17th World Conference on Earthquake Engineering, 17WCEE
Sendai, Japan - September 13th to 18th 2020

IMPLEMENTING MONTE-CARLO EARTHQUAKE CATALOGS IN CYBERSHAKE

S. Azar(1), M. Dabaghi(2)

(1) PhD Student, Dept. of Civil and Environmental Engineering, American University of Beirut, Lebanon, sna53@mail.aub.edu
(2) Assistant Professor, Dept. of Civil and Environmental Engineering, American University of Beirut, Lebanon, md81@aub.edu.lb

Abstract

Probabilistic Seismic Hazard Analysis (PSHA) is a commonly used method in evaluating the seismic hazard at an earthquake-prone site. It combines the uncertainties in the location, size and shaking levels of earthquakes to develop hazard curves at the site. While the uncertainties in the location and size of the seismic sources are accounted for by earthquake rupture forecast models, shaking levels are determined using ground motion models. Conventional PSHA employs ground motion prediction equations (GMPEs) as ground motion models. However, alternative methods such as numerical simulations can also be used in predicting the intensity of a ground motion for a given scenario. In this context, SCEC developed CyberShake, a research project that incorporates physics-based 3D ground motions simulations within seismic hazard calculations.

CyberShake starts with rupture definitions and their associated probabilities obtained from an earthquake rupture forecast. The ruptures are then converted into a discrete set of rupture variations characterized by different hypocenter locations and slip distributions. Based on a 3D velocity model, Green's functions are generated for all the faults of interest and a ground motion time series is simulated for each sampled rupture variation. This requires a considerable amount of computational time and resources due to the significant processing requirements imposed by source-based models on the one hand, and by the large number of rupture variations on the other hand.

The current paper proposes a more efficient PSHA framework that can accurately represent the seismic hazard at a site by only considering a subset of earthquake scenarios. The proposed framework is based on a Monte Carlo procedure that generates earthquake catalogs with a specified duration. In this case, ground motion time series are exclusively simulated for the scenarios selected in the earthquake catalog, and hazard calculations are limited to the intensity measures obtained from this subset of scenarios. The proposed framework is applied to a site located in Southern California. The entire set of simulated ground motions generated by CyberShake Study 15.12 that can affect the site is extracted. Hazard calculations are performed by considering a subset of these simulations based on earthquake catalogs with different lengths. To evaluate the accuracy of the method, the resulting hazard curves are then compared to the curves obtained by considering the entire set of simulated ground motions. The results show that the accuracy of the hazard curves developed based on Monte-Carlo catalogs depends on the length of the catalog and that an optimal catalog length can be selected based on the required level of accuracy. For an accuracy level larger than 95%, the 200,000-years catalog is selected as it can significantly reduce the number of earthquake scenarios and produce hazard curves with errors that are consistently smaller than 5%. The catalog is also able to reproduce the disaggregation patterns observed in the approach that considers the entire set of simulations. The obtained results demonstrate the potential of implementing Monte-Carlo catalogs in CyberShake simulation-based PSHA in which the entire set of simulations that can affect the site of interest is considered. Such an approach can reduce the computational time of the current CyberShake approach by reducing the required number of simulations while still maintaining acceptable levels of accuracy in the developed hazard curves.

Keywords: CyberShake; Monte-Carlo Simulations; Probabilistic Seismic Hazard Analysis; Synthetic Ground Motions

1. Introduction

The goal of seismic design is to ensure that a structure at a site can withstand seismic loads during an earthquake event with acceptable performance. The severity of the loads to consider in the design is a function of the intensity of the ground motion expected at the site. Thus, the analysis of earthquake-resistant structures requires an appropriate description of the level of shaking, generally determined from seismic hazard analyses. These analyses involve a quantitative estimation of the ground shaking level at the site and can be developed through deterministic or probabilistic approaches. While the first approach defines the hazard as the "worst-case" earthquake scenario, the second approach considers the uncertainties in the earthquake sources and their probabilities of occurrence, as well as the uncertainties in the ground shaking levels produced at the site. Although combining these uncertainties can add complexity to the hazard analyses, probabilistic seismic hazard analysis (PSHA) is commonly used because it provides ground motions intensities associated with their occurrence rates at the site [1-3].

