
# UNIVERSITÀ DEGLI STUDI DI CAMERINO

## School of Advanced Studies

**DOCTORAL COURSE IN**

***PHYSICAL AND CHEMICAL PROCESSES IN EARTH SYSTEMS***

**34th cycle**

# New approach for seismic hazard analysis and earthquake damage scenarios

**PhD Student**
**Hamed Dabiri**

**Supervisors**
**Prof. Andrea Dall'Asta**

**Co- supervisors**
**Prof. Emanuele Tondi**

**Abstract**

Seismic hazard causes a considerable portion of loss in many countries annually and thus is of high importance. Seismic hazard and its consequences, known as seismic risk, have been studied in the fields of geology, earthquake engineering and structural engineering in the framework of probabilistic seismic hazard analysis (PSHA). PSHA could be simply defined as the probability of occurrence of an earthquake characteristic (e.g., PGA), considering uncertainties namely magnitude, location and their resulting ground motion specified by ground motion prediction equation (GMPE). The final output of PSHA is the rate of exceeding specific intensity measure (IM), and could be expressed in terms of return period exceedance. For structural design purposes, for a given probability and structure's design life, the seismic hazard could be stated in terms of return period $T_R$, which is commonly applied in current design code (e.g., Eurocode 8).

The above-mentioned explanations of return period of exceedance are based on the Poisson relationship between time and earthquake which assumes that the probability of an earthquake occurring remains unchanged by elapsing time. In the recent decades, however, it has been claimed that earthquake occurrence could be expressed through time-dependent models which means that elapsing time since the last earthquake affects its occurrence probability.

This research focuses mainly on the effect of time-dependent seismic hazard on structural design, by evaluating the strength required by the structure (seismic capacity) for different time intervals elapsing from the last event. "Seismic capacity" is defined as the capacity required to provide a fixed failure rate. Two different seismic scenarios (i. e., a point source and a combined source including area and line source) have been investigated and results concerning different site-to-source distance, capacity dispersion of the structure and different recurrence properties of the time-dependent source have been discussed. The results obtained from the analyses highlight a significant influence of time-dependent hazard properties on the structural capacity required to attain a target reliability, and give evidence to the different roles played by the parameters considered in the analysis.

Moreover, in order to deeper investigate the effect of time-dependent seismic hazard on structural design, the influence of soil classification, period and the GMPE implemented in the analysis have been assessed and the results discussed extensively. The analysis outcomes illustrate the remarkable impact of soil and period on structural response as well as the importance of appropriate GMPE used for the time-dependent seismic hazard.

Furthermore, machine-learning (ML) based models have been proposed for deriving fragility curves of buildings. Generating fragility curves is a critical key step in the Performance-Based Earthquake Engineering (PBEE) framework which is generally time-consuming. The accuracy of the quick accurate models proved the high reliability of ML-based techniques for obtaining fragility parameters namely dispersion and median. The developed ML-based prediction models could be used for estimating capacity (both time-dependent and time-independent cases) in further studies.

**Funding**

This research is funded by the university of Camerino (Italian: *Università degli Studi di Camerino*) and National Institute of Geophysics and Volcanology (Italian: *Istituto Nazionale di Geofisica e Vulcanologia, INGV*). The funding and technical supports are gratefully acknowledged.

**Acknowledgment**

First and foremost, I would like to express my deepest gratitude to my esteemed supervisors, Prof. Andrea Dall'Asta and Prof. Emanuele Tondi for their continuous guidance, support and patience throughout my PhD research. Their immense knowledge and plentiful experience have encouraged me in all the time of my academic research.

I am extremely grateful to my supervisor, Prof. Asaad Faramarzi at the University of Birmingham, the host University of my visiting period, for his treasured support and help for doing a part of this study.

My gratitude extends to the Faculty of Geology for the funding opportunity to undertake my studies at the school of science and technology, University of Camerino. Additionally, I would like to extend my sincere thanks to The National Institute of Geophysics and Volcanology (Italian: Istituto Nazionale di Geofisica e Vulcanologia, INGV) for providing us with a partial financial support.

I owe a big thank you to Prof. Enrico Tubaldi for giving me the opportunity to do my PhD research at the University of Camerino under my respected supervisor, by recommending me for the position.

I would like to thank the referees for taking the time to review my research. Their constructive comments improved the quality of our study and the results undoubtedly.

I gratefully acknowledge the assistance of the research team, who are my best friends as well, for their help and support from the beginning of my PhD to the end.

Last but not least, from the bottom of my heart I would like to say a big thank you to my family, specifically my wife and my parents, who always supported me with their love and patience. Without their tremendous understanding and encouragement in the past few years, it would have not been possible for me to complete my study.

Hamed Dabiri

May 2022

**Contents**

## List of Tables

## List of Figures

# 1 Research background, motivation, objectives and organization

## 1.1 Definitions

Definition of the frequent used terms in this thesis are provided in this subsection. The main terms are stochastic processes, counting process, interarrival and waiting time, memory, recurrence interval, return period:

- *Stochastic processes*: a stochastic process could be defined as a system in which there are observations at certain times, and that its target outcome, that is, the observed value at each time is a random variable (Blomberg 2006). The fundamental concepts of stochastic processes are based on generating (pseudo) random numbers and the sequences of uniform random number provided by computer routines is an example of stochastic process (Slepchenko & Loew 2010, Ferrari 2001).
- *Counting process*: a counting process is a type of stochastic process with non-negative, integer, and increasing values. The most common application of a counting process is to count the number of occurrences of some event of interest by elapsing time (time-dependent events), and therefore the values are usually chosen non-negative real numbers $[0, \infty)$. It is worth explaining that in the general form, negative values $R = (-\infty, \infty)$ are also commonly used) (Olofsson 2006).
- *Arrival/ arrival time*: the appearance of seismic energy on a seismic record is known as arrival while the time at which a particular wave / phase arrives at a station or detector is defined as arrival time.
- *Recurrence Interval*: the approximate length of time between earthquakes in a specific seismically active area.
- *Recurrence interval (*or return period)*:* the average time span between earthquake occurrences on a fault or in a source zone is called recurrence interval (or return period).

## 1.2 Background

One of the main concerns in earthquake engineering is to provide structural engineers with parameters which lead to a desire level of structure performance in a given ground shaking level. But there are always uncertainties in various factors including location, size and resulting shaking intensity of future earthquakes. Probabilistic seismic hazard analysis (PSHA) is used to provide a description of the future shaking by quantifying and combining mentioned uncertainties (Baker 2008).

The outcomes of PSHA could be used for determining return period of exceedance which is defined as the probabilities of exceeding a given ground motion intensity within a specified time interval for a given rate of exceedance. Currently the Poissonian model is used for simulating exceedance probability which assumes that an earthquake occurrence probability remains constant by elapsing time since the last event.

In the view point of structural design, the probability of a ground motion level within a given design life of a structure is considered. As a result, the seismic hazard could be also expressed in terms of return period (TR), specifically in design codes and provisions (e.g. Eurocode 8) (Shahbazi et al. 2018).

Models which are used in PSHA are generally divided into two main categories: time-independent and time-dependent models. It should be taken into account that both time-dependent and time-independent models are based on the "characteristic earthquake model" (magnitude-frequency characteristic of earthquake occurrence) which assumes that the seismic activity on a given segment should be dominated by earthquakes of similar characteristics that recur at somewhat regular intervals. Both of them also require moment-balanced models which are consistent with the global plate rate models and slip rate determined on individual faults (Peterson et al. 2007).

One of the most well-known time-independent models which is mostly used by earthquake engineers is Poisson models (Eq. 1.1).

$$f_T(t) = \frac{1}{T_R} e^{-t/T_R} \qquad (1.1)$$

where $T_R$ is the return period. In Poisson models the time elapsed since the last earthquake is not considered. It should be taken into account that they can be used for small earthquakes because they do not model large earthquakes on specific faults properly (Jalalalhosseini et al.

2017). Poisson models are appropriate for sites with a single source where the time since the last earthquake is more than the meantime interval and the source has strong characteristic time behaviour (Mousavi and Salehi 2018).

On the other hand, non-Poisson models which are also known as renewal time-dependent models, are appropriate for large earthquakes (Jalalalhosseini et al. 2017). Results of time-dependent models are broadly consistent with the elastic rebound theory of earthquakes. Parameters including static elastic fault interaction, visco-elastic stress transfer and dynamic stress changes from earthquakes on close faults may affect the results of time-dependent earthquake probability. But the most important factor is the time elapsed since the last event (because of stress increment on fault) (Petersen et al. 2007).

The distributions which are commonly used in renewal models are normal, lognormal, gamma and Weibull distributions. It is noteworthy that

- Lognormal distribution is appropriate for model inter-event time distribution for large earthquakes.
- Normal distribution is not appropriate sometimes since it should be truncated at t=0 due to impossibility of negative inter-event time (Hebden and Stein 2009).
- Gamma distribution is recommended as a probability model for waiting time (Bak et al. 2002).
- Lognormal or Gamma distribution are recommended for a probabilistic treatment of crustal rapture time (Estera 1970).
- The Weibull distribution function is commonly considered for the PDF of inter-event time (Hagiwara 1974).

One of the models which has attracted researchers' attention to properly describe the probability distribution of inter-event time is Brownian Passage Time (BPT) (Matthew et al. 2002).

In the BPT model, which has been proposed to describe the probability distribution of inter-event time (Matthew et al 2002), earthquake occurrence is assumed to have periodicity (Jalalalhosseini et al. 2017).

In fact, in the BPT model, which is based on Brownian relaxation oscillator, the load state increases steadily by elapsing time, reaches a failure threshold and relaxes instantaneously back

to the ground level. BPT is defined by two parameters: μ and α, which are mean time (or time between earthquakes) and aperiodicity of meantime, respectively. α is the coefficient of variation: α=1 refers to irregular Poissonian behaviour and α=0 refers to periodic behaviour (Matthew et al. 2002).

The probability density function for BPT model is given in Eq. (1.2):

$$f_T(t) = \sqrt{\frac{T_R}{2\pi\alpha^2 t^3}} \cdot e^{\frac{(t-T_R)^2}{2T_R\alpha^2 t}} \qquad (1.2)$$

where $T_R$ is the return period and $\alpha$ is the aperiodicity of the interarrival time. As it was mentioned before, by decreasing α, the hazard function increases whilst it becomes Poisson-like when α increases (=1) (Jalalalhoseeini et al. 2017). Hazard function provides the instantaneous probability of occurrence at the time *t*, given that no event had occurred previously, and describes the hazard variation in time (Eq. 1.3).

$$h_T(t) = \frac{f_T(t)}{1 - F_T(t)} \qquad (1.3)$$

where $F_T(t)$ is cumulative distribution function (CDF). According to Nishenko and Buland (1987) lognormal distribution for $F_T(t)$ leads to more reliable results. The BPT is approximately indistinguishable from the lognormal distribution for up to three recurrence times and has several advantages in comparison to lognormal distribution. Firstly, BPT is based on a physical process which is the reflection of the stress evolution at one point. Secondly, by elapsing time since the last event, lognormal hazard rate tends toward zero probability (which is not reasonable) whilst BPT tends toward $(2\mu\alpha^2)^{-1}$. Finally, stress changes in the faults and their influence on earthquake triggering are easily incorporated in BPT. However, the problem with using BPT is that changing stress is probably not uniform on all the fault length (Smith 2006). It should be taken into consideration that one of the fundamental problems of the BPT under the effect of stress changes of real fault is that, changes in stress are not distributed uniformly across the fault due to their significant spatial extent (Boyd et al. 2008, Smith 2006, Zafarani and Ghafoori 2013).

## 1.3 Motivations, novelty and significance

Motivation and significance of this research could be denoted in two parts:

- In the first part of the study, concerning the influence of time-dependent seismic hazard on structural design, the focus is on evaluating structures' failure probability considering time-dependent seismic hazard. As mentioned previously, the final output of PSHA could be expressed as the probabilities of exceeding a given ground motion intensity within a specified time interval for a given rate of exceedance. This parameter has been used in design codes (e.g., Eurocode 8) by implementing time-independent seismic hazard defined by Poisson model (POI).

