), the whole area may not be characterized as a single unit for the study of seismic hazard. Therefore, it is reasonable to divide the region into subregions with similar seismological characteristics, called seismotectonic provinces. Such subregions can be defined as geographic areas with a comparable tectonic setting, and geophysical and seismological similarity, as well as a unified seismicity pattern. Several researchers have studied the seismotectonic provinces of Iran, including Stocklin (1968), Takin (1972), Berberian (1976), Nowroozi (1976), Tavakoli (1996), and Mirzaei et al. (1998). Mirzaei et al. (1998) divided Iran into five major seismotectonic provinces, namely Alborz-Azarbaijan, Zagros, Central-East Iran, Koppeh Dagh, and Makran, as shown in Fig. 1. This relatively simple model was derived from the most recent study of the available earthquake catalogue and the tectonic and regional geomorphology and seismicity pattern.

A homogenous earthquake catalogue is required to perform appropriate PSHA for a region. As the earthquake catalogue in Iran is based on different magnitude scales, converting these magnitude scales to facilitate interpreting all the earthquakes at the same scale is mandatory. Karimiparidari et al. (2013) employed the orthogonal regression technique to compile a seismic catalogue for Iran by converting
the diverse types of magnitude into moment magnitude. Shahvar et al. (2013) published uniform earthquake catalogues by dividing the territory of the Iranian Plateau into two domains to derive empirical relationships in order to convert the original magnitudes to a uniform scale. More details on and a discussion of their study can be found in Mirzaei et al. (2014) and Shahvar et al. (2014).

More recently, Mousavi-Bafrouei et al. (2015) provided a new earthquake catalogue of Iran with all the magnitudes de-clustered and converted into a moment magnitude ($M_w$) catalogue. The time- and distance-window algorithm of Gardner and Knopoff
(1974) and Uhrhammer (1986), and the cluster method proposed by Reasenberg (1985) were used to eliminate the dependent shocks. The entire magnitude range (EMR) method (Woessner and Wiemer 2005) was used to evaluate the magnitude of completeness of the seismotectonic provinces proposed by Mirzaei et al. (1998). In the current study, we used the earthquake catalogue provided by Mousavi-Bafrouei et al. (2015) because of their detailed analysis of each seismotectonic province. Table 1 shows the magnitude of completeness and $b$ values for instrumental earthquakes that occurred during three periods, namely 1900–1963, 1964–1996 (i.e., after the development of seismic networks in Iran), and 1997–2012 in each and all major seismotectonic provinces of Iran. This assessment of completeness of the Iranian earthquake catalogue has been confirmed by the most recent studies (e.g., Talebi et al. 2017).

3. Estimation of Seismicity Parameters

The SPs can be determined in various ways, but, after the fundamental work by Aki (1965), the maximum likelihood procedure has been the most popular method for assessing SP. Unfortunately, the simple estimator of the $b$ value, as suggested by Aki (1965), can be applied only when the seismic event catalogue is complete. Several attempts have been made to extend the Aki $b$ value estimator in instances of incomplete catalogues and, specifically, when the level of completeness varies with time (Molchan et al. 1970; Rosenblueth 1986; Rosenblueth and Ordaz 1987; Kijko and Smit 2012). Of these, the best-known and most applied method is the estimator derived by Weichert (1980). The most common approach to estimating SPs in instances where the seismic event catalogue contains historic (largest events only) and instrumental (complete catalogues, with varying level of completeness) is the procedure by Kijko and Sellevoll (1989, 1992).

Kijko and Sellevoll (1989, 1992) introduced a maximum likelihood procedure to estimate SP (referred to as KS-1 and KS-2, respectively) that combines the incomplete (historic) and complete parts of a catalogue, with an account of the uncertainty of the earthquake magnitude. In addition, the procedure can accommodate the observational gaps in the catalogue and the common phenomenon that the completeness of the catalogue varies with time. Despite KS-2 requiring numerical integration, the approach is applied widely (e.g., Bayrak et al. 2009; Kalaneh and Agh-Atabai 2016; Khodaverdian et al. 2016; Hamdache et al. 2017). However, the flexible and generic KS-1 and KS-2 procedures do not address the problem of the uncertainties associated with the applied earthquake occurrence models. This problem has been addressed recently by the introduction of simple, mixed (compound) distributions and the treatment of both the $\lambda$ and the $b$ value as random variables, each described by the gamma distribution (Kijko et al. 2016). However, such approach (utilising mixed distributions) is not unique. An alternative method to consider model parameter uncertainties is the application of the Bayesian formalism (Pisarenko and Lyubushin 1999; Tsapanos et al. 2001; Tsapanos 2003; Tsapanos and Christova 2003; Yadav et al. 2013; Yazdani and Kowsari 2013).