Hazard assessment using PSHA requires identification and characterization of all possible earthquake ruptures that can produce significant shaking levels at the site. An earthquake rupture forecast (ERF) collects information about the different potential sources from geologic and tectonic evidence, fault activity, historical and instrumental seismicity. The ERF typically results in a set of relevant earthquake sources in the region of interest described by their spatial geometries and associated probabilities of occurrence. For instance, the 2007 Working Group on California Earthquake Probabilities (WGCEP, 2007) developed the Uniform California Earthquake Rupture Forecast Version 2 (UCERF 2), which contains time-dependent and time-independent models of the seismic sources located in the California region with magnitudes $M > 5$ [4]. The developed models describe the geometry of the sources and determine the rates at which earthquakes with specified magnitudes occur on these sources. In addition to source characterization, PSHA requires a ground motion model that can determine the severity of the shaking for the different scenarios. Ground motion prediction equations (GMPEs) are traditionally used in a PSHA framework. Such models predict the probability distribution of a ground motion intensity measure due to an earthquake with given source and site parameters. However, applying GMPEs in PSHA assumes ergodicity, i.e., that the variability of ground motions across different sites from past earthquakes represents the variability of ground motions at one site for future earthquakes. Anderson and Brune [5] showed that PSHA that makes the ergodic assumption can result in an overestimation of the hazard. On the other hand, GMPEs are regression models with predefined functional forms that were developed based on seismic data. Therefore, the accuracy of the fitted models depends on the availability of the data from historical earthquakes and is hence limited to scenarios with significant recorded evidence. For instance, a comparison of the NGA-West2 GMPEs [6] showed that the largest difference between the median ground motions predicted by the different models is observed in cases where the NGA-West2 database is sparse.

More recently, numerical simulation models were implemented in PSHA to predict the ground motion intensity at a site for a given earthquake scenario. Numerical simulations incorporate physics related to the fault rupture, wave propagation and site response in the simulated time series. Consequently, these alternative ground motion models are being increasingly utilized as they offer benefits over empirical models. Such models allow the simulation of ground motions for all arbitrary conditions even for the less frequent and unobserved cases. Additionally, site-specific responses such as directivity and sedimentary basin effects are naturally included in the model using the 3D velocity model. Researchers working through the Southern California Earthquake Center's (SCEC) Community Modeling Environment developed a methodology for incorporating numerically simulated ground motion predictions in PSHA [7]. In this platform, called CyberShake, the ground motion time series are simulated using seismic reciprocity for an ensemble of rupture variations sampled from UCERF 2.0. The latter provides rupture descriptions that are then augmented with slip distributions and hypocenter locations for a complete description of the kinematic source. Assigning multiple hypocenter locations and slip models to the same rupture considers the expected variability in the source parameters and alleviates the need for the ergodic assumption. Strain Green Tensors (SGTs) are then calculated around the site of interest and post-processed using seismic reciprocity to obtain the seismograms. Intensity measures of the simulated ground motions at a site are eventually calculated and combined with the source probabilities of occurrence to determine the hazard at the site. Validation of simulated ground motions showed that CyberShake models are viable and can capture basin, directivity and basin-directivity coupling effects [7].

However, implementing the CyberShake approach in PSHA requires significant computational time and resources. The significant processing requirements are imposed by the large number of rupture variations needed for the development of hazard curves. Furthermore, the maximum frequency in the simulated time series depends on the spatial resolutions of the source and velocity model, which largely influence the computational demands. In this context, CyberShake Study 15.12 combines the deterministic low-frequency ground motion time series from the physics-based model with high-frequency stochastic seismograms generated using the SCEC Broadband Platform [8] at a transition frequency of 1 Hz. Additional computational resources are required to extend the physics-based transition frequency to higher values. The SCEC research group is currently working on extending the deterministic calculations to shorter periods by providing larger computational resources. The researchers are also working on altering the current CyberShake workflow in a way that reduces its computational demands. These modifications include running the SGTs on parallel nodes and simulating the seismograms without the need to save the calculated SGTs.

Although much research has been devoted to reducing the time-to-solution per site by adapting the CyberShake workflow, less attention has been paid to restrict the number of ground motions that need to be simulated per site. The purpose of this paper is to present a PSHA framework that reduces the number of earthquake scenarios to consider for the development of hazard curves at a site. When implemented in the CyberShake platform, the number of physics-based ground motions to be simulated is reduced, therefore requiring less computational resources. The method is based on the simulation of earthquake Monte-Carlo catalogs with a specified duration. Ground motions are then simulated only for the scenarios included in the catalog. The corresponding intensity measures are calculated and then used for the development of hazard curves. To validate the proposed approach, Monte-Carlo catalogs with different lengths are developed for a site located in downtown Los Angeles, California. The resulting hazard curves are compared with curves from CyberShake Study 15.12 for the same site. The accuracy of the curves is evaluated based on the relative error at the commonly used hazard levels between the rates obtained from the proposed method and the rates obtained by considering all the simulations. The latter represents the method that is currently used by CyberShake for developing hazard curves. An optimal catalog length is then selected based on the variation of the accuracy of the hazard curves with the simulated catalog length. Additionally, disaggregation patterns from the selected Monte-Carlo catalog length are compared to the patterns obtained by considering the entire set of simulations. The relevance in the proposed approach is that the limited number of scenarios to consider can shorten the time-to-solution per site currently required by CyberShake to develop hazard maps. Alternatively, the implementation of Monte-Carlo catalogs in CyberShake simulation-based PSHA can make the platform more practical which remains a crucial need among seismologists.