  In the current research, whereas, time-dependent seismic hazard is applied for the first time for determining structures' failure probability. It is also worth mentioning that previous studies carried out on time-dependent seismic hazard mainly aimed at evaluating and comparing various time-dependent hazard models (e.g., BPT, Weibull, Gama, etc.) and their influence on the outcomes' accuracy. This study, however, emphasizes the application of time-dependent seismic hazard analysis in the viewpoint of structural engineering. Otherwise noted, time-dependent seismic hazard is used to figure out the possible effects on structural design. Briefly noted, the significance of the first part is (i) implementing time-dependent seismic hazard for the purpose of structural engineering for the first time and consequently, (ii) providing the effect of time-dependent seismic hazard on the design process to be considered in the next revisions of design codes.

- As far as the second part is taken into account, the research motivation could be expressed as the need for proposing models for easier obtaining fragility curves which is a key step in performance-based earthquake engineering (PBEE). It should be explained that the main objective of PBEE is to design structures that are capable of demonstrating anticipated desirable performance objectives, in contrast to the conventional approach of designing structures to strictly satisfy the codes' provisions. PBEE is based on accurate prediction of seismic capacities and demands. It utilizes the pre-defined performance objectives which combines the damage or performance limit states with the seismic hazard level. By the application of PBEE, making decisions concerning the choice of structural systems and detailing levels based on life-cycle performance and cost analysis would be possible (AlHamaydeh et al, 2019). It is worth recalling that the conventional methods used for obtaining fragility curves fundamentally includes steps namely (i) collecting appropriate seismic records, (ii)

scaling the collected records, (iii) developing numerical model of a structure which could simulate its nonlinear behavior, (iv) analyzing the numerical model for a series of scaled records until reaching the building failure, and (v) reaping step iv for each record. The above-mentioned steps are recognized as Incremental Dynamic Analysis (IDA). The most significant drawback of IDA is being time-consuming, specifically in the case of analyzing tall and complicated buildings. A few studies, therefore, have been carried out to date in order to develop and present alternative methods using machine learning-based approaches. IDA analysis, however, still needs to be conducted in the recent ML-based models.

As a result, considerably faster and more accurate models are developed and introduced here. IDA is eliminated and the fragility curves could be obtained by defining structures' properties (i. e., plan area, height, period, construction material, lateral resisting system, soil classification of the building location, damage state and design code).

## 1.4 Aims and objectives

Earthquakes are one of the natural hazards which pose the greatest threat to life and could lead to remarkable financial losses. The main objective of this research is therefore evaluating performance of the structures subjected to time-dependent seismic load in order to enhance their response to dynamic loads. In other words, this study aims at designing buildings more accurately by considering realistic scenarios which generally involve time-dependent seismic hazard. Besides, the application of machine learning-based techniques for analysing structures' behaviour is assessed for the sake of easiness and accuracy. More specifically, the main research objectives are as below:

→ *State-the-art-review on time-dependent seismic hazard and structural response*

Previous studies conducted on time-dependent hazard analysis were carefully evaluated in terms of models introduced so far, their adjustable variables and their results. More clearly, the review aimed at finding out the most accurate time-dependent model implemented for simulating seismic occurrence. The appropriate variables which led to the most reliable results are determined and used for this study. The most recent studies on time-dependent

structural response were also reviewed in order to fill the research gap and provide practical conclusions to be considered in the design codes and analysis process.

→ *Influence of time-dependent seismic hazard on structural design*

Considering earthquake occurrence as a periodical model, known as time-dependent seismic hazard, has been extensively assessed by researchers. The influence of time-dependent seismic hazard on structures' behavior has not been yet analyzed sufficiently. The main concern of the present research, as a result, is to clearly realize how time-dependent seismic hazard affect structural behavior by elapsing time. More clearly, the response of structures subjected to time-dependent and time-independent seismic hazard was compared together. The effect of site-to-source distance, structural dispersion and aperiodicity parameter of time-dependent hazard model was also investigated analytically.

→ *Influence of soil classification, period and GMPE on time-dependent seismic hazard and structural response*

In order to deeper assess the influence of time-dependent hazard on structures' behaviour, other potential effective parameters were considered in our analytical study. Otherwise mentioned, the aim of this phase was to assess if site location (soil classification), period and the GMPE used in the analysis could have a significant effect on the final outcomes.

→ *Reviewing both analytical and machine learning-based models for deriving fragility curves and developing ML-based models for obtaining fragility parameters and fragility curves consequently*

This section was mainly aimed at developing a model for obtaining fragility curves easier. To this end, the models introduced based on either analytical analysis or ML-based techniques are reviewed and their shortcomings are determined. Accordingly, ML-based models are developed for yielding fragility parameters of buildings accurately and in a quite short time. It is worth explaining that the time-consuming analytical analysis required for Incremental Dynamic Analysis (IDA) could result in inaccurate results sometimes. In the prediction models provided in this section the IDA is eliminated and therefore they pose the benefits namely (i) easier implementation, (ii) quicker performance and (iii) more accurate results in comparison to conventional approaches. These models return fragility

parameters of reinforced concrete (RC), steel and masonry buildings by considering buildings properties and site characteristics.

## 1.5 Thesis organization

The present thesis is organized in two main parts: part I (sections 2, 3, & 4) which focuses on time-dependent hazard analysis and its influence on structural design, Part II (section 5) which aims at developing ML-based methods for deriving fragility curves of buildings. It should be explained that each section is organized independently of other sections and readers do not need to refer to previous sections. Organization of each section is summarized below:

→ **Chapter 2**: *State-of-the-art review on time-dependent seismic hazard and Time-dependent structural response:*

In the second section of the thesis, the most significant and recent studies carried out on time-dependent seismic hazard are collected from highly ranked international publications. Their aim, methodology and remarkable conclusions are presented and discussed. The review focused on the time-dependent hazard models implemented so far and the influence of their parameters on the analysis outcomes. More importantly, time-dependent structural response assessed in previous researches are provided and their conclusions are discussed as well.

→ **Chapter 3**: *Influence of time-dependent seismic hazard on structural design*

In the third part of the research, the implemented methodology which was fundamentally based on the application of PSHA and fragility curves in different time intervals is extensively explained. Next, two seismic sources are defined: (i) a point source (both time-dependent and time-independent) and (ii) a more realistic source combined of an area source (time-independent) and a line source (time-dependent). The results of a time-dependent seismic source were compared and analyzed to those of a time-independent source. Furthermore, the effect of site-to-source distance (source characteristics), structural response dispersion (structure property) and aperiodicity of time-dependent hazard model (BPT) on both seismic hazard and structure behavior (seismic capacity) was assessed. It is worth explaining that "seismic capacity" is defined here as the capacity required to provide a fixed reliability level, measured by the failure rate. Then, the structure response variation in different times elapsed since the last event (earthquake) is discussed. The conclusions of the first part, highlighted a remarkable effect of time-dependent seismic hazard on structural design.

→ **Chapter 4**: *Influence of soil classification, period and GMPE on time-dependent seismic hazard and structural response*

The influence of time-dependent seismic hazard on structural design is deeper investigated in the fourth section of the current thesis by considering different GMPEs, soil classifications and periods. GMPEs proposed by (Ambraseys et al. 1996) and (Sabetta and Pouliese 1996) ware applied in the analysis and the results compared together. The effect of soil and period on the results was also assessed by changing the Ambrasyes GMPE which are dependent on both soil type and period. Seismic hazard and structural capacity obtained for each set of the above-mentioned variables was analyzed in different time elapsed since the last earthquake and the outcomes are discussed.

→ **Chapter 5**: Fragility curves of buildings; a critical review and a ML-based evaluation

The fifth section aimed at proposing machine learning-based models for obtaining fragility curves of structures. To this end, firstly the most recent analytical and ML-base models developed are reviewed in order to find out their shortcomings and research gap. Then, a database was collected from high-ranked international publications for training and testing the proposed models. Building properties (plan area, height, period, resisting system) soil classification, design code and damage state were the input for estimating the target outputs (dispersion and median). Various ML-based techniques namely nonlinear regression, decision Tree, Random Forest, K-Nearest Neighbors and Artificial Neural Network were used for developing prediction models. The accuracy of the models is assessed by performance metrics and Taylor diagram. The results proved the high reliability of developed ML-based models for obtaining fragility curves using the defined input variables. More importantly, a sensitivity analysis conducted in order to realize the influence of input parameters on fragility behavior of buildings.

→ **Chapter 6**: Summary and conclusions

A summary of the main aims and methodologies of the research is provided in the last section. The remarkable outcomes are presented as well and the route to further studies is suggested regarding the outcomes of this thesis.

The research outcomes have been published in journal papers and presented in conferences as listed below:

- Dall'Asta, A., **Dabiri, H.**, Tondi, E., & Morci, M. (2021). *Influence of time-dependent seismic hazard on structural design*. *Bulletin of Earthquake Engineering*, 19(6), 2505–2529. doi:10.1007/s10518-021-01075-3
- **Dabiri, H.**, Faramarzi, A., Dall'Asta, A., Tondi, E., Morici, M., (2022), *Predicting Fragility Curves of RC structures Using ML-based techniques, Journal of Building Engineering, under review*
- **Dabiri, H.**, Dall'Asta, A., Tondi, E., Morici, M., (2019), *Evolution of structural capacity in the case of time-dependent point source*, GNGNTS 2019, 12-14 November, Rome, Italy.
- **Dabiri, H.**, Dall'Asta, A., Tondi, E., Morici, M., (2019), *Preliminary study on the impact of time-dependent seismic hazard on design capacity*, XVIII ANIDIS congress, 15-19 September, Ascoli Piceno, Italy.

# References

AlHamaydeh, M., Aly, N., Najib, M., & Alawnah, S. (2019). INSPECT-PBEE: A performance-based earthquake engineering GUI for IDARC-2D. SoftwareX, 9, 132-144.

Allin Cornell, C. and S. R. Winterstein (1988). "Temporal and magnitude dependence in earthquake recurrence models." Bulletin of the Seismological Society of America 78(4): 1522-1537.

Abrahamson, N. A. and W. J. Silva (1997). "Empirical response spectral attenuation relations for shallow crustal earthquakes." Seismological Research Letters 68(1): 94-127.

Bak, P., et al. (2002). "Unified scaling law for earthquakes." Physical Review Letters 88(17): 178501.

Baker, J. W. (2008). "An introduction to probabilistic seismic hazard analysis (PSHA)." White paper, version 1: 72.

Blomberg, C. (2006). Fluctuations for good and bad: the role of noise in living systems. *Physics of Life Reviews*, *3*(3), 133-161.

Boyd, O. S., et al. (2008). "Toward a time-dependent probabilistic seismic hazard analysis for Alaska." Active Tectonics and Seismic Potential of Alaska, Geophysical Monograph 179: 399-416.

Eurocode 8, E. S. E.-. (2004). Eurocode 8. Design of structures for earthquake resistance - Part 1: General rules, seismic actions and rules for buildings.

Ferrari, P. A. (2001). Stochastic Models, https://doi.org/10.1016/B0-08-043076-7/00592-1.

Hagiwara, Y. (1974). "Probability of earthquake occurrence as obtained from a Weibull distribution analysis of crustal strain." Tectonophysics 23(3): 313-318.

Jalalalhosseini, S. M., et al. (2018). "Time-dependent seismic hazard analysis for the Greater Tehran and surrounding areas." Journal of Seismology 22(1): 187-215.

Matthews, M. V., et al. (2002). "A Brownian model for recurrent earthquakes." Bulletin of the Seismological Society of America 92(6): 2233-2250.

