

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa



Evidence in support of seismic hazard following Poisson distribution



J.P. Wang^{a,*}, Su-Chin Chang^b

- a Department of Civil & Environmental Eng., The Hong Kong University of Science and Technology, Hong Kong
- ^b Department of Earth Sciences, The University of Hong Kong, Hong Kong

HIGHLIGHTS

- Evidence in support of seismic hazard following Poisson distribution.
- A novel application of Monte Carlo Simulation to seismic hazard assessment.
- Detailed procedure of the unique Monte Carlo Simulation.

ARTICLE INFO

Article history: Received 8 July 2014 Received in revised form 15 October 2014 Available online 13 January 2015

Keywords: Seismic hazard The Poisson distribution Monte Carlo Simulation

ABSTRACT

Unlike earthquake frequency that was proved following the Poisson distribution, seismic hazard (the annual rate of earthquake ground motions) is assumed to be the same type of random variables without tangible support. Instead of using total-probability algorithms currently employed, this study applied Monte Carlo Simulation (MCS) to obtain the probability function of seismic hazard, and then compared it to the Poisson distribution to see if it is really close to the model prediction as assumed. On the basis of a benchmark calculation, the analysis shows a very good agreement between the two, providing some evidence for the first time that seismic hazard should follow the Poisson distribution, although the relationship has been commonly employed in earthquake studies.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

Given that earthquakes are not predictable [1,2], seismic hazard analysis has become one of the practical approaches for earthquake hazard mitigation [1,2]. But before introducing probabilistic seismic hazard analysis (PSHA), it is worth clarifying the definition of seismic hazard: rather than casualty or economic loss associated with earthquakes, seismic hazard refers to the annual rate of a given ground motion of exceedance, e.g., $PGA > 0.1 \text{ g} = 0.01 \text{ per year.}^{1}$

With the mean hazard rate calculated, the next step of PSHA is to estimate the probability for the seismic hazard (e.g., PGA > 0.1 g) to occur in a given period of time (e.g., 50 years), by assuming it is a random variable following the Poisson distribution [3]. For example, if the mean rate (λ) of PGA > 0.1 g is estimated at 0.005 per year, the probability for the event to occur in next 50 years is equal to 22%:

$$Pr(T \le 50 \text{ years}; \lambda = 0.005 \text{ per year}) = 1 - e^{-\lambda t} = 1 - e^{-0.005 \times 50} = 0.22$$
 (1)

where $Pr(T \le t; \lambda) = 1 - e^{-\lambda t}$ is the cumulative density function of the exponential distribution that can satisfactorily model the event's temporal probability function, as the stochastic process is governed by the Poisson model [3,4].

E-mail addresses: jpwang@ust.hk (J.P. Wang), suchin@hku.hk (S.-C. Chang).

^{*} Corresponding author.

¹ PGA = peak ground acceleration.

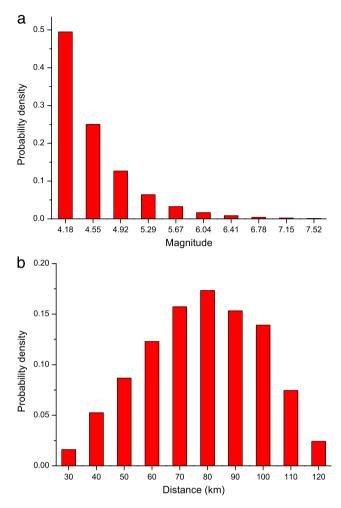


Fig. 1. The probability functions of earthquake magnitude and source-to-site distance of a benchmark example in the literature [3].

However, without any tangible support from the literature, the "seismic-hazard-and-Poisson" assumption has been an engineering judgment at best. As a result, the key scope of this study is to examine if seismic hazard should follow the Poisson distribution as commonly adopted in some earthquake studies.

Monte Carlo simulation (MCS) is one of the common approaches used in a variety of probabilistic analyses [5–11]. But unlike total-probability and FOSM² algorithms, Monte Carlo Simulation allows us to obtain a variable's probability function in addition to its mean value and standard deviation. Therefore, instead of using total-probability algorithms that are currently employed, this study applied a novel MCS calculation to probabilistic seismic hazard assessments, aiming to obtain the hazard's probability function, and to see if it is really following the Poisson distribution as assumed.