2. Methodology

This paper presents a PSHA framework that involves numerical simulations instead of employing the conventional GMPEs as ground motion model. The framework also limits the number of ground motion time series that need to be simulated. In the remainder of the paper, we refer to the proposed approach as Monte-Carlo catalog method and to the typical approach that considers the entire set of simulations as the full-fledged method. The latter is a representation of the method that is currently used in CyberShake as described in [7] and is hence considered as a basis for comparison. In the full-fledged method [7], the rates of exceedance needed to develop the hazard curves at a site are obtained by considering the intensity levels from all the scenarios that affect the site. Conversely, in the Monte-Carlo catalog method, the exceedance rates are determined by only considering the subset of the earthquake scenarios sampled by the catalogs generated using Monte-Carlo simulation.

2.1. Seismic Source Model

The first step in a PSHA framework is to define a seismic source model that identifies the geometry, magnitude and probability of occurrence of the possible earthquake ruptures affecting a given site. Study 15.12 of CyberShake uses the time-independent model of UCERF 2.0 as seismic source model. This model provides the probability of occurrence of discretized fault ruptures that can generate earthquakes with $M_w > 5$ and describes their geometries with a 200-m rupture surface resolution. CyberShake limits the ruptures to $M_w > 6$, ignores background seismicity and only considers ruptures that are within 200 km of the site. For a complete description of the kinematic ruptures, CyberShake transforms the ruptures from UCERF 2.0 into rupture variations by differing slip distributions and hypocenter locations for the same rupture geometry. This process results in an average of 415,000 scenarios to be considered at each site.

2.2. Earthquake Catalog

As mentioned before, the proposed approach selects a subset of earthquake scenarios to develop hazard curves at the site. For a given site, the I relevant ruptures, their associated annual probabilities $P_i, i = 1 \dots I$ and the number of rupture variations k_i are used to generate earthquake catalogs over a period of Y years. The number of rupture variations k_i represents the number of hypocenter locations and slip distributions assigned to rupture i. The simulation of the earthquake catalog is based on the Monte-Carlo technique and is similar to the methodology presented in Azar, Dabaghi, and Rezaeian [9]. First, assuming that the number of earthquakes from rupture i follows a Poisson distribution, the mean annual rate of occurrence λ_i of each relevant rupture i can be determined:

$$\lambda_i = -\log(1 - P_i) \tag{1}$$

The next step is to determine N_i, the number of times a relevant rupture i is expected to occur during the duration Y of the catalog. The number N_i is determined by simulating a random number from a Poisson distribution with a mean rate equal to $\lambda_i \times Y$. Each of these N_i occurrences of rupture i can have a different hypocenter location and slip distribution i.e. a different rupture variation. To determine the rupture variation to be considered for each of these occurrences $n_i = 1 \dots N_i$, a random number is drawn from a uniform distribution ranging from 1 to k_i. Sampling from a uniform distribution assumes equiprobability of rupture variations for the same rupture, consistent with CyberShake assumption. After determining the N_i different earthquake scenarios caused by the i^{th} rupture, the same procedure is applied to the I relevant ruptures. The result is a simulated catalog that contains an ensemble of $\sum_i N_i$ earthquake scenarios, which are considered as one possible realization of the seismic events that may affect the site over a duration of Y years. If the duration Y of the catalog is longer, a larger number of seismic events are expected to occur, therefore, the number of earthquake scenarios $\sum_i N_i$ in a simulated catalog increases with Y.

2.3. Simulated Ground Motions and Hazard Curves

After developing an earthquake catalog with a duration of Y years, ground motion time series are simulated at the site of interest for each earthquake scenario in the catalog. In this study, the CyberShake physics-based simulation model is employed. The result is a set of ground motion time series that may affect the site in Y years. In this paper, the hazard curves are developed and compared for the $S_{a\,RotD50}(T)$ intensity measure suggested by Boore [10]. Therefore, the $S_{a\,RotD50}(T)$ of the ground motion time series are calculated at specified periods of interest. To develop the hazard curves, the mean annual rate of exceedance λ_α for an intensity threshold α is calculated by counting the number of ground motions for which $S_{a\,RotD50}(T) > \alpha$ and dividing the total number by the catalog duration Y. The hazard curves are obtained by plotting the mean annual rate of exceedance of different $S_{a\,RotD50}(T)$ thresholds.

2.4. Developing different catalogs lengths

The goal of this paper is to evaluate the effect of the length Y of the Monte-Carlo catalog on the accuracy of the resulting hazard curves. The procedure described above refers to the simulation of a single catalog with a duration of Y years and the development of the corresponding hazard curves. For the purpose of this paper, multiple earthquake catalogs are simulated for various values of Y, and the resulting hazard curves are compared with the outputs from the full-fledged method. Since both the full-fledged and the Monte-Carlo catalog methods start with the same seismic source model and use the same ground motion model, comparing the hazard curves from these two methods allows evaluating the effect of catalog length on the computed rates and on the accuracy of the proposed approach.