Mousavi, M. and M. Salehi (2018). "Temporal distribution of earthquakes using renewal process in the Dasht-e-Bayaz region." Journal of Seismology 22(1): 153-159.

Olofsson, P. (2006). Counting process. *Encyclopedia of Environmetrics*.

Petersen, M. D., et al. (2007). "Time-independent and time-dependent seismic hazard assessment for the State of California: Uniform California Earthquake Rupture Forecast Model 1.0." Seismological Research Letters 78(1): 99-109.

Sabetta, F. and A. Pugliese (1996). "Estimation of response spectra and simulation of nonstationary earthquake ground motions." Bulletin of the Seismological Society of America 86(2): 337-352.

Shahbazi, S., et al. (2018). "Effect of soil classification on seismic behavior of SMFs considering soil-structure interaction and near-field earthquakes." Shock and Vibration 2018.

Slepchenko, B. M., & Loew, L. M. (2010). Use of virtual cell in studies of cellular dynamics. *International review of cell and molecular biology*, *283*, 1-56.

Smith, D. E. (2006). A new paradigm for interpreting stress inversions from focal mechanisms: How 3D stress heterogeneity biases the inversions toward the stress rate, California Institute of Technology.

Zafarani, H. and S. Ghafoori (2013). "Probabilistic assessment of strong earthquake recurrence in the Iranian plateau." Journal of Earthquake Engineering 17(3): 449-467.

# 2 Literature review

Various studies have been conducted on using time-independent and time-dependent hazard analysis. Some of them are presented in this section.

## 2.1 Jalalalhosseini et al. (2017)

Jalalalhoseini et al. (Jalalalhosseini, Zafarani et al. 2018) analysed time-dependent seismic hazard for Tehran city (located in Iran) and surrounding area. They utilized the Brownian Passage Time model (Matthews, Ellsworth et al. 2002) to predict the seismic performance of active faults in Tehran. They presented the total hazard curves by combining the results of the BPT model (with different values of aperiodicity parameter) with the hazard from small to moderate background seismicity. The results of their study were presented by seismic hazard maps which demonstrated the probabilities of exceeding different ground motion levels at any site due to the earthquake in seismic sources in a special period.

In their study, to evaluate the influence of aperiodicity parameter on the hazard, three values of $\alpha$ (0.25, 0.5 and 0.75) were considered and investigated. Based on their research, $\alpha$=0.5 leads to more reliable results rather than other values.

One of the most significant conclusions they reached in their study was that there is a notable difference (10%-20%) between time-dependent maps and time-independent maps near a fault source. Otherwise stated, time-independent maps are similar to time-dependent maps in the areas which are far from sources. The other notable point of their study was that the areas with relatively short period since the last earthquake, the hazard of time dependent map is less than the hazard of time-independent map.

## 2.2 Mousavi and Salehi (2018)

In Mousavi and Salehi study (Mousavi and Salehi 2018), the recurrence intervals for Dasht-e-Bayaz region, eastern Iran was evaluated by applying Weibull, Gamma, lognormal and Brownian Passage Time. According to their research, the Weibull model has the benefit of overcoming the limitation of both exponential and normal distribution. On top of this, based on their research, the outcomes obtained by Weibull distribution demonstrated a better consistency.

Primary required variables to estimate model parameters were $T_R$ (mean recurrence interval) and $S_R$ (standard deviation of recurrence interval) which were considered 9 and 8 years, respectively. The parameters assumed for the four models and the standard deviation were as Table. 2.1 and 2.2:

***Table 2-1**. Parameter estimation from the Maximum Likelihood Estimation method.*

| *Probability model* | *Parameter* | *MLE value* |
|---|---|---|
| *Lognormal model* | *m* | *1.75* |
| | *σ* | *1.1* |
| *Weibull model* | *α* | *0.077* |
| | *β* | *1.12* |
| *Gamma model* | *r* | *1.17* |
| | *c* | *0.12* |
| *BPT* | *μ* | *9.33* |
| | *α* | *0.72* |

***Table 2-2.** Standard deviation of models based on Utsu (1984) and Matthews et al. (2002) studies.*

| *model* | *Variance ($V[T_i]$)* | *Standard deviation* |
|---|---|---|
| *Weibull* | $\alpha^{\frac{-2}{\beta}}\{\Gamma(\frac{2}{\beta}+1)-\Gamma^2(\frac{1}{\beta}+1)\}$ | *8.47* |
| *Gamma* | $r^2/c^4$ | *81.25* |
| *Lognormal* | $e^{2m+\sigma^2}(e^{\sigma^2}-1)$ | *16.16* |
| *BPT* | $(\mu\alpha)^2$ | *45.13* |

To find the most proper model for the site, two methods were used to choose the best distribution: Bayesian information criterion (BIC) as given in Eq. 2 and maximum likelihood criterion which uses the maximum logarithmic of likelihood value to find the best model.

$$BIC = KLn(N) - 2Ln(L) \qquad (2.1)$$

where N is the number of observed data and K is the number of parameters which describes the model. The model with lower absolute values of Ln L and BIC can be selected as the best model. Hazard function for Dasht-e-Bayaz was calculated and drawn in Fig. 2.1.

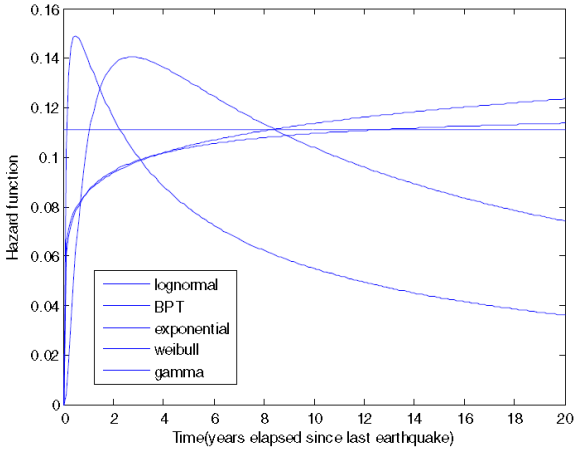



***Figure 2-1****. Time-dependent and time-independent hazard functions for the Dasht-e-Bayaz region in east of Iran Mousavi and Salehi (2018).*

Considering Fig. 2, it can be said that hazard function values of the lognormal model increase near the mean recurrence time (1.75), while it decreases sharply by elapsing time since the last event. This reduction in the occurrence rate in lognormal distribution can be considered as one of its defects. The occurrence rate in Weibull and Gamma models increases constantly by elapsing time since the last event.

## 2.3 Zafarani and Ghafoori (2013)

In this study (Zafarani and Ghafoori 2013), the earthquake recurrence intervals for the Iranian Plateau were evaluated by considering three models: Gamma, lognormal and the BPT.

The approach in which the parameters of the models are estimated by likelihood function for lognormal and Gamma distribution can be found in (Utsu 1984).

The likelihood function of the BPT model is as Eq. 2.2:

$$L = f_T(t_1,t_2,t_3,...,t_n \mid \mu,\alpha) = \prod_{i=1}^{n} \left(\frac{\mu}{2\pi\alpha^2 t_i^3}\right)^{1/2} \exp\left(\frac{-(t_i-\mu)^2}{2\alpha^2 \mu t_i}\right) \qquad (2.2)$$

where n is the number of earthquakes.

The maximum likelihood estimates of model parameters μ and α can be obtained by calculating following Eq. 2.3-2.6:

$$\frac{\partial \ln L}{\partial \mu} = \frac{n}{2}(\frac{1}{\mu}) + (-\frac{1}{2})[\sum_{i=1}^{n} \frac{-2(t_i - \mu)}{\alpha^2 \mu t_i} - \sum_{i=1}^{n} \frac{-(t_i - \mu)^2}{\alpha^2 \mu^2 t_i}] \quad (2.3)$$

$$\frac{\partial \ln L}{\partial \alpha} = -\frac{n}{\alpha} + (\frac{1}{\mu \alpha^3})[\sum_{i=1}^{n} \frac{(t_i - \mu)^2}{t_i}] \quad (2.4)$$

$$\partial LnL / \partial \mu = 0 \quad (2.5)$$

$$\partial LnL / \partial \alpha = 0 \quad (2.6)$$

Fig. 2.2 demonstrates the cumulative distribution of the observed inter-event time calculated for 15 sources and the curves of cumulative function obtained using three models.

Based on the study conducted by Zafarani and Ghafoori, considering the logarithmic of likelihood function (Ln L), the lognormal model leads to more reliable results for regions with intermittent occurrence characteristics in Iran. On the other hand, Gamma distribution was the worst in comparison to the other models.

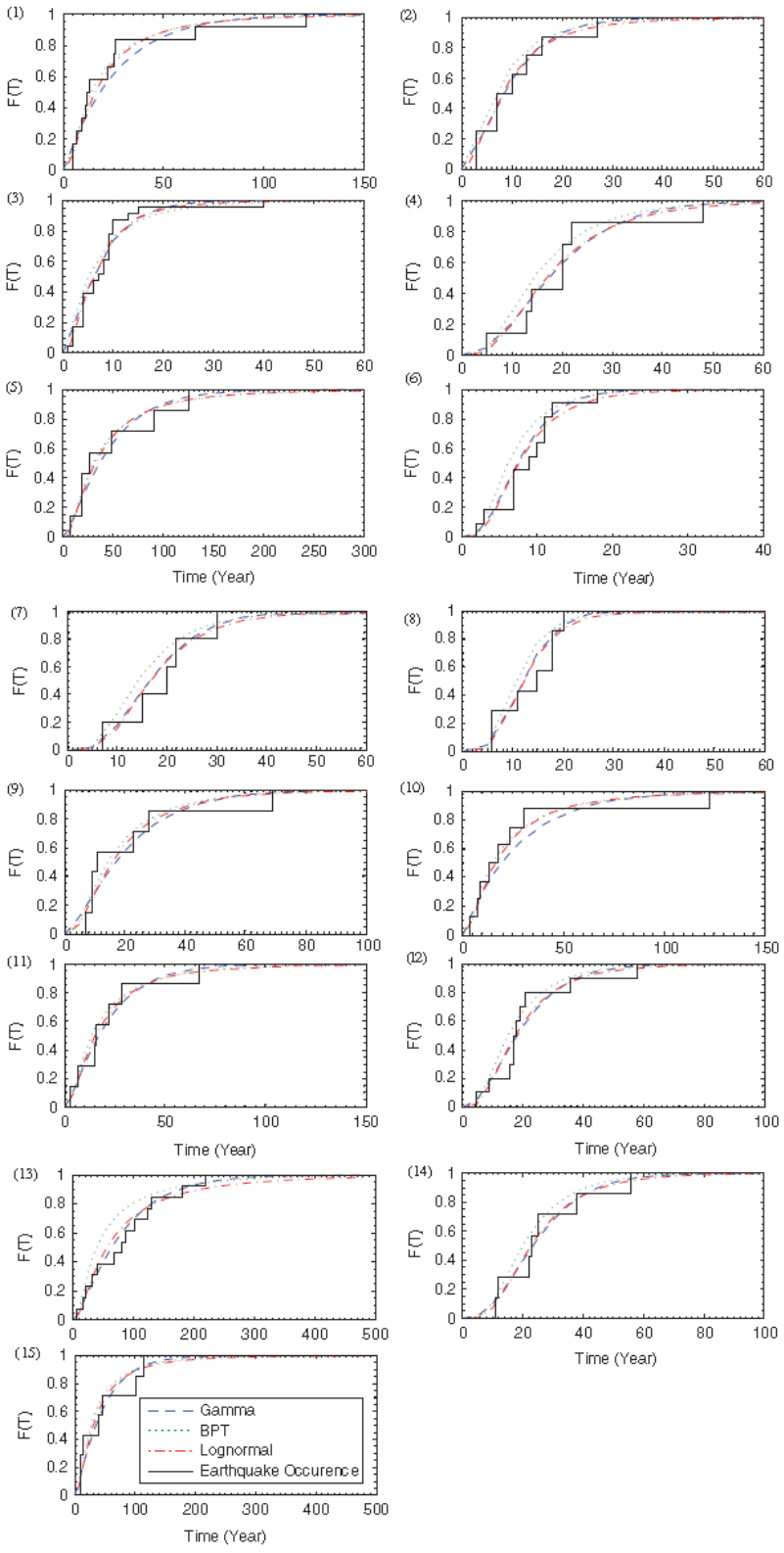



***Figure 2-2.*** *Cumulative distribution of the observed time intervals and the curves of FT (t) = P(T < t) using three models for 15 regions (Zafarani and Ghafoori 2013).*

## 2.4 Akinci et al. (2009)

In this research (Akinci, Galadini et al. 2009), time dependent models were characterized using BPT model to assess probability seismic hazard for central Apennines, Italy. The influence of aperiodicity parameter, α, on probabilistic ground motion and its deaggregation was also evaluated in their study. Maps for PGA and SA1 on rock for 10% probability of exceedance in a time period of 50 year were shown to compare the separate contribution of smoothed seismicity and fault components.