The paper is organized as follows: an overview of probabilistic seismic hazard analysis, followed by the introduction to the novel MCS application to this study. Next, the MCS calculations were demonstrated with a benchmark example from the literature, as well as the statistical tests examining the (Poisson) model's goodness-of-fit to the probability distribution of seismic hazard.

2. Probabilistic seismic hazard analysis (PSHA)

The framework of probabilistic seismic hazard analysis was first proposed in the late 1960s [12], and in the past decade it has become a common approach for developing site-specific earthquake-resistant designs [13–19]. Different from deterministic assessments, PSHA takes the uncertainties of earthquake magnitude, location, and motion attenuation into account. For instance, Fig. 1 shows the uncertainties or probability functions of earthquake magnitude and source-to-site distance for a benchmark example Ref. [3].

² FOSM: first-order second-moment.

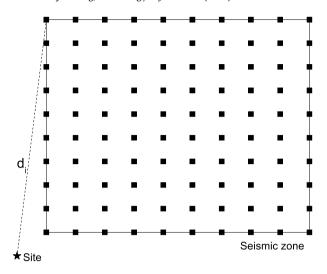


Fig. 2. The schematic diagram showing how to develop source-to-site distance distributions by "counting" a number of "point-to-site" distances.

It must be noted that the source-to-site distance distribution was developed based on the presumption that earthquake would occur at any location with equal probability [3]. We made a schematic diagram in Fig. 2 to further demonstrate the principle and the function development: by discretizing a seismic zone into many "uniformly distributed" points, a series of "point-to-site" distances can be calculated at first. Then the source-to-site distance distribution can be developed by "counting" those (point-to-site) distances. (For example, if 10 and 100 point-to-site distances out to 1000 are in ranges of 10–20 and 100–110 km, respectively, the probabilities for the two distance ranges are 0.01 and 0.1.) Most importantly, this is the reason why the source-to-site probability function is not a (perfectly) uniform distribution, although it is developed with the "uniform" presumption that earthquake can occur at any location with an equal probability.

The governing equation or performance function of PSHA is a ground motion prediction equation, which is a regression model describing the correlation between earthquake ground motion and (usually) magnitude and distance combined, like the following model used in the benchmark example Ref. [20]:

$$\ln PGA = 6.74 + 0.859M - 1.8 \ln(D + 25) + \varepsilon; \ \sigma_{\varepsilon} = 0.57$$
 (2)

where PGA is in gal, source-to-site distance D is in km, and M denotes earthquake magnitude. (Because the example was used for demonstrating PSHA algorithms, the unit of earthquake magnitude was not specified in the Ref. [3].) Besides, the term ε is the error of the model, which follows the normal distribution based on the fundamental of regression analysis [4]. To be more specific, the error of the regression model (i.e., Eq. (2)) follows the normal distribution with mean = 0 and standard deviation = 0.57.

Next, by applying total-probability computation to the performance function, the probability of ground motion of exceedance can be calculated as follows [3]:

$$\Pr(PGA > y^*) = \sum_{i=1}^{n} \sum_{j=1}^{k} \Pr(PGA > y^* | M = m_i, D = d_j) \times \Pr(M = m_i) \times \Pr(D = d_j)$$
(3)

where n and k denote the number of data bins in magnitude and distance probability functions, responsibility (see Fig. 1). Then with earthquake rate = v per year, the annual rate for PGA > y* can be calculated as follows [3]:

$$\lambda(PGA > y^*) = v \times \Pr(PGA > y^*). \tag{4}$$

As mentioned previously, by assuming that seismic hazard follows the Poisson distribution, the probability of seismic hazard (i.e., PGA $> y^*$) occurring in t years can be calculated with the exponential distribution [3]:

$$Pr(PGA > y^* \text{ in } t \text{ years}) = 1 - e^{-\lambda t}$$
(5)

where λ denotes the mean rate of PGA > y^* estimated with Eqs. (3) and (4).