3. Case Study: Los Angeles, California

The Monte-Carlo catalog method is applied to a site located in downtown Los Angeles (LADT) at a latitude of 34.05204, a longitude of -118.25713 and with a V_{s30} of 390 m/s. An ensemble of 7,019 seismic ruptures located within 200 km from LADT are extracted from the CyberShake database with their associated probabilities of occurrence. Similarly, the full set of 476,920 corresponding earthquake scenarios are identified and extracted from the database.

Earthquake catalogs are developed using the ruptures, their probabilities of occurrence and their corresponding earthquake scenarios based on the methodology described in section 2.2. As a first step, 10 Monte-Carlo catalogs having a duration of 10,000 years are generated. Each simulated catalog is a representation of the earthquake scenarios that could affect LADT over a 10,000 years period. Multiple catalogs are simulated for the same time period to account for the variability in Monte-Carlo simulations. The 10 earthquake catalogs of 10,000 years contain an average of 644 ruptures and 1,048 earthquake scenarios. In the next step, synthetic ground motion time series should be simulated to represent each earthquake scenario in the 10 catalogs. However, for the purpose of this study, members from the SCEC research group had provided the numerical simulations for the 476,920 earthquake scenarios at LADT that were used in CyberShake Study 15.12. Therefore, the synthetic ground motions for the different scenarios in the simulated catalogs are selected (instead of being simulated) from the collection of all the synthetic ground motions generated for LADT. The $S_{a\,RotD50}(T)$ values of the selected synthetic ground motions are then calculated at different periods for a fixed viscous damping ratio of 5%. Hazard curves that correspond to each simulated catalog are developed as described in Section 2.3.

To evaluate the effect of the duration Y of the catalog, additional catalogs are simulated using the same method applied to the 10,000 years catalog. As a result, 10 earthquake catalogs are generated for each of the following durations: 10,000; 20,000; 50,000; 100,000; 200,000; 500,000 and 1,000,000 years.

3.1. Number of earthquake scenarios

The mean number of ruptures and earthquake scenarios across the 10 catalogs for the different catalog lengths are summarized in Table 1. As expected, the number of earthquake scenarios in the Monte-Carlo catalogs increases with the length of the catalog since more earthquake events are expected to occur over a longer time period. As mentioned before, the source model identifies 7,019 significant ruptures that are located within 200 km of LADT and CyberShake defines 476,920 earthquake scenarios accordingly. In the full-fledged method, the complete set of ruptures and earthquake scenarios is considered which requires the simulation of a ground motion for each earthquake scenario. Table 1 shows that by using the Monte-Carlo catalog method, the number of ground motions that need to be simulated decreases, thus reducing the computational time. The reduction in the number of ruptures depends on the length of the catalogs and varies from 19% for the 1,000,000-years catalogs to 91% for the 10,000-years catalogs. Similarly, the number of earthquake scenarios required in the 1,000,000-years catalogs is 90% less than what is required by the full-fledged method and is almost 100% less for the 10,000-years catalogs.

Table 1 – Mean number of ruptures and scenarios in the simulated catalogs and the percent reduction compared to the full-fledged method

Catalog Duration (years)		10,000	20,000	50,000	100,000	200,000	500,000	1,000,000
Ruptures	Mean number	644	1025	1773	2464	3379	4708	5699
	Reduction (%)	91	85	75	65	52	33	19
Earthquake Scenarios	Mean number	1048	2078	5161	10359	20773	52032	103804
	Reduction (%)	100	100	99	98	97	94	90

Graves et al. [7] explains that the large number of numerical simulations to be performed by CyberShake and that is illustrated in the paper by the full-fledged method, requires significant computational time and resources. By choosing a shorter catalog (i.e. reducing the number of earthquake scenarios for which synthetic ground motions need to be simulated), the time-to-solution per site required by CyberShake is shortened. However, the choice of the optimal length of the catalog also depends on the desired accuracy of the resulting hazard curves.

3.2. Hazard Curves

$S_{a\,RotD50}(T)$ hazard curves corresponding to the different simulated catalogs are computed at periods $T = 0.1, 0.2, 0.5, 1, 2$ and $5s$. Fig. 1 shows the $S_{a\,RotD50}$ hazard curves at LADT for a period $T = 2s$ computed using the proposed framework. Each plot in the figure shows the 10 hazard curves computed from catalogs having the same length and specifies the most commonly used hazard levels that correspond to a 2%, 5% and 10% probability of exceedance in 50 years. The computed hazard curves are compared with results from the full-fledged method at the same site and period. The figure shows that for $Y = 10{,}000$ years, the hazard curves from the different catalogs are variable especially at high hazard levels. The number of scenarios in a 10,000-years earthquake catalog does not seem to be sufficient to represent the hazard at different levels. Moreover, as the $S_{a\,RotD50}$ threshold increases, the number of ground motions with high intensity measures that exceed the threshold decreases resulting in less accurate rates at low hazard levels. However, as the length of the catalog increases, the variability in the hazard curves from the different Monte-Carlo catalogs is less pronounced. Similarly, the computed hazard curves are closer to the full-fledged method hazard curves for a catalog with a longer duration even at low hazard levels. Fig. 1 also shows that hazard curves computed from catalogs with duration of 200,000 years or longer are good representations of the full-fledged method hazard curves, in which all the scenarios are considered.