Table 1

<table>
<thead>
<tr>
<th>Seismotectonic province</th>
<th>Magnitude of completeness ($M_c$)</th>
<th>$b$ value$^a$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alborz-Azarbayejan</td>
<td>5.0</td>
<td>4.4</td>
</tr>
<tr>
<td>Central-East Iran</td>
<td>5.7</td>
<td>4.5</td>
</tr>
<tr>
<td>Kopeh Dagh</td>
<td>4.8</td>
<td>4.7</td>
</tr>
<tr>
<td>Makran</td>
<td>5.4</td>
<td>4.7</td>
</tr>
<tr>
<td>Zagros</td>
<td>5.6</td>
<td>4.9</td>
</tr>
<tr>
<td>Iran</td>
<td>5.7</td>
<td>4.7</td>
</tr>
</tbody>
</table>

$^a_b$ value = $b$/ln(10)
After adopting the assumption that the earthquake occurrence in time follows the Poisson distribution and the earthquake magnitude is distributed according to a doubly truncated G–R relation, the seismic activity rate can be written as (McGuire and Arabasz 1990):

\[
\lambda(m) = \lambda(m_{\text{min}}) \exp[-\beta(m - m_{\text{min}})] - \exp[-\beta(m_{\text{max}} - m_{\text{min}})] / (1 - \exp[-\beta(m_{\text{max}} - m_{\text{min}})])
\]

where \( m_{\text{min}} \) is the level of magnitude completeness and \( m_{\text{max}} \) is the area-characteristic maximum possible earthquake magnitude. Usually, earthquake catalogues are incomplete and riddled with errors in respect of magnitude determination. However, worldwide, in most earthquake catalogues, both incompleteness and the magnitude determination errors decrease with time.

Among the SPs, \( m_{\text{max}} \) is considered a key parameter in hazard studies. Any PSHA requires an appropriate estimate of \( m_{\text{max}} \) for each source zone to avoid the inclusion of unrealistically large earthquakes (Wheeler 2009). Although there is no generally accepted technique to estimate \( m_{\text{max}} \), the two approaches currently used are the deterministic and the probabilistic methods (Kijko 2004). Selecting either the deterministic or the probabilistic method depends on the available information related to factors such as the seismicity, geology, and tectonics of the investigated region. In the deterministic approach, in most cases, the estimates of \( m_{\text{max}} \) are obtained through an empirical relationship, referred to as the source-scaling relationship, which describes the tectonic features and the seismological and geological information for the region of interest (Wells and Coppersmith 1994; Anderson et al. 1996; Stein and Hanks 1998; Field et al. 1999; Somerville et al. 1999; Hanks and Bakun 2002; Shaw 2009; Leonard 2010; Shaw 2013). On the other hand, in the probabilistic approach, \( m_{\text{max}} \) can be estimated by applying an appropriate statistical procedure, as well as the information related to the seismicity of the region being investigated. Based on the available information on earthquake magnitude distribution, Kijko (2004) derived a generic equation for the estimation of \( m_{\text{max}} \) in three different instances. These are (1) when the magnitudes are distributed according to the ‘classic’, doubly truncated G–R relation, (2) when the empirical magnitude distribution deviates moderately from the G–R relation, and (3) when no specific type of magnitude distribution is assumed. The ‘classic’ Kijko–Sellevoll probabilistic procedure can be estimated as follows:

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

where \( n \) denotes the number of seismic events, \( n_1 = n/[1 - \exp(-\beta(m_{\text{max}} - m_{\text{min}}))] \), \( n_2 = n_1 \exp(-\beta(m_{\text{max}} - m_{\text{min}})) \), and \( E_1(z) \) is an exponential integral function, which can be approximated as \( E_1(z) = \frac{1.8333z^{\frac{1}{2}} + 0.25}{z^{\frac{1}{2}} + 3.3333 + 1.68} \exp(-z) \). Therefore, the maximum possible earthquake in the region is equal to the maximum observed magnitude plus a positive correction factor. The value of the correction depends on the investigated area—the longer the catalogue is, the smaller is the correction factor (Kijko and Singh 2011).