With the given probability functions (i.e., Fig. 1), the ground motion model (i.e., Eq. (2)), and earthquake rate = 1.99 per year, Fig. 3 shows the annual rates for PGA > 0.01, 0.05, and 0.1 g for this benchmark example, equal to 1.022, 0.079, and 0.017 per year, respectively [3]. Note that the mean earthquake rate of 1.99 per year is one of the input data of the benchmark example; to be more specific, it is a function of four other input data (i.e., "a-value" = 3.5, "b-value" = 0.8, magnitude threshold = 4.0, and maximum magnitude = 7.7), and its calculation is detailed in the Appendix.

The reason of using the benchmark calculation as a demonstration to this study is that the data of the benchmark example, from inputs to outputs, are available and complete, which makes our calculations more objective and traceable. Similarly, the benchmark example with complete data available was used in a study for verifying new PSHA computer codes, for the same reason in order to enhance the study's objectivity and traceability [19].

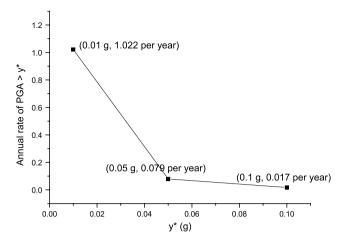


Fig. 3. The three hazard estimates for the benchmark example calculated with total-probability algorithms.

3. Seismic hazard analysis with Monte Carlo Simulation

This section will show how to apply Monte Carlo Simulation to solve the same probabilistic analysis to calculate the annual rate of earthquake ground motions. More importantly, we will show the hazard's probability function from the MCS analysis, and compare it to the theoretical Poisson distribution.

3.1. Data randomizations

Like any other MCS study, the first step to conduct the following analyses is to randomize the four earthquake parameters (i.e., magnitude, distance, frequency, and model error) based on their probability distribution given. In addition to the magnitude and distance functions shown in Fig. 1, the probability distributions of model error and earthquake frequency are shown in Fig. 4. As mentioned previously, the model error following the normal distribution is governed by the fundamentals of a regression model, and earthquake frequency following the Poisson distribution is a common presumption in earthquake studies [21,22].

For the three discrete distributions (magnitude, distance, and frequency) shown in Figs. 1 and 4(b), the randomizations of such a discrete variable X were assisted with the in-house algorithm as follows:

$$X = \begin{cases} x_1; & 0 < U \le P_1 \\ x_i; & P_{i-1} < U \le P_i \end{cases}$$
 (6)

where U is a uniformly distributed random number from 0 to 1, and P denotes cumulative probability in a given cumulative density function. To better demonstrate the algorithm, we made a schematic diagram in Fig. 5. Based on the algorithm and the given probability functions (Fig. 5), the randomly generated x is equal to 10 in corresponding to U = 0.05 (between 0 and 0.1), or x is equal to 30 as U = 0.6 (between 0.5 and 0.7), and so on.

As for the continuous distribution (model error) shown in Fig. 4(a), the so-called inversed algorithm was followed in this study to generate a random model error ε_R for the MCS [23]:

$$\varepsilon_R = f^{-1}(U) \tag{7}$$

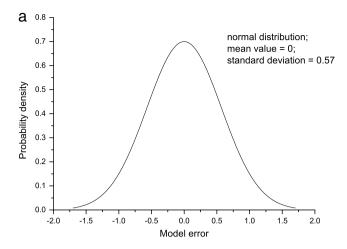
where f^{-1} is the inversed normal distribution; to be more specific, it is with mean value = 0 and standard deviation = 0.57 based on the input data (i.e., Eq. (2)) of the benchmark example.

From Eqs. (6) and (7), it is understood that the randomizations are depending on a random number from 0 to 1 (or *U*). Therefore, such a random number generator is needed for this study. Specifically, this study used function *RAND* in Excel for the data randomizations.

In summary, with the four given probability distributions (Figs. 1 and 4), the two randomization algorithms (Eqs. (6) and (7)), and the random number generator in Excel (function *RAND*), we can randomly generate earthquake magnitudes, distances, frequencies, and model errors in the first place, and then substitute them into the performance function (Eq. (2)) to obtain a random ground motion for seismic hazard assessments. More details regarding the MCS calculation are given in the following.