To quantify the accuracy of the computed hazard curves, relative errors at the frequently used hazard levels are calculated. These values estimate the difference between using Monte-Carlo catalogs and considering all the scenarios in a simulation-based PSHA. The errors are calculated for each computed hazard curve at the different periods. Fig. 2 represents the boxplots of these errors averaged across the different periods and over all sets of catalogs for a single catalog duration. The errors in the figure correspond to the 5% probability of exceedance in 50 years hazard level. The figure clearly shows that the median error values for the different catalog lengths are close to zero. This implies that the results from the Monte-Carlo catalog method are unbiased and that the proposed approach is functional in a PSHA framework. On the other hand, Fig. 2 highlights the significant effect of the catalog length on the variability of the resulting hazard curves. While the relative error can reach 17% with a 10,000-years earthquake catalog, it can be reduced to less than 2% when using a 1,000,000-years catalog.

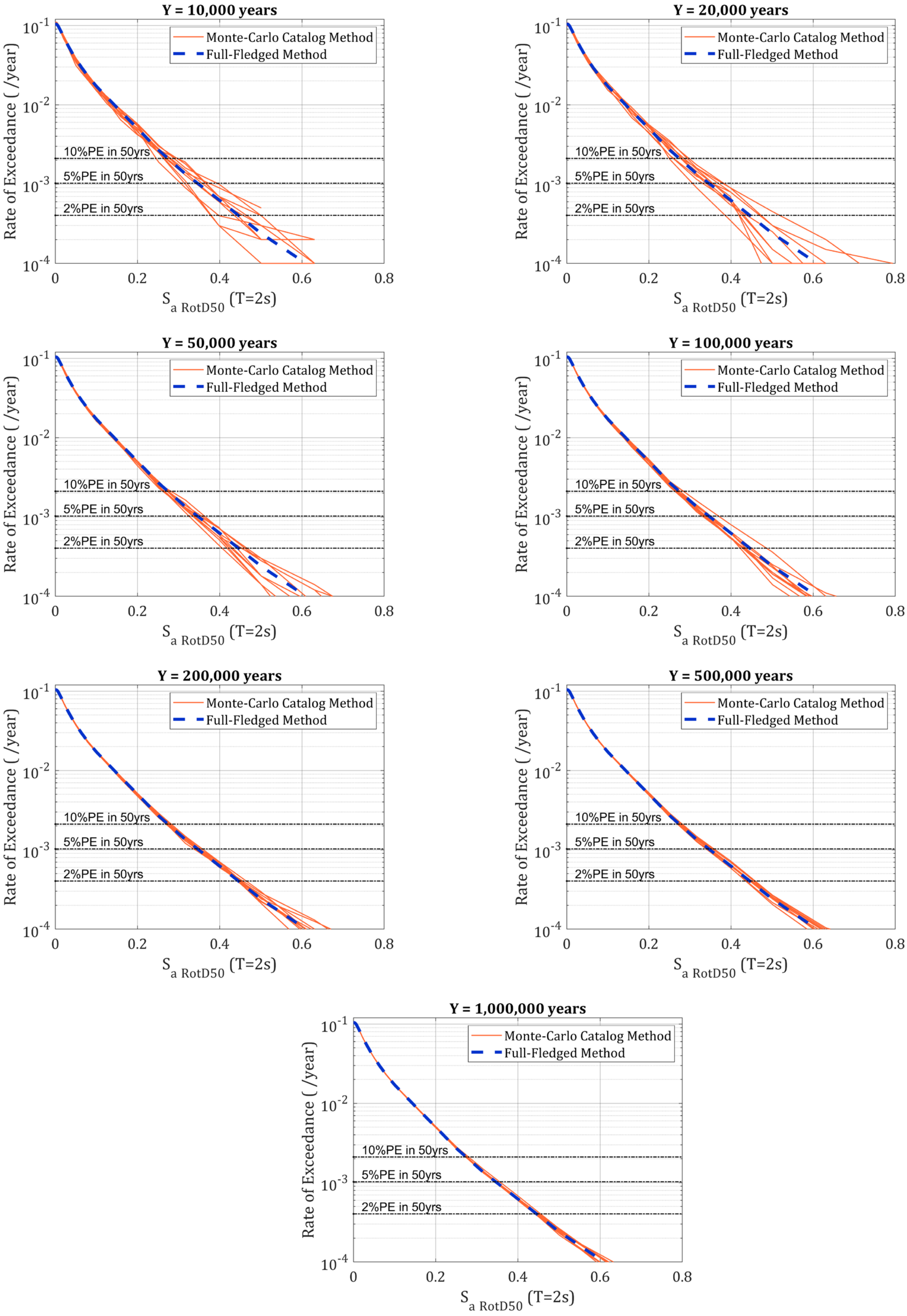

Fig. 1 - Comparison of $S_{a\ RotD50}$ hazard curves from the full-fledged method with hazard curves developed from Monte-Carlo catalogs at Downtown Los Angeles for $T = 2s$