Despite statistical procedures (particularly based on the mathematical formalism of extreme values, see, e.g., Vermeulen and Kijko 2017; Beirlant et al. 2018) providing powerful tools to evaluate \( m_{\text{max}} \), they have one weak point, namely that the available seismic event catalogues are often too short and are not able to provide reliable estimates of \( m_{\text{max}} \). Therefore, Bayesian procedures are considered superior, as they allow the inclusion of alternative independent information such as local geology, tectonics, geophysical data, palaeoseismicity, similarity to another seismic area, and the like. In most instances, these procedures can provide assessments of \( m_{\text{max}} \) that are more reliable than any other procedure. Unfortunately, the frequently used Bayesian procedure derived by Cornell (1994) contains a mathematical flaw and should therefore be used with particular caution (Kijko 2012).

4. Sensitivity Analysis

SA can reveal the relative contribution of each input variable (parameter) to the output variability, either qualitatively or quantitatively. To simplify the
analysis problem, the uncertain variables that do not contribute much to the uncertainty in the model output can be fixed at their best estimates, rather than be treated as random variables. On the other hand, the variables that contribute significantly to the overall output uncertainty should be analysed properly to reduce the variability (Porter et al. 2002). These SA methods are categorised often as either local or global. In the local analysis, the sensitivity of each uncertain input is computed by keeping the other variables fixed and varying only the certain input variable in a local area around its nominal value (Gustafson et al. 1996). The local SA methods are used frequently because of their relatively low computational burden. However, one of the negatives of the local SA approach is that the sensitivity study is performed at the central estimate of input variables, whereas the results could be quite different at another point nearby. The local SA methods do not account for the interactions between the input variables (Saltelli et al. 2010). Moreover, these methods do not quantify the difference in the importance of a given unknown input variable compared with that of another. In other words, they provide qualitative sensitivity measures (i.e., the input factors are ranked simply in order of significance; Feyissa et al. 2012). In contrast, the global SA apportions the output variability to the variability of the input variables when they vary over their entire uncertainty domain. Therefore, the quantitative relative importance of each input variable can be measured by using the global SA approach.

In this study, a global SA method based on information-theory tools was used to quantify the relationship between the input parameters and the output distribution (Auder and Iooss, 2008). One of the major advantages of our chosen SA approach is its shorter computational time compared with that of the most often used variance-based global SA method.

Following Shannon (2001), let us assume that an observation $x$ from the domain $X$, with probability $p(x)$, contains information of $- \log_2 p(x)$ bits. The entropy $H(X)$ of the probability distribution $p(x)$ that describes the uncertainty in a random variable because of input errors is of the form

$$H(X) = - \int_{x \in X} p(x) \log_2 p(x) dx. \quad (3)$$

Following (3), the conditional entropy of the model parameters $\Theta$, given all possible observations $x$, can be quantified as

$$H(\Theta|X) = \int_{x \in X} p(x) H(\Theta|x) dx = - \int_{x \in X} \int_{\theta \in \Theta} p(x)p(\theta|x) \log_2 p(\theta|x) d\theta dx \quad (4)$$

The mutual information, $I$, quantifies the amount of information held in a random variable through the other random variable. In other words, it is defined as the difference in output uncertainty with and without knowledge of $X$

$$I(\Theta, X) = \int_{\Theta} \int_{X} p(x, \theta) \log_2 \left( \frac{p(x, \theta)}{p(x)p(\theta)} \right) dx d\theta, \quad (5)$$

where $p(x, \theta)$ is the joint probability distribution equal to $p(\theta|x)p(x)$.

Then the mutual information takes the simple form

$$I(\Theta, X) = \int_{\Theta} \int_{X} p(\theta|x)p(x) \log_2 \left( \frac{p(\theta|x)p(x)}{p(x)p(\theta)} \right) dx d\theta$$

$$= \int_{\Theta} \int_{X} p(\theta|x)p(x) \log_2 p(\theta|x) dx d\theta$$

$$- \int_{\Theta} \int_{X} p(\theta|x)p(x) \log_2 p(\theta) dx d\theta$$

$$= - H(\Theta|X) - \int_{\Theta} \log_2 p(\theta) \left[ \int_{X} p(\theta|x)p(x) dx \right] d\theta$$

$$= - H(\Theta|X) \int_{\Theta} \log_2 p(\theta) d\theta = - H(\Theta|X) + H(\Theta) \quad (6)$$