3.2. The calculations

The flowchart shown in Fig. 6 summarizes the MCS procedure for seismic hazard calculations. As mentioned previously, after the randomizations of the four parameters, we could obtain random PGA motions (e.g., 0.05, 0.12, and 0.2 g) and calculate the rate of seismic hazard (e.g., PGA > 0.1 g) from those randomizations. Next, by repeating the calculations for



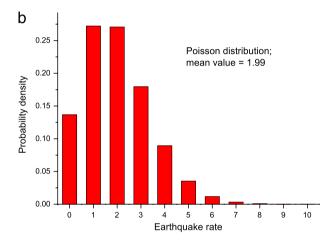


Fig. 4. The probability functions of two other parameters of the benchmark probabilistic analysis: (a) the distribution of the model error of the ground motion prediction equation, and (b) the distribution of earthquake frequency with mean rate = 1.99 per year.

a number of trials, we could obtain many seismic hazard samples, and then calculate the mean value as the final hazard estimates with the MCS.

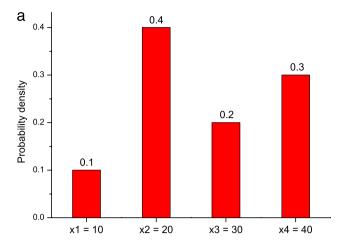
For better explaining the calculation, Table 1 summarizes the data in two trials, including the four random earthquake parameters, and those associated U values. In the first trial, given random earthquake rate = 3 per year, three sets of random magnitudes, distances, and model errors were generated in the first place, and then substituted into the ground motion model to obtain three random motions. Accordingly, given that one of them is greater than 0.01 g, the rate of PGA > 0.01 g is equal to 1 per year from this trial. By contrast, since there is no ground motion greater than 0.01 g in the second trial, the hazard rate is equal to zero from this simulation.

3.3. Sample size

A large sample size is the key to the reliability of Monte Carlo Simulation, and the larger the size, the more reliable the result will be (e.g., Refs. [7,23]). Although some suggestions are available for calculating the thresholds of sample size depending on the number of variables appearing in the performance function [23], a more straightforward approach is to adopt a very large sample size (e.g., 10,000) to ensure that the MCS is accountable [7]. As a result, this study simply used 10,000 as the sample size of the MCS, ensuring that the MCS is accountable with such a sample size much larger than the recommendations [23].

4. Evidence: seismic hazard following the Poisson distribution

This section shows the results of the novel MCS calculations for the benchmark seismic hazard assessments and illustrates their comparison to those from total-probability algorithms currently employed. As expected, the hazard estimates from



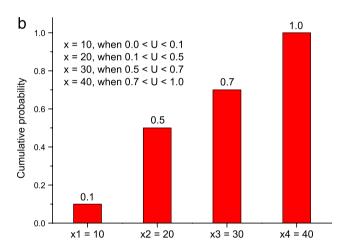


Fig. 5. A demonstration to the in-house algorithm for randomizing a discrete parameter based on its probability functions, where *U* is a uniformly distributed random value from 0 to 1.

Table 1Summary of the first two trials of the Monte Carlo Simulation to estimate the seismic hazard in PGA > 0.01 g; the numbers in the parentheses are random *U* values (from 0 to 1) generated with Excel.

	Random earthquake rate	Random magnitude	Random distance (km)	Random model error	Random PGA motion (g)	Random rate of PGA > 0.01 per year
Trial 1	3 (0.779)	4.56 (0.501) 4.19 (0.117) 4.19 (0.157)	70 (0.372) 70 (0.365) 100 (0.809)	0.61 (0.859) -1.27 (0.013) -0.23 (0.341)	0.022 0.002 0.004	1
Trial 2	2 (0.446)	4.19 (0.461) 4.56 (0.646)	90 (0.671) 70 (0.280)	-1.27 (0.012) -0.42 (0.228)	0.002 0.008	0

both algorithms are close to each other, supporting that the MCS calculation is robust. More importantly, the MCS can provide the tangible evidence about the hazard's probability function, the key scope and objective of the study.

4.1. Robustness of the MCS

With a large sample size of 10,000, Fig. 7 shows the scatter plot for the 10,000 seismic hazard samples in PGA > 0.01 g. The result shows that most trials return a hazard rate in less than three events per year. More importantly, the mean rate of the 10,000 samples is equal to 1.022 per year, which is almost identical to the solution from total-probability calculation (see Fig. 3), attesting the robustness of our randomization procedures and the MCS calculation, and supporting the common proposition that a probabilistic analysis can be solved with different algorithms.