The results underline the tradeoff that exists between the length of the catalog and the accuracy of the computed rates. In other words, a longer catalog results in more accurate hazard curves while a shorter catalog requires less computational resources. The selection of the optimal catalog length should balance between the required computational resources and the errors expected in the resulting hazard curves. Among the simulated catalogs, the 200,000-years catalog can be considered as the most convenient catalog for implementing Monte-Carlo catalogs in a simulation-based PSHA. The selected catalog length reduces the number of earthquake scenarios by more than 97% compared to the full-fledged method while still maintaining an acceptable error rate, with a median of 0.5% and a standard deviation of 1.6%.

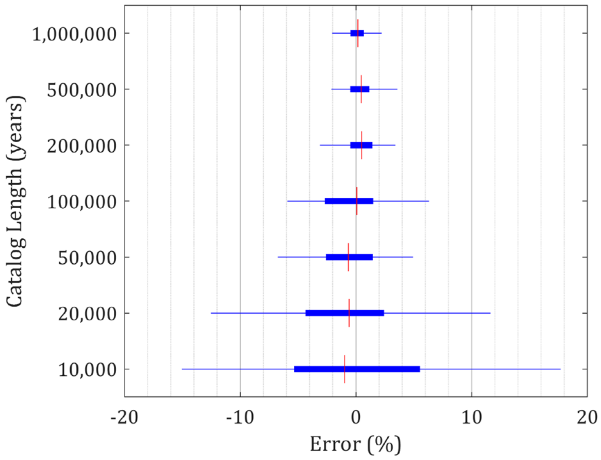

Fig. 2 – Boxplots of errors between annual rates of exceedance from the Monte-Carlo catalog method and the full-fledged method at the 5% probability of exceedance in 50 years hazard level

3.3. Disaggregation

To further validate the proposed method, a disaggregation of the magnitude M_w and the closest source-to-site distance R_{rup} of the contributing scenarios is applied to the 200,000-years catalogs at LADT at the different periods and hazard levels. As mentioned before, a set of 10 catalogs is generated for each catalog duration. In this section, we select a random catalog from the set to perform the disaggregation. Fig. 3 and Fig. 4 represent the $M_w - R_{rup}$ disaggregation at the 5% probability of exceedance in 50 years hazard level at periods of 0.5s and 5s, respectively. In both figures, the left plot corresponds to a disaggregation from the full-fledged method that uses the full set of scenarios, while the right plot is based on the Monte-Carlo catalog method. For the full-fledged method, the seismic hazard at $T = 0.5s$ is controlled by earthquake scenarios with magnitudes between 6.25 and 7.25 at short distances ($R_{rup} < 10\ km$). The same disaggregation pattern is observed for the 200,000-years catalog shown on the right side of Fig. 3.

The disaggregation patterns of the two methods are also similar at $T = 5s$, as illustrated in Fig. 4. At longer periods, earthquakes from larger distances are expected to be contributing. The first mode of the disaggregation from both methods is characterized by a R_{rup} between 50 and 60 km and a mean M_w equal to 8. At the same period, scenarios with magnitudes between 6.75 and 7.25 and $R_{rup} < 10km$ are also contributing to the hazard with similar proportions from both methods. Furthermore, contribution of scenarios with larger distances that exceed 100 km is also observed in both histograms.

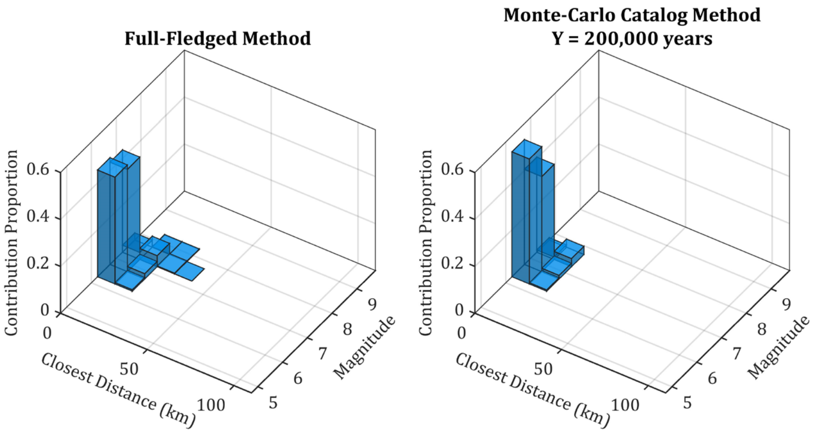

Fig. 3 - Magnitude-Distance Disaggregation at LADT for 0.5s $S_{a\,RotD50}(T)$ at 5% probabibility of exceedance in 50 years. Left histogram is derived from the full-fledged method and right histogram is from the Monte-Carlo catalog method using a 200,000-years catalog