Akinci et al. investigated the differences in the results of earthquake-included ground-motion hazard using Poissonian and non-Poissonian models. It should be noted that (Pace, Peruzza et al. 2006) also developed PSHA for the same location with several differences in their data and assumption. Akinci et al. used historical catalogues working in (Cruppo di Lavoro 2004) while Pace et al. use CSIT catalogues (Augliera, Cattaneo et al. 2001). The other difference was that in Akinci et al. study, the faults with magnitude greater than 5.9 were considered whilst the ounces greater than 5.5 were considered by Pace et al. In spite of Pace et al. study, Akinci et al. estimated aperiodicity for three faults as a guide for finding the range of α and investigating its influence on PSHA. Fig. 2.3 illustrates the basic procedure for preparing hazard maps schematically.

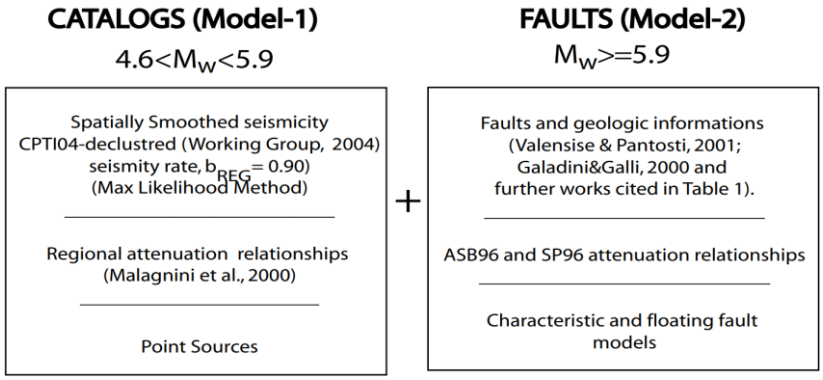



***Figure 2-3**. Scheme used to make hazard calculations for the central Apennines (Akinci et al. 2009).*

In their study, it was assumed that earthquake hazard is due to the background earthquake (seismicity of unknown faults, 4.6≤M<5.9) and geological data (M>5.9). Details of these two mentioned models can be found in (Akinci, Galadini et al. 2009).

Some studies have been conducted to find the value of α. (Ellsworth, Matthews et al. 1999) concluded that "1) the limited worldwide earthquake recurrence data have the α values of 0.46±0.32. 2) The 35-recurrence interval sequence examined are compatible with a shape factor

of 0.5 and 3) 35 earthquakes had no systematic difference when grouped by tectonic style." Fig. 2.4 shows the effect of α values on the time-dependent results.

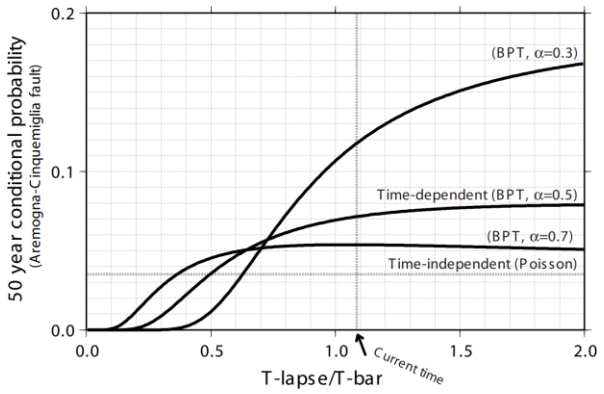



***Figure 2-4**. Graph showing 50 yr probability of the Aremogna–Cinquemiglia fault/earthquake occurrence as a function of elapsed time ratio. Curves are for Poisson model and BPT model with indicated α values (Ellsworth, Matthews et al. 1999).*

According to the study results, it could be said that generally "the smaller the α, the nearer the rise in hazard above the Poisson level that the average recurrence time occurs". Moreover, for the lowest values of α, elapsed time affects the time-dependent probabilities significantly.

Based on the obtained results, the probabilities increase by elapsing time except for faults with recently occurred earthquakes. Fig. 2.5 illustrates the map of probabilistic PGA with 10% probability of exceedance in 50 years using BPT and Poisson models.

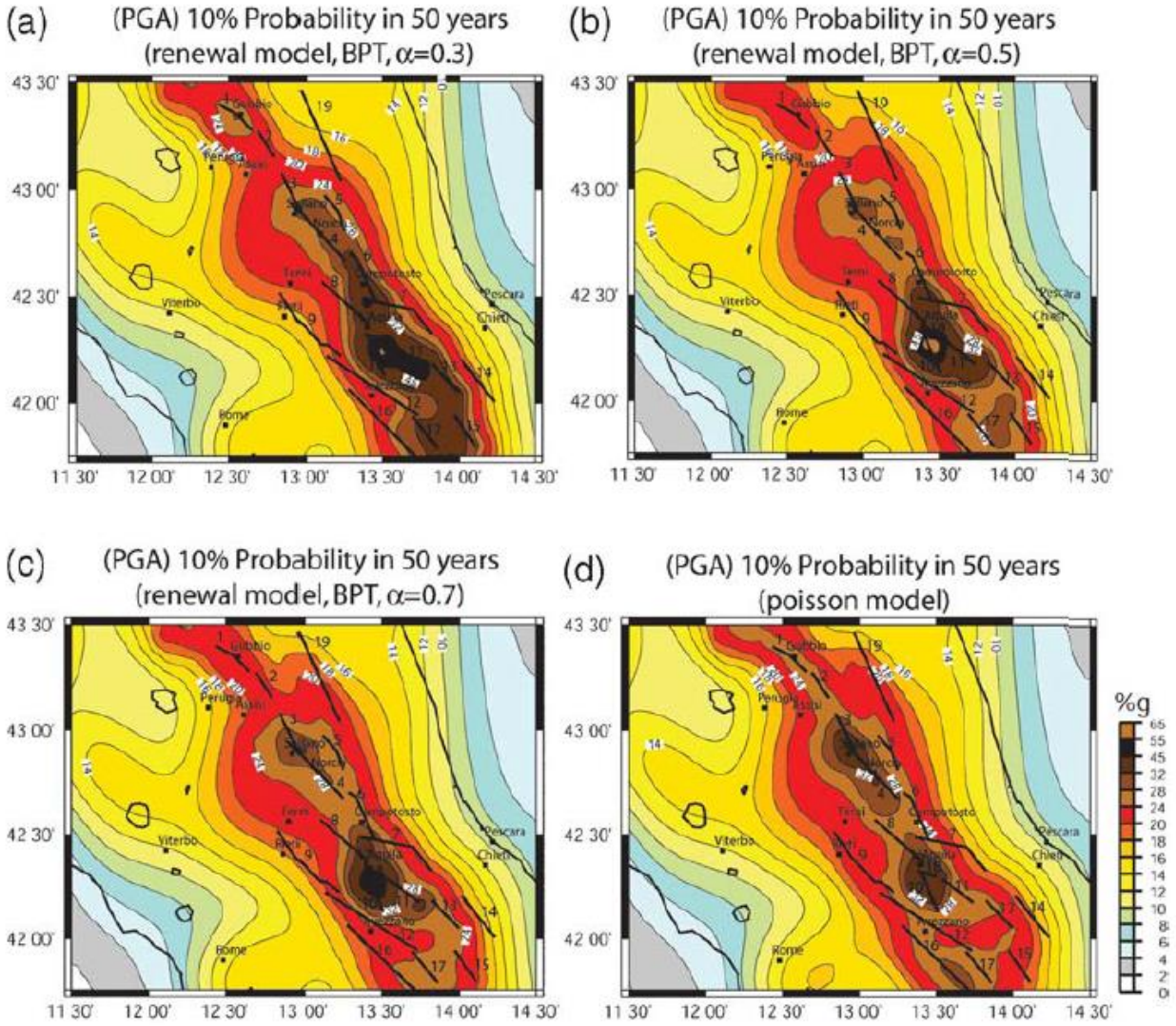



***Figure 2-5.*** *Maps of probabilistic PGA having 10% probability of exceedance in 50 years, derived from both gridded seismicity and faults BPT renewal model using the (a) α=0.3, (b) 0.5, (c) 0.7, and for (d) Poisson model (Ellsworth, Matthews et al. 1999, Zafarani and Ghafoori 2013, AlHamaydeh, Aly et al. 2019).*

Considering Fig. 2.5, following conclusions were drawn:

- Hazard increases continuously: one fault having elapsed time greater than the average recurrence interval dominates at the site.
- Hazard decreases and then levels off: the domination of a fault having elapsed time shorter than average recurrence interval decreases to the point where the background seismicity dominates.
- Hazard maintains a steady level: background seismicity always dominates.
- Hazard stays level and then increases: background seismicity loses domination to a fault with elapsed time longer than the average recurrence interval.

- Hazard stays level and then decreases: a fault having elapsed time shorter than average recurrence interval.
- Hazard decreases and then increases: initially a fault having recurrence time shorter than the average recurrence interval dominates but then loses domination to another fault having elapsed time longer than average recurrence interval.
- Hazard increases and then decreases: a dominating fault has elapsed time near 0.6 times the average recurrence time, a value where the probability of occurrence increases and then decreases as α goes from 0.7 to 0.5 to 0.3.

## 2.5 Panthi et al. (2011)

In Panthi et al. research (Panthi, Shanker et al. 2011), a model was proposed for earthquake occurrence in the region of northeast India. Based on their results, the time interval between two main shocks depends on the preceding main shock magnitude while it is independent of the following main shock. They suggested a linear relationship between the logarithmic of repeat time (T) of two events and the magnitude of the preceding main shock (Eq. 2.7).

$$Log\ T = cM_p + a \qquad (2.7)$$

Where c is the positive slope of line, a is the function of magnitude of the considered earthquake and $M_p$ is preceding main shock magnitude. c and a are assumed to be 0.21 and 0.35 in the region of the study (northeast India and its surrounding regions). They believed that their study outcomes can be considered for long term seismic hazard in the delineated seismogenic region.

It should be mentioned that their study was based on Raid's theory of the elastic rebound theory which assumes that an earthquake occurs when stress reaches a critical value in a fault of seismogenic source. Fig. 2.6 illustrates two values of stresses: $\tau_1$ and $\tau_2$, which affect the performance of a fault. The constant value of $\tau_1$ means that the model is time predictable (stress drop changes to different shocks). On the other hand, when the value of $\tau_2$ is constant, the model assumed to be slip predictable (earthquakes start a variable states of stress).