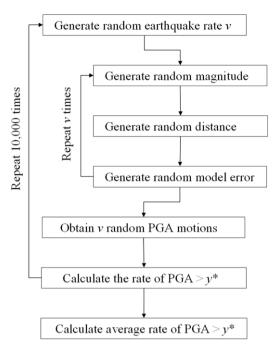


Fig. 6. The flowchart summarizing the MCS procedure for probabilistic seismic hazard analysis.

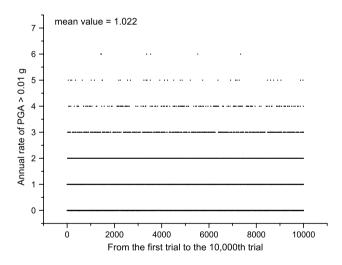


Fig. 7. The 10,000 seismic hazard samples in PGA > 0.01~g from the Monte Carlo Simulation.

4.2. The probability distribution of seismic hazard

Next, by "counting" the 10,000 data in Fig. 7, we developed the hazard's probability function as shown in Fig. 8. In addition, given the mean rate (i.e., the model parameter of the Poisson distribution) calculated as 1.022, we also show the theoretical Poisson distribution in Fig. 8 for the comparison. Excitingly, the two are in a very good agreement from the analyses.

As many other studies [21,22], we further performed Chi-square tests to examine the model's goodness-of-fit in a more quantitative manner. (Details of the statistical test are given in the Appendix.) Accordingly, the Chi-square value (χ^2) of the test data is equal to 6.95, lower than the critical value of 11.07 given a 5% level of significance employed by the test. Therefore, with the tangible, quantitative support from MCS to statistical analysis, this study offers some evidence for the first time that seismic hazard should follow the Poisson distribution as the relationship has been commonly used in earthquake studies.

4.3. Two more verifications

We carried out two additional analyses for two other levels of seismic hazard (i.e., PGA > 0.05 and 0.1 g) to further attest the conclusion from the previous case. With the same procedure, Fig. 9 shows the histograms of the 10,000 data from MCS,

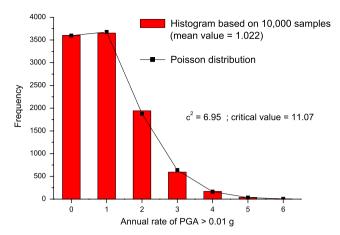


Fig. 8. The histogram of the 10,000 MCS data in terms of PGA > 0.01 g, and the theoretical Poisson distribution based on its mean value.

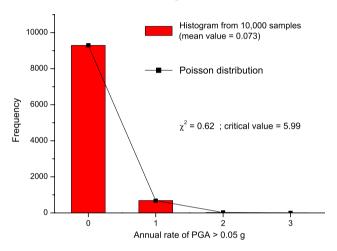


Fig. 9. The histogram of the 10,000 MCS data in terms of PGA > 0.05 g, and the theoretical Poisson distribution based on its mean value.

indicating that the mean rate for PGA > 0.05 g is equal to 0.075 per year, also very close to the total-probability calculation (see Fig. 3). More importantly, the statistical test also supports that this larger seismic hazard should follow the Poisson distribution as the test data shown in Fig. 9.

Fig. 10 shows the probability distribution for an even larger seismic hazard in PGA > 0.1 g. The MCS shows that the hazard estimate is about 0.016 per year, also very close to 0.017 from the total-probability calculation (see Fig. 3). In addition, as the comparison and statistical test data shown in Fig. 10, this seismic hazard in PGA > 0.1 g should be following the Poisson distribution as well, with the support from MCS to statistical analysis.

5. Discussion: earthquake, seismic hazard, and the Poisson distribution

As mentioned previously, earthquake frequency following the Poisson distribution is commonly used in earthquake analyses, including in this study. Nevertheless, it must be noted that the "earthquake-and-Poisson" relationship was supported by earthquake observations in California or in Taiwan [21,22], making it more of a fact than an assumption.