Therefore, the parameter

$$H(X) = - \int_{x \in X} p(x) \log_2 p(x) dx. \quad (3)$$

Following (3), the conditional entropy of the model parameters $\Theta$, given all possible observations $x$, can be quantified as

$$H(\Theta|X) = \int_{x \in X} p(x) H(\Theta|x) dx = - \int_{x \in X} \int_{\theta \in \Theta} p(x)p(\theta|x) \log_2 p(\theta|x) d\theta dx \quad (4)$$

The mutual information, $I$, quantifies the amount of information held in a random variable through the other random variable. In other words, it is defined as the difference in output uncertainty with and without knowledge of $X$

$$I(\Theta, X) = \int_{\Theta} \int_{X} p(x, \theta) \log_2 \left( \frac{p(x, \theta)}{p(x)p(\theta)} \right) dx d\theta, \quad (5)$$

where $p(x, \theta)$ is the joint probability distribution equal to $p(\theta|x)p(x)$.

Then the mutual information takes the simple form

$$I(\Theta, X) = \int_{\Theta} \int_{X} p(\theta|x)p(x) \log_2 \left( \frac{p(\theta|x)p(x)}{p(x)p(\theta)} \right) dx d\theta$$

$$= \int_{\Theta} \int_{X} p(\theta|x)p(x) \log_2 p(\theta|x) dx d\theta$$

$$- \int_{\Theta} \int_{X} p(\theta|x)p(x) \log_2 p(\theta) dx d\theta$$

$$= - H(\Theta|X) - \int_{\Theta} \log_2 p(\theta) \left[ \int_{X} p(\theta|x)p(x) dx \right] d\theta$$

$$= - H(\Theta|X) \int_{\Theta} \log_2 p(\theta) d\theta = - H(\Theta|X) + H(\Theta) \quad (6)$$

Therefore, the parameter
\[
\eta = \frac{I(\Theta, X)}{H(\Theta)} = \frac{-H(\Theta|X) + H(\Theta)}{H(\Theta)} = 1 - \frac{H(\Theta|X)}{H(\Theta)},
\]

can be treated as a sensitivity index (Krzykacz-Hausmann 2001). The defined sensitivity index takes the values between 0 and 1 and describes the effect (contribution) of input random variable \(X\) to the variability of the output.

5. Uncertainty Analysis in PSHA

Conventionally, PSHA is carried out for a specified region to estimate the annual probability of exceeding the ground motion levels for the parameters of interest, e.g., peak ground acceleration (PGA) or spectral acceleration. This process is dependent on various inputs and models related to the identification of possible earthquake sources, description of the seismic activity for each source, selection of appropriate ground motion models (GMMs), and, finally, the integration of all possible earthquake scenarios. However, the uncertainties of the most basic parameters are not given and, likely, are not considered. Therefore, it is vital to have a clear understanding of which parameters affect the hazard results under various circumstances, particularly for applications that build on PSHA, which can be performed by SA and UA.

In PSHA, the aleatory variability can be handled by integrating over the distribution of ground motion amplitudes about the median. On the other hand, the epistemic uncertainty has been modelled by the use of alternative models in a logic tree framework. Another way to account for the epistemic uncertainty in PSHA is the assignment of probability distributions to the model inputs (Molkenthin et al. 2017). Accordingly, in this study, we employed uncertainty analysis by using the Monte Carlo simulation technique and a Latin hypercube sampling method (McKay et al. 1979; Robert 2004) to present an example showing how the epistemic uncertainties associated with the SPs affected the PSHA results. Figure 2 shows schematically the procedure employed for this analysis. Different sources of uncertainties can be identified and quantified in all the PSHA steps, including the characteristics of the seismic sources, definition of the SPs, and the selection of GMMs (Sabetta 2014). In particular, the selection of appropriate GMMs is crucial since the largest uncertainties in PSHA estimations can be caused by uncertainties in GMMs (e.g., Hintersberger et al. 2007; Scherbaum et al. 2005). The uncertainty of the median predicted by different GMMs is epistemic and it is generally incorporated into PSHA by using multiple GMMs in a logic tree or backbone approaches (Atkinson et al. 2014; Bommer et al. 2015; Akkar et al. 2018; Douglas 2018a, b). Therefore, to avoid the inclusion of uncertainties associated with seismic sources and GMMs, a single line source and one GMM were considered in this example. The line source, defined as a uniform earthquake potential source, is shown in step 1, along with a plan view of the site and its orientation to the source. In step 2, different sets of SPs were generated, with the corresponding probability distributions. Subsequently, for each generated SP, the ground motion intensity measure of interest was estimated by the selected GMM. Finally, PSHA calculations were performed for each set of generated ground motions and the different hazard curves were obtained accordingly.