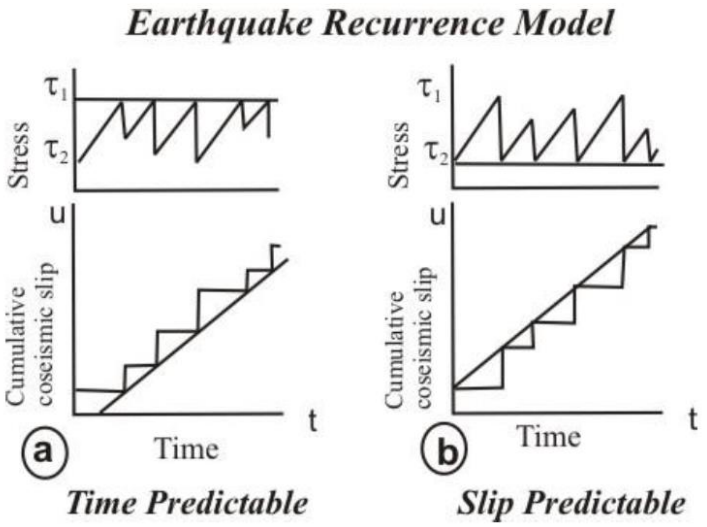



***Figure 2-6**. Earthquake Recurrence Model: **(a)** time-predictable model showing stress build-up to a certain value ($\tau_1$) and non-uniform stress drop; and **(b)** slip predictable-model illustrating non-uniform stress build-up and stress drop to a certain minimum value ($\tau_2$) (after Shimazaki and Nakata, 1980).*

## 2.6 Papaioannov and Papazachos (2000)

In this study (Papaioannou and Papazachos 2000), time-dependent and time-independent hazard for 144 broad sites of Greece were evaluated. Probability of occurrence of strong ground motion with macro seismic intensity I≥VII (in MM scale) in the period of 1996-2010 was presented using time-dependent hazard analysis.

They considered the results of previous studies (Byerlee and Brace 1968, Brune 1973) which is based on elastic rebound theory. According to this theory, when two sides of a fault move in opposite direction, they lock. After reaching a sufficient level of shear stress, slip occurs by releasing energy and then the fault locks again. The sequence of stress accumulation and release suggests that the time and magnitude of occurrence of an earthquake may not be stochastically independent.

The equation suggested by (Papazachos and Papazachou 2003) was considered in their study (Eq. 2.8) to assess seismic hazard:

$$Log\ T_t = 0.19M_{min} + 0.33M_p - 0.39\ log\ m_0 + 7.81 \qquad (2.8)$$

Where $T_t$ is the intervent time of the main shock of every source. $M_{min}$ is the magnitude of the smallest main shock considered ($M_{min}\geq 5.5$), $M_p$ is the magnitude of the proceeding main shock and $m_0$ is the moment rate. Time-dependent seismic hazard was assessed by a computer program and the probability of occurrence of a seismic motion with intensity I≥VII for each site was calculated and plotted accordingly at the corresponding sites in the map of Fig. 2.7.

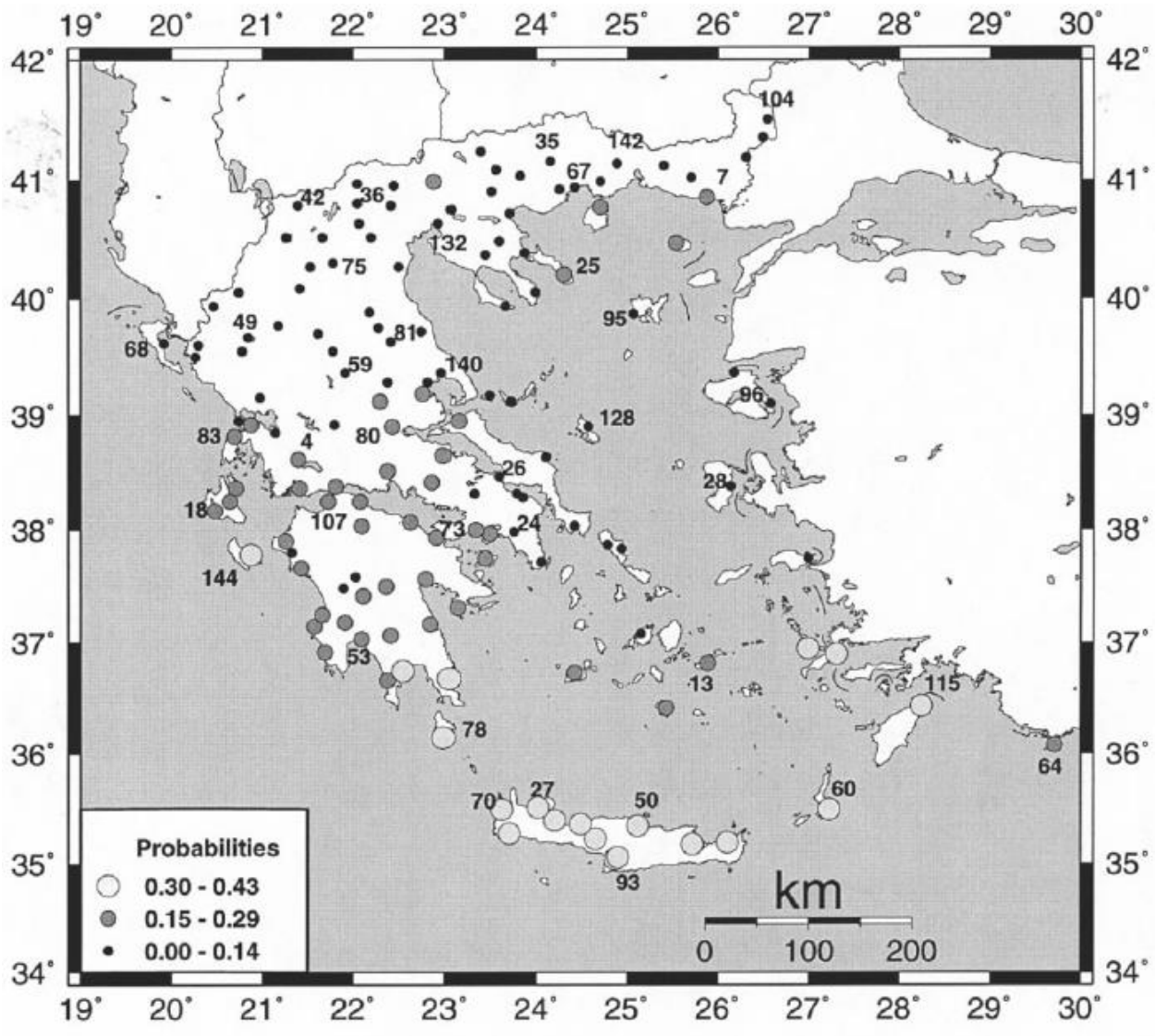



***Figure 2-7****. Probabilities for the occurrence of strong-ground motion (I ≥ VII) at each one of the 144 examined sites in the area of Greece during the period 1996–2010 (Papaioannou and Papazachos 2000).*

Based on the comparison made between assessed time dependent hazard with observed macro-seismic intensities of the period of 1950-1995, time-dependent models lead to reliable results.

## 2.7 Chat et al. (2013)

Long term and short term probabilistic seismic hazard by considering ground motion prediction equations for crustal and subduction earthquakes were assessed by (Chan, Wu et al. 2013). An appropriate approach for time-dependent probabilistic seismic hazard determination was presented using an updated earthquake catalogue. In their study, (Kagan and Knopoff 1978) model which is known as ETAS was considered. Based on the time-space Epidemic Type

Aftershock (ETAS), each earthquake is regarded as a main shock that may trigger a consequent earthquake. In this study an approach for the short-term PSHA was also proposed.

## 2.8 Matthews et al. (2002)

In this study (Matthews, Ellsworth et al. 2002), a model was proposed for rapture time on a recurrence earthquake source. The events interval was assumed to have a BPT distribution because of its noteworthy distribution:

- The probability of immediate rapture is zero.
- Between t=0 and mean recurrence time, hazard rate increases from zero to a finite maximum and then decreases to a quasi-stationary level, in which the conditional probability of event becomes time-independent.
- When coefficient of variation is less than, equal or greater than $1/\sqrt{2} \approx 0.707$, the quasi-stationary failure will be greater than, equal or less than mean failure rate.

Based on the elastic-rebound model, which was proposed first by (Lawson and Reid 1910), great tectonic earthquakes may reoccur at regular time intervals. (Hagiwara 1974) concluded that Gaussian distribution cannot be an appropriate choice for stochastic recurrence model since it assigns positive probability to negative intervals. He investigated using Gamma, lognormal and Weibull distributions and finally applied Weibull on grounds of “practical convenience”. And its popularity in “probabilistic quality control”. (Nishenko and Buland 1987) concluded that lognormal provided the best fit to the distribution of normalized intervals. They also concluded that coefficient of variation in lognormal is almost constant across sequence from different regions with different characteristic time scales.

(Matthews, Ellsworth et al. 2002) used the BPT model and compared it with other ones. Moreover, they assessed applying this model to the fault which their last rupture is unknown. Some models which are considered for probability distributions are shown in Fig. 2.8. Fig 2.9 also illustrates the BPT distribution and hazard rate for different values of α.

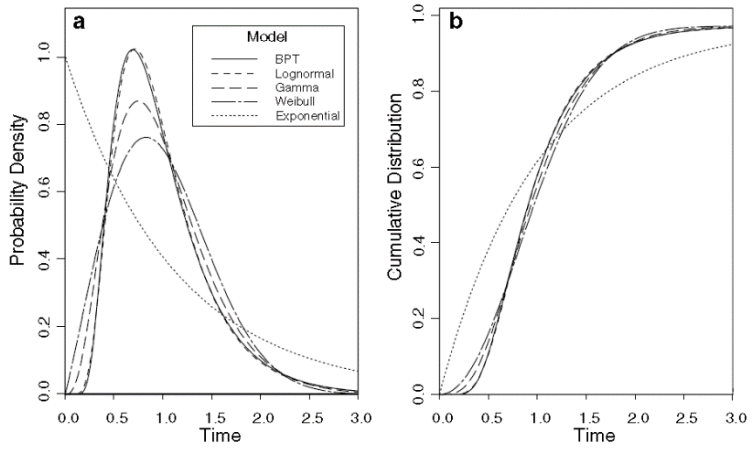



***Figure 2-8****. Probability density (a) and cumulative distribution (b) functions of exponential (Poisson), BPT, log–normal, gamma, and Weibull models. All distributions have mean 1 and standard deviation 0.5 (except the exponential distribution) (Matthews, Ellsworth et al. 2002).*

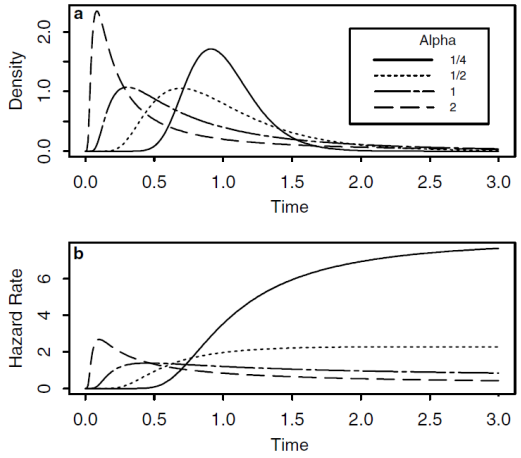



***Figure 2-9****. Probability functions for BPT (1, α), α= 1/4, 1/2, 1, 2: probability densities (a) and hazard rates (b) (Matthews, Ellsworth et al. 2002).*

As it could be seen in Fig. 2.9, small values of α lead to nearly symmetrical densities with pronounced control density near the mean value, while larger values of α result in highly skewed to the right densities which sharply peak at a value left of the mean. As shown in hazard rate diagram, the Brownian failure process reaches a quasi-stationary state in which residual time to failure becomes independent of elapsed time. Common models which are considered for earthquake rupture periodicity were compared by(Matthews, Ellsworth et al. 2002).

## 2.9 S. Hebden and Stein (2009)

Damage due to future earthquakes can be estimated by ground shaking hazard maps. Hazard and risk maps are prepared by incorporating earth science and engineering to estimate the

probability of occurrence rate of earthquake, ground shaking level and building response to a ground motion (Petersen, Cao et al. 2007). Using these maps, larger predicted motion will lead to higher predicted seismic hazard. Seismic hazard maps are also used to revise the codes which are utilized to design structures.