By contrast, seismic hazard, which is a function of earthquake magnitude, distance, frequency, and ground motion models, is a different random variable than earthquake. But without any tangible support in the literature, the "seismic-hazard-and-Poisson" assumption has been an engineering judgment at best, until this study offering some evidence for the first time from the novel application of Monte Carlo Simulation.

6. Summary

Seismic hazard is usually considered a variable following the Poisson distribution in earthquake studies. But without any support in the literature, the key scope of this study is to evaluate this engineering assumption. The novelty of the study is to apply Monte Carlo Simulation to solve the probabilistic analysis, instead of using total-probability algorithms currently employed.

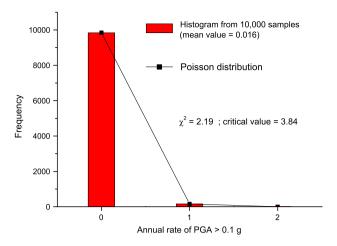


Fig. 10. The histogram of the 10,000 MCS data in terms of PGA > 0.1 g, and the theoretical Poisson distribution based on its mean value.

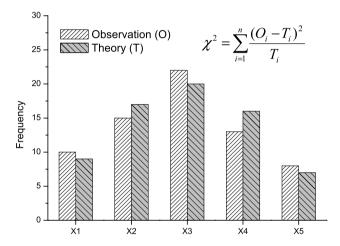


Fig. A.1. The schematic diagram explaining the essentials of the Chi-square test.

On the basis of a benchmark calculation in the literature, this study shows the hazard estimates from MCS and total-probability algorithms are almost identical to each other, attesting the robustness of our randomization procedures and the new MCS application to earthquake study.

Most importantly, with a number of seismic hazard samples from Monte Carlo Simulation, the statistical analyses attested its probability function should follow the Poisson distribution, providing tangible evidence for the first time in support of this engineering presumption that has been commonly used in earthquake analyses.

Acknowledgments

We appreciate the valuable comments from the Editor and reviewers, making the submission and the study much improved in so many aspects.

Appendix A. Calculation of the annual earthquake rate

Earthquake annual rate λ that is an input parameter of PSHA is usually calculated with the Gutenberg–Richter recurrence law along with four other input parameters [3]:

$$\lambda = 10^{a-b \times m_0} - 10^{a-b \times m_{\text{MAX}}} \tag{A.1}$$

where a and b are the model parameters of the Gutenberg–Richter relationship, m_0 is the magnitude threshold, and m_{MAX} is the maximum magnitude. For the benchmark example, the data are as follows: a = 3.5, b = 0.8, $m_0 = 4.0$, and $m_{MAX} = 7.7$. By substituting them into Eq. (A.1), one can calculate the earthquake rate as 1.99 per year as it is given in the literature [3].

Appendix B. Overview of the Chi-square test

As a schematic diagram shown in Fig. A.1, the Chi-square test is to compute the Chi-square value quantifying the difference between theoretical and observational frequencies, and then comparing the value to the critical value at a given level of significance. When the Chi-square value of test data is less than the critical value, the observation is considered following the model suggested; otherwise the hypothesis should be rejected [4].