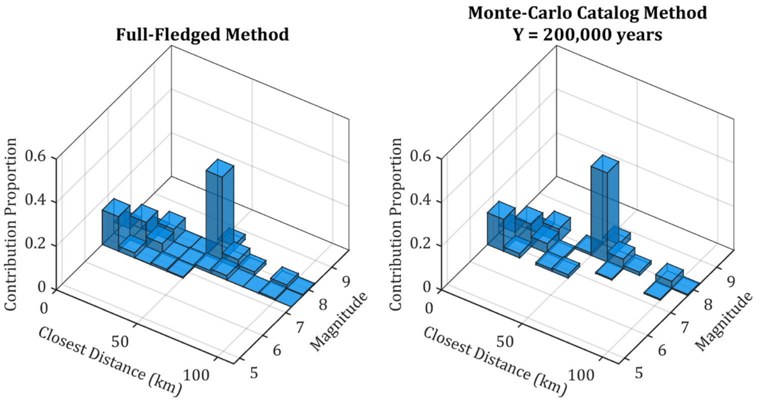

Fig. 4 - Magnitude-Distance Disaggregation at LADT for 5s $S_{a\,RotD50}(T)$ at 5% probability of exceedance in 50 years. Left histogram is derived from the full-fledged method and right histogram is from the Monte-Carlo catalog method using a 200,000-years catalog

Fig. 5 and Fig. 6 illustrate the source disaggregation of the 5% probability of exceedance hazard level obtained using both approaches at $T = 0.5s$ and $2s$, respectively. At $T = 0.5s$, the contributing sources from the full-fledged method and their contribution proportions are close to results from the Monte-Carlo catalog method. For instance, both plots in Fig. 5 show that the most contributing fault is the upper segment of the Elysian Park fault, followed by the Los Angeles section of the Puente Hills fault. For the longer period, a larger number of faults contribute to the hazard, including more distant faults. The results from both methods are consistent for the three most contributing sources as seen in Fig. 6. As the contribution proportion decreases, some differences are observed between the two methods, such as the case of the Puente Hills and Raymond faults. These minor differences are expected since the Monte-Carlo catalog method does not consider all the earthquake scenarios.

In summary, for LADT, the magnitude-distance and source disaggregation at short and long periods obtained using a 200,000-years Monte-Carlo catalog result in patterns that are mostly similar to the patterns obtained from the disaggregation of the complete set of scenarios. This similarity between the two methods is a key validation of the proposed Monte-Carlo catalog method.

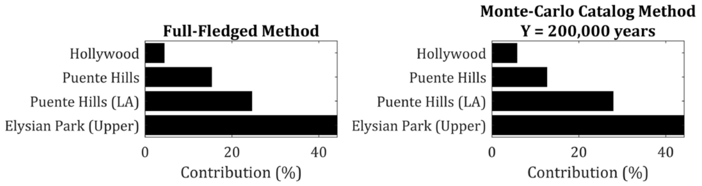

Fig. 5 - Source Disaggregation at LADT for 0.5s $S_{a\ RotD50}(T)$ at 5% probability of exceedance in 50 years. Left bar plot is derived from the full-fledged method and right histogram is from the Monte-Carlo catalog method using a 200,000-years catalog

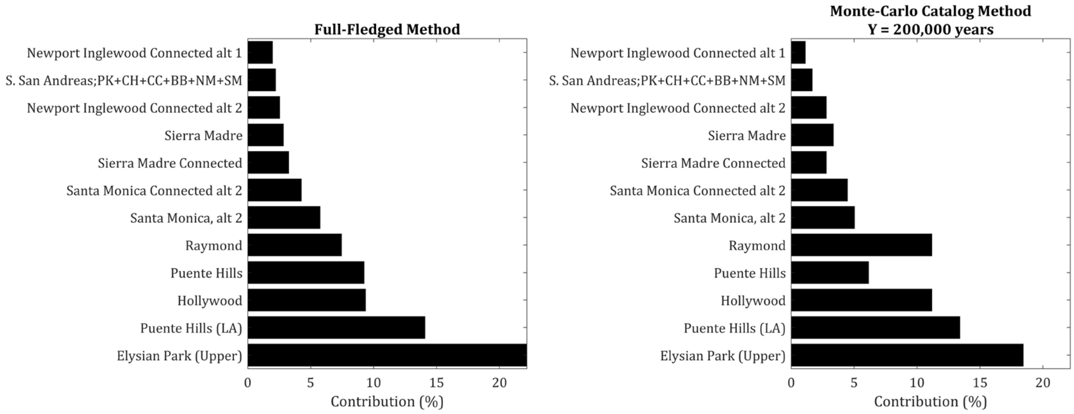

Fig. 6 - Source Disaggregation at LADT for 2s $S_{a\ RotD50}(T)$ at 5% probability of exceedance in 50 years. Left bar plot is derived from the full-fledged method and right histogram is from the Monte-Carlo catalog method using a 200,000-years catalog