6. Results and Discussion

Any deficiencies in PSHA assessments have their roots in the application of incorrect models, as well as the use of uncertain inputs (Klügel 2008; Musson 2012). Therefore, we investigated the influence of the uncertainties associated with the SPs on the PSHA results. Using the Monte Carlo methods and information theory, UA and SA, respectively, were conducted for the five different seismotectonic provinces of Iran. As regards UA, we performed thousands of MC simulations, with the data for each run in the input catalogue randomly sampled from the predefined probability distributions. The Latin hypercube sampling method was used in the computations to generate random stochastic inputs from the respective probability distributions. The superior efficacy of this method has been proven in terms of convergence and robustness (McKay et al. 1979). However, the analysis should be carried out for
independent input parameters. The relevant scatter diagrams of the SPs and their Pearson correlation coefficients ($\rho$) for Central-East Iran are shown in Fig. 3. Only the results for Central-East Iran are presented here for the sake of brevity; however, the same behaviour was observed for the other seismotectonic provinces. The values close to $\pm 1$ indicated a strong correlation between the two variables. The value close to $+1$ indicated that as one variable increased, the value of the second variable also increased, i.e., the variables were correlated positively. Conversely, the value close to $-1$ showed a negative correlation, meaning that as one variable increased, the value of the other variable decreased.

To perform the information-theoretic sensitivity analysis, a probabilistic approach is used by evaluating the model for multiple sets of randomly and independently selected input values drawn (Lüdtke et al. 2008). From Fig. 3, the high correlation between $b$, $k$, and $m_{\text{min}}$ indicates that these parameters should not be analysed together because many of the SA methods are intended for cases where the inputs are mutually independent (Eggels and Crommelin 2018).

The uncertainty analysis deals with the propagation of uncertainty through a model, from inputs to outputs. In other words, the UA investigates the effect of input parameter uncertainties on the variability of the outputs. In this vein, the uncertainty in the SPs is modelled using a normal distribution with a mean and
standard deviation proposed in Mousavi-Bafrouei et al. (2015). The results of the UA on the SPs are shown in Fig. 4. This figure shows the uncertainty in the output (i.e., the SPs) due to the variability of inputs along with the histograms of SPs and the corresponding fitted, log-normal distributions. However, only the histograms of the SPs in Alborz-Azarbayejan are presented here for the sake of brevity. The results can be used for Monte Carlo PSHA, where the generated synthetic sub-catalogues are based on SP distributions (Musson 1999, 2004; Atkinson 2012; Assatourians and Atkinson 2013).
We used the information theory approach to investigate the influence of the uncertainty in each input variable on the SPs. To estimate the Krzykacz–Hausmann sensitivity indices of entropy and the mutual parameter dependence, we used the nearest-neighbour procedure (Kozachenko and Leonenko 1987; Kraskov et al. 2004). A major advantage of the chosen procedure is its relatively short computational time compared with that of the density estimation method (Huoh 2013). It should be noted that the applied nearest-neighbour algorithm requires a reasonably large sample size. Some preliminary tests showed that a sample size of 5000 is required to assure the reliable assessments of the sensitivity values.

In our analysis, we used the earthquake catalogue provided by Mousavi-Bafrouei et al. (2015). The complete part of the catalogue was divided into periods with different threshold magnitudes, namely 1900–1963 ($M_{c1}$), 1964–1996 ($M_{c2}$), and 1997–2012 ($M_{c3}$). As the applied catalogue is highly inhomogeneous, a special effort was made to assess the uncertainty of the earthquake magnitude determination, as it plays a crucial role in the estimation of the SPs. Different magnitude determination uncertainties were assumed for the historical and the instrumental parts of the catalogue. Regarding the instrumental part of the catalogue, the uncertainty of the $M_w$ magnitudes of some earthquakes has been reported occasionally by various researchers (e.g., Ambraseys 2001) and agencies. Consequently, the effect of the magnitude determination uncertainty was also considered, denoted as MU in Table 2. The results showed that the last complete part of the catalogue ($M_{c3}$), which contains the minimum observed magnitude, played a crucial role in the estimation of the hazard parameters. In other words, the minimum magnitude, which is defined as the lower limit of integration over the earthquake magnitudes, was the crucial parameter input among the threshold magnitudes (Bommer and Crowley 2017).