It should be kept in mind that in the locations where the recurrence rate of large damaging earthquake is low hazard estimation is rather complicated because of poor knowledge of required date.

Based on the previous studies, high hazard in central and eastern United States (CEUS) is due to four main assumptions: the first one is the earthquake magnitude which is assumed for the future event (known as characteristic earthquake). Second one is the relation between ground acceleration at a given distance for an earthquake at a given size (GMPE). The third factor is the time window or probability level chosen to define the hazard. 2% probability of exceedance in 50 year or once in 2500 years was considered by (Frankel, Mueller et al. 1996, Frankel, Petersen et al. 2002) to illustrate the hazard as the maximum acceleration predicted at a geographic point. This consideration leads to the much higher hazard in comparison to the former assumption (1% in 50 years or one in 500 years by (Algermissen, Perkins et al. 1982)) and is because of lack of knowledge of large earthquakes. The fourth factor is the recurrence interval of large earthquakes.

Fig. 2.10 compares time-dependent and time-independent models for earthquake recurrence.

According to this study, when the time elapsed since the last earthquake is less than approximately 2/3 of the assumed mean recurrence interval, time-dependent models predict lower probabilities.

The effect of the model chosen depends on the ratio of elapsed time since the last earthquake to the assumed mean recurrence time and the assumed probability distribution and variability of the recurrence time.

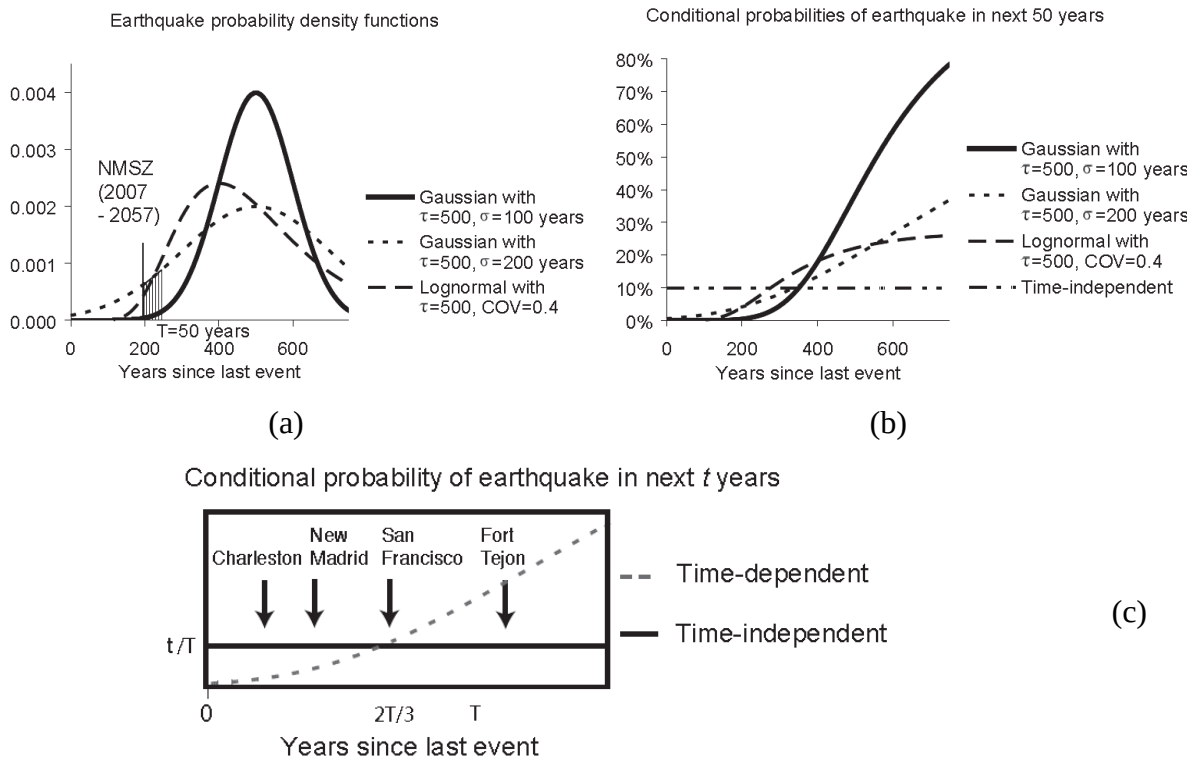



***Figure 2-10.*** *Comparison of time-dependent and time-independent models for earthquake recurrence. (a): Representative probability density functions for the distribution of recurrence times of characteristic earthquakes in the New Madrid zone. Recurrence times are described by Gaussian distributions with a mean of 500 years and a standard deviation of either 100 or 200 years, or a lognormal distribution with a similar mean and coefficient of variation. Time zero corresponds to the date of the past major earthquake in 1811. (b): Comparison of the conditional probability of a large earthquake in the New Madrid zone in the next 50 years, assuming that the mean recurrence time is 500 years. In the time-independent model the probability is always 10%. In the time-dependent models ((a)) the probability is small shortly after the past one and then increases with time. Because the time since 1811 is less than 2/3 of the assumed mean recurrence interval, these models predict lower probabilities of a large earthquake in the next 50 years at present and for the next hundred years. (c): Schematic comparison of time-independent and time-independent models for different seismic zones. Charleston and New Madrid are "early" in their cycles, so time-dependent models predict lower hazards. The two model types predict essentially the same hazard for a recurrence of the 1906 San Francisco earthquake, and time-dependent models predict higher hazard for the nominally "overdue" recurrence of the 1857 Fort Tejon earthquake. The time-dependent curve is schematic because its shape depends on the probability distribution and its parameters (Hebden and Stein 2009).*

In this study, hazard maps prepared using both time-dependent and time-independent models were compared. Parameters including assumed maximum magnitude of the largest earthquake, GMPE model, and probability level were considered the same to compare only the effect of different models. Based on the results, the time-dependent model predicted considerably lower hazard for the 50-year period. Fig. 2.11 shows the effect of using time-dependent and time-independent models. Fig. 2.12 also compares the results of using two models.

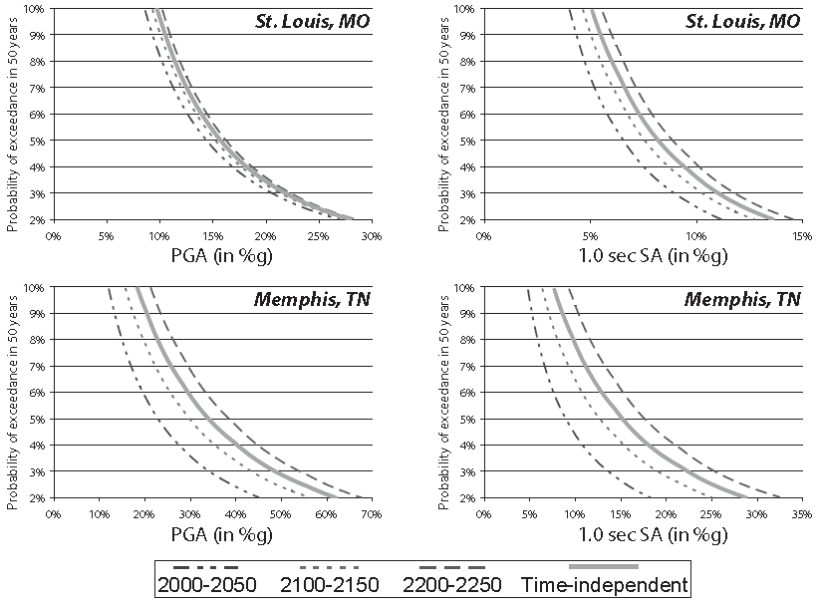



***Figure 2-11**. Comparison of predicted hazard values for Memphis and St. Louis for a range of return periods or, equivalently, probabilities of exceedance, for time-independent and time-dependent models. These are shown for peak ground acceleration in the left column and for acceleration with a period of 1 s in the right column (Hebden and Stein 2009).*

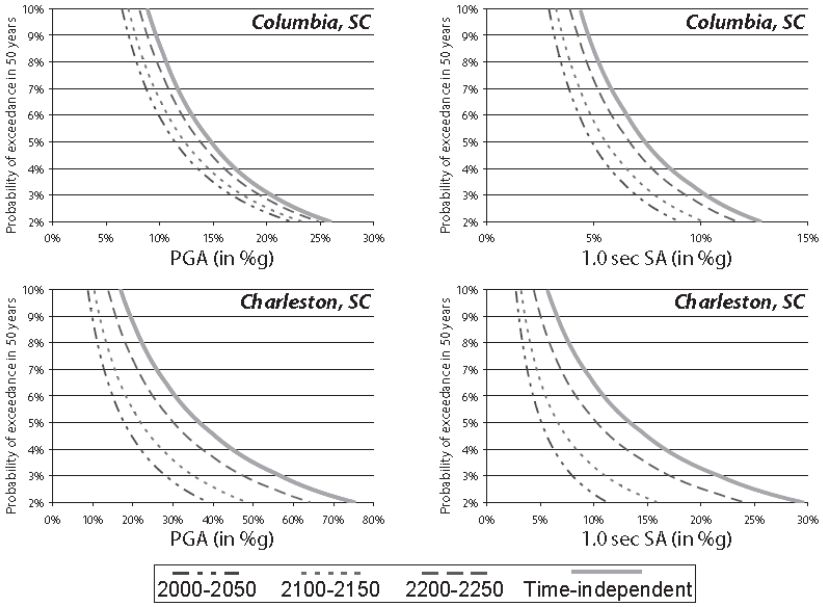



***Figure 2-12**. Comparison of predicted hazard values for Charleston and Columbia for a range of return periods or, equivalently, probabilities of exceedance, for time-independent and time-dependent models. These are shown for peak ground acceleration in the left column and for acceleration with a period of 1 s in the right column (Hebden and Stein 2009).*

As it can be seen in Fig. 2.12, time-dependent model predicts lower hazard at present and in the future time period shown.

## 2.10 Boyd et al. (2008)

(Boyd, Zeng et al. 2008) presented a time-dependent seismic hazard analysis for Alaska and the Aleutians. It is noteworthy that time-independent maps consider all sources statistically independent. On the other hand, in time-dependent analysis, BPT model is used to calculate conditional probability of occurrence for the next 50 years.

Based on their study, two notable events can change time-dependent probabilistic seismic hazard estimation: occurring earthquake on characteristic fault segments and stress changes on the fault due to regional earthquake. It should be taken into account that significance of stress transfer depends on the location, orientation and sense of slip-on target faults. Previous studies showed that changing co-seismic stress may have a lasting influence on probability. Others also suggest that transferring visco-elastic stress transfer can play a significant role over the long term. Based on BPT, changing co-seismic stress cannot affect the earthquake probability for a long term and continuing post-seismic relaxation should be considered in earthquake triggered models.

Based on this research, smaller and more frequent earthquakes in the Gutenberg-Richer component of a seismic hazard model have a more contribution to seismic hazard than an equal contribution from a characteristic component. They concluded that when a time-dependent model without stress changing is considered, the probability of earthquake occurrence alters notably (decreasing to approximately zero or increasing to several times the value of time-independent ones). Moreover, co-seismic stress changes can have a local influence on earthquake probabilities, while post seismic effects can be far-reaching in both time and space, finally, combining time-dependent and time-independent sources, the seismic hazard does not change considerably.

## 2.11 Petersen et al. (2007)

The hazard map of peak ground acceleration for 10% probability of exceedance in 30 years from the 2002 national seismic hazard models obtained considering time-independent and time-dependent PSHA are shown in Fig. 2.13 and Fig. 2.14, respectively. Fig. 2.15 is also provided to show the difference between Fig 2.13 and Fig. 2.14 (a ratio map) (Petersen, Cao et al. 2007).