4. Conclusions

In the current CyberShake PSHA approach, ground motion time series are generated for an average of 415,000 earthquake scenarios that can affect a site. Strain Green tensors are generated for all the faults and post-processing is applied to simulate a seismogram for each earthquake scenario. This method requires considerable computational resources due to the large number of earthquake scenarios and to the important processing requirements imposed by the numerical simulations. This paper presented a methodology that can reduce the computational requirements of the CyberShake platform by reducing the number of earthquake scenarios used in the analysis. The methodology is based on the development of earthquake catalogs of a specified duration at the site of interest. These catalogs specify the subset of earthquake scenarios to select among the complete set of scenarios that can affect the site.

To evaluate the accuracy of the proposed method, earthquake catalogs with different lengths were developed at a site located in Downtown Los Angeles. The hazard curves that correspond to the different simulated catalogs were compared to the hazard curves developed from the full-fledged method which, similary to CyberShake, considers the entire set of simulations. The comparison showed that the proposed method can result in accurate hazard curves and that the level of accuracy increases with the duration of the catalog. A Monte-Carlo catalog with a duration of 200,000 years resulted in hazard curves with mean errors of 0.5% at an annual rate of exceedance of 5% in 50 years. The number of ground motions required by this 200,000-years earthquake catalog is 97% less than what is currently required by the CyberShake approach. If a lower level of accuracy is acceptable, the catalog duration can be further decreased and a larger reduction in the number of earthquake scenarios can be achieved. The similarity in the disaggregation patterns between the proposed and the full-fledged method confirmed the validity of the Monte-Carlo catalog method in hazard calculations.

Implementing Monte-Carlo catalogs in the current CyberShake framework offers a reliable and practical tool. By reducing the time-to-solution per site, hazard maps can be developed for a larger region of interest. Additionally, more studies of CyberShake can be conducted and fast improvement of the platform can be achieved. Future research should further develop and confirm these initial findings by applying the same methodology to other sites. Similarly, future work should quantify the reduction in the CyberShake computational requirements in terms of data and CPU hours as a function of the length of the catalog.

5. Acknowledgments

We would like to thank Scott Callaghan (Southern California Earthquake Center) for his assistance with the CyberShake platform and for providing us with the simulations needed for this study.

6. References

[1] Cornell, C. A. Engineering seismic risk analysis. *Bulletin of the Seismological Society of America, 58*(5), 1583-1606. (1968)

[2] Kramer, S. Geotechnical Earthquake Engineering. Prentice-Hall, Inc. *New Jersey*, 348-422. (1996)

[3] McGuire, R. K. Probabilistic seismic hazard analysis: Early history. *Earthquake Engineering & Structural Dynamics, 37*(3), 329-338. (2008)

[4] Field, E. H., Dawson, T. E., Felzer, K. R., Frankel, A. D., Gupta, V., Jordan, T. H., Parsons, T., Petersen, M. D., Stein, R. S., Weldon, R. J., II, & Wills, C. J. Uniform California Earthquake Rupture Forecast, Version 2 (UCERF 2). *Bulletin of the Seismological Society of America, 99*(4), 2053-2107. (2009)

[5] Anderson, J. G., & Brune, J. N. Probabilistic Seismic Hazard Analysis without the Ergodic Assumption. *Seismological Research Letters, 70*(1), 19-28. (1999)

[6] Gregor, N., Abrahamson, N. A., Atkinson, G. M., Boore, D. M., Bozorgnia, Y., Campbell, K. W., Chiou, B. S.-J., Idriss, I., Kamai, R., & Seyhan, E. Comparison of NGA-West2 GMPEs. *Earthquake Spectra, 30*(3), 1179-1197. (2014)

[7] Graves, R., Jordan, T. H., Callaghan, S., Deelman, E., Field, E., Juve, G., Kesselman, C., Maechling, P., Mehta, G., Milner, K., Okaya, D., Small, P., & Vahi, K. CyberShake: A Physics-Based Seismic Hazard Model for Southern California. *Pure and applied geophysics, 168*(3), 367-381. (2011)

[8] Maechling, P. J., Silva, F., Callaghan, S., & Jordan, T. H. SCEC Broadband Platform: System architecture and software implementation. *Seismological Research Letters, 86*(1), 27-38. (2014)

[9] Azar, S., Dabaghi, M., & Rezaeian, S. (2019). *Probabilistic Seismic Hazard Analysis Using Stochastic Simulated Ground Motions*. Paper presented at the 13th International Conference on Applications of Statistics and Probability in Civil Engineering, ICASP13, Seoul, South Korea.

[10] Boore, D. M. Orientation-independent, nongeometric-mean measures of seismic intensity from two horizontal components of motion. *Bulletin of the Seismological Society of America, 100*(4), 1830-1835. (2010)