Various definitions of the maximum possible earthquake magnitude exist; however, in this study, it refers to the upper limit of the earthquake size in a given region. This implies that no earthquake is expected to exceed this maximum possible earthquake magnitude (Kijko 2004). Moreover, there are several probabilistic procedures to estimate the $m_{max}$ (for more details, see Kijko and Singh 2011). All the procedures are based on the underlying principle that the estimated $m_{max}$ value is equal to the maximum observed magnitude plus a positive correction factor. Therefore, in the SA, the effect of model selection is also considered as one of the inputs. The other inputs are the maximum observed magnitude, magnitude uncertainty and the earthquake catalog with different threshold magnitudes for three time intervals, namely 1900–1963 ($M_{c1}$), 1964–1996 ($M_{c2}$), and 1997–2012 ($M_{c3}$). The contribution of each input on the $m_{max}$ is shown in Fig. 5. The results showed that the $m_{max}$ values tended to be controlled by the maximum observed value in the catalogue, as has been shown already by Musson (2004). As most catalogues cover a relatively short time period and the occurrence of $m_{max}$ is rare relative to the period of observation, the probabilistic approaches might underestimate the $m_{max}$ estimates in a region under study. Thus, paleoseismic investigations can be used to get more reliable estimates of the maximum earthquake magnitude (McCalpin 2009). Salamat et al. (2017) calculated the confidence intervals for the maximum magnitude in six seismotectonic zones of Iran. However, their analysis did not provide clear results, as, in some regions, $m_{max}$ was estimated with acceptable confidence, whereas, in others, the uncertainty of the estimated $m_{max}$ was unrealistically high. As the magnitude uncertainty is another

<table>
<thead>
<tr>
<th>Output</th>
<th>Input variables</th>
<th>AA</th>
<th>CE</th>
<th>KD</th>
<th>ZG</th>
<th>MK</th>
<th>IR</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta$</td>
<td>$M_{c1}$</td>
<td>1.48</td>
<td>2.37</td>
<td>0.67</td>
<td>1.37</td>
<td>3.22</td>
<td>1.77</td>
</tr>
<tr>
<td></td>
<td>$M_{c2}$</td>
<td>3.58</td>
<td>2.27</td>
<td>1.27</td>
<td>2.50</td>
<td>10.75</td>
<td>4.07</td>
</tr>
<tr>
<td></td>
<td>$M_{c3}$</td>
<td>91.18</td>
<td>95.12</td>
<td>97.52</td>
<td>95.46</td>
<td>84.95</td>
<td>92.84</td>
</tr>
<tr>
<td></td>
<td>MU</td>
<td>3.77</td>
<td>0.23</td>
<td>0.54</td>
<td>0.66</td>
<td>1.09</td>
<td>1.30</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>$M_{c1}$</td>
<td>4.93</td>
<td>7.48</td>
<td>17.80</td>
<td>20.01</td>
<td>1.95</td>
<td>16.10</td>
</tr>
<tr>
<td></td>
<td>$M_{c2}$</td>
<td>0.68</td>
<td>1.41</td>
<td>2.49</td>
<td>0.41</td>
<td>3.47</td>
<td>1.74</td>
</tr>
<tr>
<td></td>
<td>$M_{c3}$</td>
<td>74.54</td>
<td>84.69</td>
<td>71.61</td>
<td>75.67</td>
<td>91.88</td>
<td>75.92</td>
</tr>
<tr>
<td></td>
<td>MU</td>
<td>18.11</td>
<td>4.22</td>
<td>7.21</td>
<td>1.32</td>
<td>1.10</td>
<td>3.24</td>
</tr>
<tr>
<td>$m_{maxobs}$</td>
<td></td>
<td>1.75</td>
<td>2.19</td>
<td>0.89</td>
<td>2.56</td>
<td>1.59</td>
<td>2.97</td>
</tr>
</tbody>
</table>

AA Alborz-Azarbayejan, CE Central-East Iran, KD Kopeh Dagh, ZG Zagros, MK Makran, IR Iran, MU magnitude uncertainty in earthquake catalogue
important parameter that affects the estimation of $m_{\text{max}}$, the earthquake catalogue has to be analysed, checked, and prepared meticulously. The choice of the earthquake occurrence model, as well as the choice of the estimation procedure (Kijko and Singh 2012; Vermeulen and Kijko 2017; Beirlant et al. 2018), can also affect the estimated value of $m_{\text{max}}$. As can be expected, the choice of the level of completeness has a marginal effect on the estimated $m_{\text{max}}$.