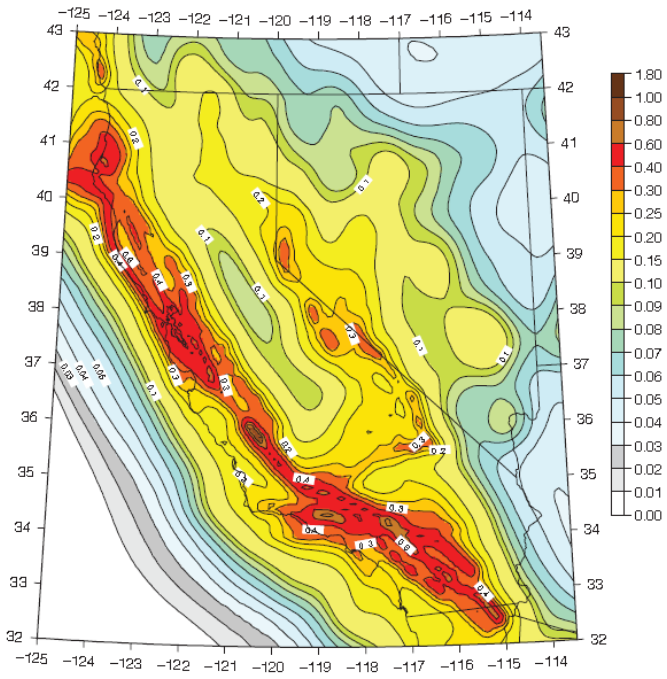



***Figure 2-13**. Time-independent (Poisson) map for rock site condition and a 10% probability of exceedance in 30 years. This map was developed from the 2002 national seismic hazard model but also includes the new Poisson model for T-I Model 3 (Petersen, Cao et al. 2007).*

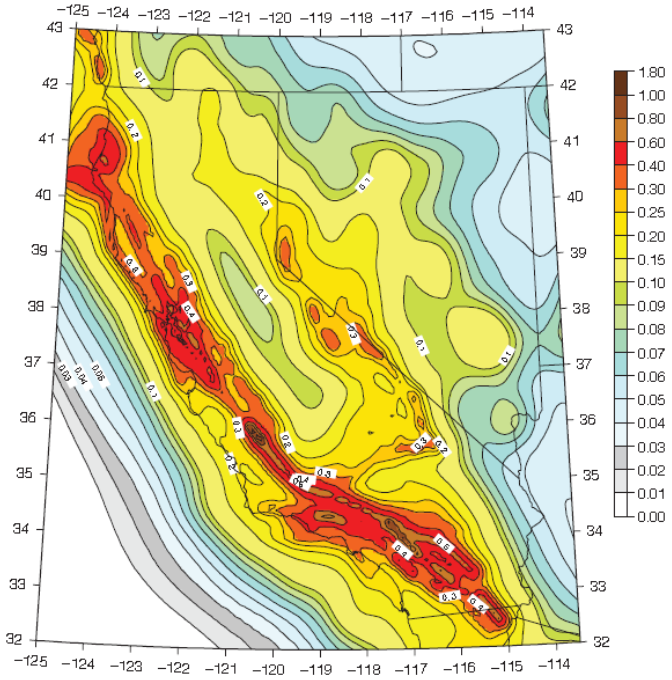



***Figure 2-14**. Time-dependent map for rock site condition and a 10% probability of exceedance in 30 years. This map was developed by equally weighting three time-dependent models (T-D model 1, 2, and 3) (Petersen, Cao et al. 2007).*

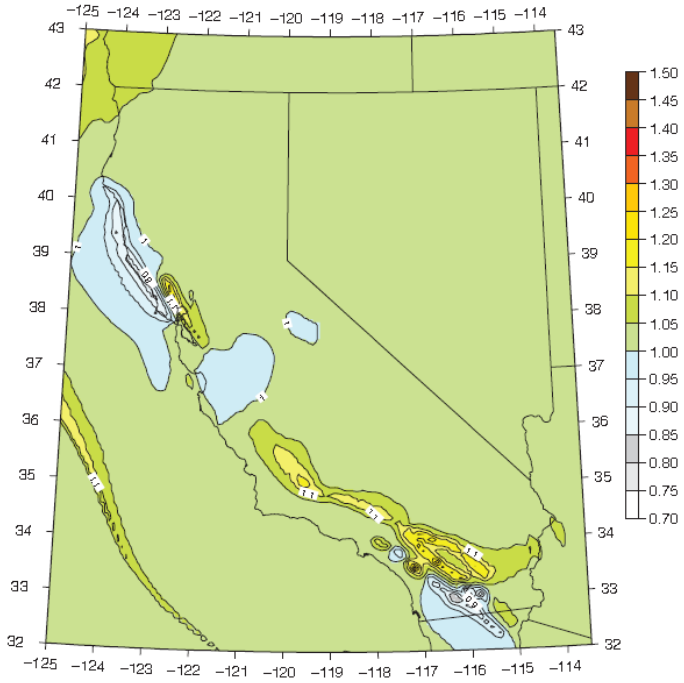



***Figure 2-15**. Ratio of the time-dependent map (figure 2-13) and the time-independent map (figure 2-14) for rock site conditions and a 10% probability of exceedance in 30 years (Petersen, Cao et al. 2007).*

Time-dependent and time-independent hazard maps (PGA) for California with a 10% probability of exceedance for a time period of 30 years was presented by Petersen et al. Based on the results of their research, time-dependent maps differ by approximately 10-15% from the time-independent map for a site which is located near a fault. For the sites which are located well away from the time-dependent sources the maps were almost the same. Moreover, for the sites, where a long time elapsed since the last earthquake, time-dependent hazard was considerably more than time-independent ones. Finally, for the sites where a short time elapsed since the last event, time-dependent hazard was lower than time-independent hazard (sites close to faults in the first half of their seismic life).

## 2.12 Douglas et al. (2013)

(Petersen, Cao et al. 2007) tested risk-targeting procedure of (Luco, Ellingwood et al. 2007) for mainland France using the hazard model presented by (Martin, Combes et al. 2002) and the PGA map to define the demand spectrum of the French seismic design code. They compared the current PGA with the outcomes of risk-targeted approach. It should be noted that their study focused on loading of new structures not retrofitting or risk analysis of existing buildings.

Two performance level mentioned in (Eurocode 8 2004) are “no-collapse requirement” and “damage-limitation requirement”: “no-collapse” (the structure withstand the seismic design action without collapse, retains its integrity and have residual load-bearing capacity after the

earthquake) should be considered for an event with a 10% probability of exceedance in 50 years (475-year return period) and "damage limitation" (structure withstand a more frequent seismic action without damage or limitation of use) should be met for an event with 50% probability of exceedance in 50 years (95-year return period).

Seismic hazard curves and fragility curves are convolved to calculate seismic risk, as given in Eq. 2.9 and Eq. 2.10 (Kennedy 2011):

$$y(a_0) = -\int_0^{\infty} P_{a0}(a)\frac{dH(a)}{da}da \tag{2.9}$$

$$y(a_0) = \int_0^{\infty} H(a)\frac{DP_{a0}(a)}{da}da \tag{2.10}$$

Where $P_{a0}(a)$ is the fragility curve (conditional probability of failure given the ground motion level a for a design-level of $a_0$ and H(a) is the seismic hazard curve (the exceedance frequency corresponding to ground motion level a). It is noteworthy that derivation of hazard curves is less accurate than derivation of fragility curves. The main reason is that hazard is obtained in PSHA for a small number of ground-motion whilst fragility curves are obtained by analytical equations. Lognormal distribution is commonly used for obtaining fragility curves and is mainly based on two parameters: mean, μ, and standard deviation, β. ( $P(a) = \phi\{[Ln(a) - Ln(\mu)]/\beta\}$).

It should be noticed that the whole hazard curve is considered by risk integral (not the design ground motion on the PGA). As a result, the slope of hazard curve affects the results considerably. It means that the relative ranking in design PGA changes just when hazard curve slopes vary.

It should be taken into account that the fragility curves should be based on the results of structure analysis using (Eurocode 8 2004).

As it was mentioned before, the value of β can affect the results. (Luco, Ellingwood et al. 2007) considered β=0.8 while chapter 21 of (ASCE 7-10 2013) standard 7-10 uses β=0.6 for its site-specific ground motion procedure for seismic design. (Crowley, Colombi et al. 2011) also reported the value of β around 0.5. (Douglas, Ulrich et al. 2013) considered 0.5 for this parameter.

Based on the conclusions drawn by (Douglas, Ulrich et al. 2013), choosing high values for β and low probability of collapse leads to unrealistic fragility curves because it predicts low collapse probability even for very high shaking levels.

They concluded that by choosing reasonable values for the free input parameters, the obtained maps will be approximately similar with the maps which are currently used. In brief, the risk-targeted method is strongly depended on the input parameters and it could be used if the input parameters adopted reasonably.

## 2.13 P. Kennedy (2011)

ASCE standard 43-05 approach is used to define the risk-consistent for site specific response spectrum (SSRS). (Kennedy 2011) studied the assumptions made in the ASCE Standard 43-05 by amplifying the commentary of (ASCE 2005).

It should be kept in mind that when Uniform Risk Response Spectrums are considered as the SSRSs, plants at different sites should have consistent seismic risks. While, risk-consistency goal is not achieved in case of using uniform hazard response spectrum (UHRS); the UHRS does not consider that different slopes, which are critical to obtaining risk-consistent seismic designs, of the seismic hazard curves is due to different sites.

Based on ASCE Standard 43-05 (ASCE 2005), the seismic demand and structural capacity are aimed at having sufficient conservatism to reasonably achieve both of the following:

"1. less than about a 1% probability of unacceptable performance for the design basis earthquake ground motion and

2. less than about a 10% probability of unacceptable performance for a ground motion equal to 150% of the design basis earthquake ground motion."

Seismic fragility, which is defined as the conditional probability of failure versus ground motion levels, typically has a lognormal distribution. It is mainly described by two parameters, such as a seismic margin factor, $F_P$, corresponding to a conditional probability of failure, $P_{FC}$, and an estimate of the capacity variability (i.e., the logarithmic standard deviation, β).

ASCE Standard 43-05 target levels of conservatism result in the seismic margin factors $F_{1\%}$, $F_{5\%}$, $F_{10\%}$, $F_{50\%}$, and $F_{70\%}$ for a 1%, 5%, 10%, 50%, and 70% conditional probability of unacceptable behaviour, respectively, which are provided in Table 3. When β is less than 0.39,

the second of the two conditional probability goals control the fragility, whilst, for β greater than 0.39, the first goal controls. By specifying both goals, the following margins are achieved:

- $F_{1\%} \geq 1.0$.
- $F_{5\%} \geq 1.3$.
- $F_{10\%} \geq 1.5$.
- $F_{50\%}$ increases with increasing β.

According to EPRI (1994) and past performed seismic probabilistic risk assessments (SPRA) studies, β value for structures and major passive mechanical components located on the ground or at low elevations within structures, is in the range of 0.3 -0.5. While, it is 0.4–0.6 for active components mounted at high elevations in structures. Therefore, the range 0.3–0.6 covers the practical range for β.