References

- [1] R.J. Geller, D.D. Jackson, Y.Y. Kagan, F. Mulargia, Earthquakes cannot be predicted, Science 275 (1997) 1616.
- [2] J.P. Wang, L. Brant, Y.M. Wu, H. Taheri, Probability-based PGA estimations using the double-lognormal distribution: including site-specific seismic hazard analysis for four sites in Taiwan, Soil Dyn. Earthq. Eng. 42 (2012) 177–183.
- [3] S.L. Kramer, Geotechnical Earthquake Engineering, Prentice Hall, New Jersey, 1996, pp. 131-134.
- [4] A. Ang, W.H. Tang, Probability Concepts in Engineering: Emphasis on Applications to Civil and Environmental Engineering, second ed., John Wally & Sons Inc., New Jersey, 2007.
- [5] C. Xuan, J. Channell, Testing the relationship between timing of geomagnetic reversals/excursions and phase of orbital cycles using circular statistics and Monte Carlo simulations, Earth Planet. Sci. Lett. 268 (2008) 245–254.
- [6] G. Weatherill, P.W. Burton, An alternative approach to probabilistic seismic hazard analysis in the Aegean region using Monte Carlo simulation, Tectonophysics 492 (2010) 253–278.
- [7] J.P. Wang, C.W. Lin, H. Taheri, W.S. Chen, Impact of fault parameter uncertainties on earthquake recurrence probability by Monte Carlo simulation—an example in central Taiwan, Eng. Geol. 126 (2012) 67–74.
- [8] J. Zhang, W.H. Tang, L.M. Zhang, H.W. Huang, Characterising geotechnical model uncertainty by hybrid Markov Chain Monte Carlo simulation, Comput. Geotech. 43 (2012) 26–36.
- [9] F. Cadini, J. De Sanctis, I. Bertoli, E. Zio, Monte Carlo simulation of radionuclide migration in fractured rock for the performance assessment of radioactive waste repositories, Reliab. Eng. Syst. Saf. 111 (2013) 241–247.
- [10] Y. Noh, K. Chang, Y. Seo, D. Chang, Risk-based determination of design pressure of LNG fuel storage tanks based on dynamic process simulation combined with Monte Carlo method, Reliab. Eng. Syst. Saf. 129 (2014) 76–82.
- [11] M.E. Kartel, H.B. Basaga, A. Bayraktar, Probabilistic nonlinear analysis of CFR dams by MCS using response surface method, Appl. Math. Model. 35 (2011) 2752–2770.
- [12] C.A. Cornell, Engineering seismic risk analysis, Bull. Seismol. Soc. Amer. 58 (1968) 1583–1606.
- [13] M.D. Petersen, J. Dewey, S. Hartzell, C. Mueller, S. Harmsen, A. Frankel, K. Rukstales, Probabilistic seismic hazard analysis for Sumatra, Indonesia and across the Southern Malaysian Peninsula, Tectonophysics 390 (2004) 141–158.
- [14] C.T. Cheng, S.J. Chiou, C.T. Lee, Y.B. Tsai, Study of probabilistic seismic hazard maps of Taiwan after Chi-Chi earthquake, J. GeoEngin. 2 (2007) 19–28.
- [15] M. Beroya, A. Aydin, Seismic hazard analysis of Laoag city, Northern Philippines for liquefaction potential assessment, Eng. Geol. 96 (2008) 28–42.
- [16] S.W. Han, Y.S. Choi, Seismic hazard analysis in low and moderate seismic region-Korean peninsula, Struct. Saf. 30 (2008) 543–558.
- [17] V.Y. Sokolov, F. Wenzel, R. Mohindra, Probabilistic seismic hazard assessment for Romania and sensitivity analysis: a case of joint consideration of intermediate-depth (Vrancea) and shallow (crustal) seismicity, Soil Dyn. Earthq. Eng. 29 (2009) 364–381.
- [18] A.Q. Bhatti, S. Hassan, Z. Rafi, Z. Khatoon, Q. Ali, Probabilistic seismic hazard analysis of Islamabad, Pakistan, J. Asian Earth Sci. 42 (2011) 468-478.
- [19] J.P. Wang, D. Huang, C.T. Cheng, K.S. Shao, Y.C. Wu, C.W. Chang, Seismic hazard analysis for Taipei city including deaggregation, design spectra, and time history with Excel applications, Comput. Geosci. 52 (2013) 146–154.
- [20] C.A. Cornell, H. Banan, A.F. Shakal, Seismic motion and response prediction alternatives, Earthq. Eng. Eng. Dyn. 7 (1979) 295-315.
- [21] J.K. Gardner, L. Knopoff, Is the sequence of earthquakes in southern California, with aftershocks removed, Poissonian? Bull. Seismol. Soc. Amer. 64 (1974) 1363–1367.
- [22] J.P. Wang, D. Huang, S.C. Chang, Y.M. Wu, New evidence and perspective to the Poisson process and earthquake temporal distribution from 55,000 events around Taiwan since 1900, Nat. Hazards Rev. ASCE 15 (2014) 38–47.
- [23] L.W. Abramson, S.L. Thomas, S. Sharma, G.M. Boyce, Slope Stability and Stabilization Methods, John Wally & Sons Inc., New Jersey, 2002.