In this study, PSHA was applied to a hypothetical site located in Iran. To show the effect of the uncertainties associated with the SPs, we used a single line source and specified GMM proposed by Akkar and Bommer (2010). The site was assumed to be located 30 km from a single fault, which could produce earthquakes up to magnitude 7 (Fig. 2). Figure 6 shows the results of the uncertainty analysis on PSHA for spectral accelerations corresponding to 10 and 2% PE in 50 years. The figure illustrates the mean $\pm$ 2 standard deviations of the uniform acceleration response spectra. The values of the standard deviation varied from 0.11 to 0.35 in the base 10 logarithms, indicating that the uncertainty associated with the SPs increased with an increase in the
structural periods. In other words, the predicted ground motion was more sensitive at higher periods than at lower ones. Such a dependency on the structural periods has been pointed out previously, e.g., by Sokolov et al. (2004, 2009). In addition, we observed that when the ground motion levels changed (i.e., for 10% and 2% PE in 50 years), the epistemic uncertainties of the PSHA estimates remained unchanged. As a result, the uncertainties of the SPs affected the hazard curves significantly and led to considerable variability in the intensity measures of ground motion.

7. Conclusion

In PSHA, all possible earthquake magnitudes, distances, and predicted ground motions are considered, along with their associated uncertainties. However, it is important to assess the effect of the PSHA uncertainties associated with the uncertainty of the input data. Furthermore, estimating the SPs requires knowledge of the different information, which is always uncertain. In the first part of this study, UA and SA were performed to identify the most dominant factors that affected the estimated SPs. Both UA and SA were conducted for five major seismotectonic provinces in Iran that have diverse seismological characteristics. However, the SA showed quite similar behaviour in all the analysed regions. Among the input variables, including different threshold magnitudes, maximum magnitude, and magnitude uncertainty in the earthquake catalogue, it was demonstrated that the last (most complete) part of the catalogue ($M_{c3}$) had a most significant effect on the estimates of the $\beta$ and $\lambda$ parameters. However, regarding $m_{\text{max}}$, the maximum observed earthquake and the magnitude determination uncertainty were the most influential factors.

In the second part of the study, the effect of the uncertain input parameters on the PSHA was investigated by using the Monte Carlo simulation. The analysis showed that the uncertain SPs had a significant effect on the estimated hazard characteristics. The effect of the uncertainty of the SP parameters on the seismic hazard curves for two analysed ground motion levels was fairly similar; however, this increased gradually with the increasing structural periods.
In conclusion, the quality of the input data is the key factor that determines the quality of the seismic hazard estimates. Unfortunately, this obvious factor is often ignored in calculating and interpreting seismic hazard characteristics. In earthquake engineering, accounting for uncertainties in a seismic hazard is costly; therefore, the value of this study lies in helping to improve the understanding of the source and degree of such uncertainties.

REFERENCES


Kijko, A. (2004). Estimation of the maximum earthquake magni-
Khodaverdian, A., Zafarani, H., Rahimian, M., & Dehnamaki, V.
Klante, S., & Agh-Atabai, M. (2016). Spatial variation of earth-
Kazemian, J., & Hatami, M. R. (2017). Temporal variations of se-
Khadaverdian, A., Zafarani, H., Rahimian, M., & Dehnamaki, V.
Kijko, A. (2004). Estimation of the maximum earthquake magni-
Kijko, A. (2012). On Bayesian procedure for maximum earthquake
Kijko, A., & Singh, M. (2011). Statistical tools for maximum possi-
Madahizadeh, R., Mostafazadeh, M., & Ansari, A. (2016). Long-
quake Engineering Research Institute.
Mirzaei, N., Shabani, E., & Bafrouei, S. H. M. (2014). Comment on “A Unified Seismic Catalog for the Iranian Plateau (1900–2011)” by Mohammad P. Shahvar, Mehdi Zare, and Sil-


(Received July 25, 2018, revised November 1, 2018, accepted November 14, 2018, Published online November 30, 2018)