These required damage frequency (DF) factors are compared with ASCE Standard 43-05 DF given by Eq. (6). The required DF were calculated using Eq. 2.11-2.14. The range of $A_P$ were chosen 1.5-6.0. The calculated DF were compared with the DF values according to ASCE Standard 43-05.

$$DF = \frac{[R_P e^{-f}]^{1/K_H}}{F_{P\%}} \tag{2.11}$$

$$f = X_P K_H \beta - \frac{1}{2}(K_H \beta)^2 \tag{2.12}$$

$$K_H = \frac{1}{\log(A_R)} \tag{2.13}$$

$$A_R = \frac{SA_{0.1H}}{SA_H} \tag{2.14}$$

Where $X_P = 2.326$ , $SA_{0.1H}$ is the spectral acceleration at the mean exceedance fragility H ( $H = mean 1\times10^{-4}/yr$ ) and $SA_{0.1H}$ is the spectral acceleration at 0.1H.

$$DF = Maximum(DF_1, DF_2) \tag{2.15}$$

$$DF_1 = 1.0$$

$$DF_2 = 0.6(A_R)^{0.80} \quad (2.16)$$

According to (Kennedy 2011) study (Table 2.3), choosing the value of β equal to 0.4 and 0.5 lead to more reliable results than those for β of 0.3 and 0.6. In their study, the fragility β values were approximately in the 0.4–0.5 range while β of 0.3 and 0.6 were extreme low and high, respectively.

***Table 2-3**. Design factor required to achieve a probability ratio 10 (Kennedy 2011).*

| $A_R$ | *DF* | | | | *DF Eq. (6)* |
|---|---|---|---|---|---|
| | $F_{1\%}$=1.1 B=0.30 | $F_{1\%}$=1.0 B=0.40 | $F_{1\%}$=1.0 B=0.50 | $F_{1\%}$=1.0 B=0.60 | |
| 1.50 | 0.88 | 0.93 | 0.95 | 1.03 | 1.00 |
| 1.75 | 0.96 | 0.96 | 0.91 | 0.91 | 1.00 |
| 2.00 | 1.05 | 1.03 | 0.95 | 0.90 | 1.04 |
| 2.25 | 1.16 | 1.11 | 1.00 | 0.93 | 1.15 |
| 2.50 | 1.27 | 1.21 | 1.07 | 0.97 | 1.25 |
| 2.75 | 1.38 | 1.30 | 1.14 | 1.03 | 1.35 |
| 3.00 | 1.50 | 1.40 | 1.22 | 1.08 | 1.44 |
| 3.25 | 1.61 | 1.50 | 1.30 | 1.14 | 1.54 |
| 3.50 | 1.73 | 1.60 | 1.38 | 1.21 | 1.63 |
| 3.75 | 1.84 | 1.70 | 1.46 | 1.27 | 1.73 |
| 4.00 | 1.96 | 1.80 | 1.54 | 1.34 | 1.82 |
| 4.25 | 2.07 | 1.90 | 1.62 | 1.40 | 1.91 |
| 4.50 | 2.19 | 2.01 | 1.70 | 1.47 | 2.00 |
| 4.75 | 2.30 | 2.11 | 1.79 | 1.54 | 2.09 |
| 5.00 | 2.42 | 2.21 | 1.87 | 1.60 | 2.17 |
| 5.25 | 2.54 | 2.31 | 1.95 | 1.67 | 2.26 |
| 5.50 | 2.65 | 2.42 | 2.04 | 1.74 | 2.35 |
| 5.75 | 2.77 | 2.52 | 2.12 | 1.80 | 2.43 |
| 6.00 | 2.88 | 2.62 | 2.20 | 1.87 | 2.52 |

## 2.14 Martins et al. (2015)

One of the uncertainties in structure collapse, which is considered in designing Codes, is the building's ultimate capacity and is defined as the maximum ground motion intensity. It is, therefore, considered as a random variable. As an explanation, if the ground motion level, a, exceeds the structure capacity, $a_c$, the structure will collapse. Structure capacity is modelled as a lognormal distribution, which is defined by two parameters: logarithmic standard deviation, β, and the median, $\hat{a}_c$, (for which $P[a_c<\hat{a}_c]=0.5$).

(Martins, Silva et al. 2015) investigated the effect of structural fragility on the risk-targeted hazard assessment. They analysed 3-story and 5-story RC structures to drive fragility curves in accordance with Eurocode 8 (Eurocode 8 2004) using both of the fundamental period of vibration, $S_a$, and peak ground acceleration, PGA.

Figure 2.16 and Figure 2.17 demonstrate the obtained fragility curves using $S_a(T_1)$ and PGA, respectively, as ground motion intensity measure (IM).

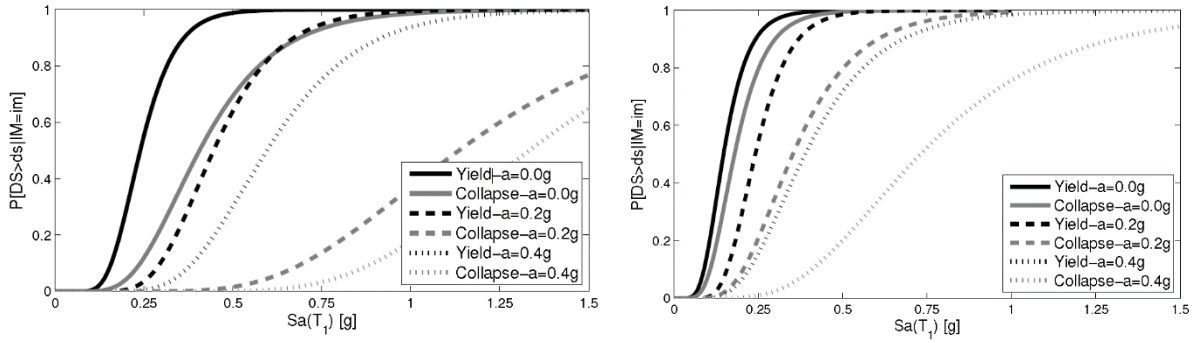



***Figure 2-16**. Case study fragility curves: Right: 3 storey structures; Left: 5 storey structures. (IM=$S_a(T_1)$) (Martins, Silva et al. 2015).*

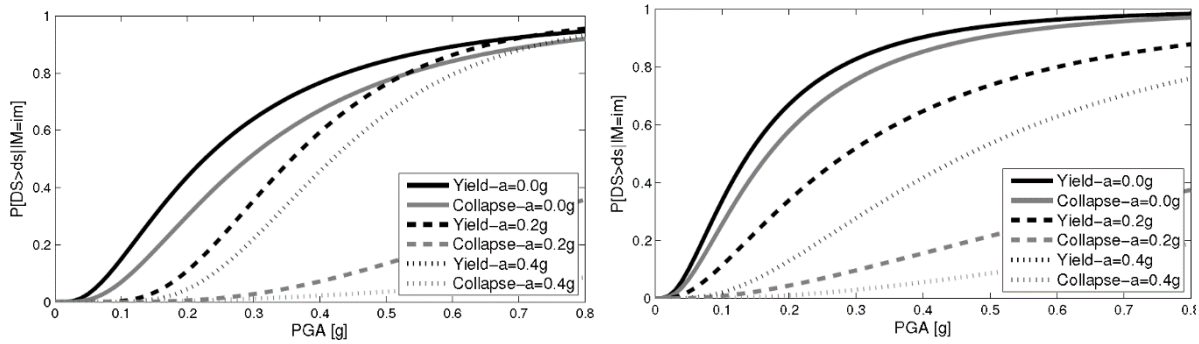



***Figure 2-17**. Case study fragility curves Right: 3 storey structures; Left: 5 storey structures. (IM=PGA) (Martins, Silva et al. 2015).*

In the previous studies, values of β were chosen between 0.5-1 (Luco, Ellingwood et al. 2007, Douglas, Ulrich et al. 2013, Ulrich, Negulescu et al. 2014, Silva, Crowley et al. 2016). (Douglas, Ulrich et al. 2013) considered β=0.5 as an appropriate lower boundary fragility curve, whilst, (Luco, Ellingwood et al. 2007) proposed values in the order of 0.8. (Martins, Silva et al. 2015) Resulted that if $S_a$ is used as IM, the values of β will be in a range of approximately 0.35-0.45. While it will be up to 0.8 if PGA is considered. This is because PGA is a less efficient IM than $S_a(T_1)$ in predicting maximum inter-storey drift.

It should be also noted that the fragility curves obtained by considering PGA as IM, the β values are within the mentioned interval but in other curves the β values were lower than the minimum suggested by the literature.

The other notable conclusion of their study was that the probability of collapse at the design ground motion, $a_{des}$, for buildings designed according to the newest regulation should be around $10^{-3}$ to $10^{-2}$ (similar to the results of (Luco, Ellingwood et al. 2007)). (Ulrich, Negulescu et al. 2014), on the other hand, reached the values ranging from $10^{-7}$ to $10^{-5}$ for the probability of collapse at the design ground motion.

## 2.15 Gkimprixis et al. (2019)

A new consistent terminology method for risk-targeting approach was developed by Gkimprixi et al. (2019). It should be explained that risk targeting has recently emerged as a tool for overcoming the limitations of current design code in which the structures are designed based on a uniform-hazard spectrum for a fixed return period. By implementing risk-targeting, consistent performance levels for structures with different properties through the definition of uniform-risk design maps would be obtained.

The authors assessed the effect of the linearization of the hazard curve in their study. Then, risk-targeted behaviour factors (RTBF) method was verified by comparing the obtained results with those of uniform-risk design spectral accelerations for single-degree-of-freedom systems with elastic-perfectly plastic behaviour for two different sites. Eventually, RTBF was utilized for developing uniform-risk design maps for Europe. The presented maps illustrated the differences of seismic design levels due to difference between uniform-hazard and uniform-risk concept.

## 2.16 Gkimprixis et al. (2020)

Gkimprixi et al. (2019) investigated alternative approaches for the seismic design of structures in a further study. The assessed the safety level and cost associated to different design approaches: (i) the design used in the current codes and is based on uniform-hazard response spectra, reduced to account for inelastic structural behaviour, (ii) a risk-targeting approach which aims at designing structures with the same risk of collapse throughout regions of different seismicity, and (iii) minimum-cost approach which aims to minimize the sum of the initial construction cost and the cost of expected losses caused by future earthquakes. The comparison of the above-mentioned methods was made through a 4-stor reinforced concrete structure located in different locations in Europe.

# References

Akinci, A., et al. (2009). "Effect of time dependence on probabilistic seismic-hazard maps and deaggregation for the Central Apennines, Italy." Bulletin of the Seismological Society of America **99**(2A): 585-610.

Algermissen, S. T., et al. (1982). Probabilistic estimates of maximum acceleration and velocity in rock in the contiguous United States, US Geological Survey.

ASCE 7-10, A. S. A. S. E. I. (2013). Minimum design loads and associated criteria for buildings and other structures

ASCE, A. S. o. C. E. (2005). Seismic Design Criteria for Structures, Systems and Components in Nuclear Facilities. . ASCE Standard 43-05.

Augliera, P., et al. (2001). "Catalogo strumentale dei terremoti'italiani'dal 1981 al 1996, Versione 1.0 (5.3. 2001)."

Boyd, O. S., et al. (2008). "Toward a time-dependent probabilistic seismic hazard analysis for Alaska." Active Tectonics and Seismic Potential of Alaska, Geophysical Monograph **179**: 399-416.

Brune, J. N. (1973). "Earthquake modeling by stick-slip along precut surfaces in stressed foam rubber." Bulletin of the Seismological Society of America **63**(6-1): 2105-2119.

Byerlee, J. D. and W. Brace (1968). "Stick slip, stable sliding, and earthquakes—effect of rock type, pressure, strain rate, and stiffness." Journal of Geophysical Research **73**(18): 6031-6037.

Chan, C.-H., et al. (2013). "Time-dependent probabilistic seismic hazard assessment and its application to Hualien City, Taiwan." Natural Hazards and Earth System Sciences **13**(5): 1143-1158.

Crowley, H., et al. (2011). "Fragility functions for common RC building types in Europe." Deliverable D3 **1**: 223.

Cruppo di Lavoro, C. (2004). "Catalogo Parametrico dei Terremoti Italiani, versione 2004 (CPTI04)." http://emidius. mi. ingv. it/CPTI/.

Douglas, J., et al. (2013). "Risk-targeted seismic design maps for mainland France." Natural Hazards **65**(3): 1999-2013.

Ellsworth, W. L., et al. (1999). "A physically-based earthquake recurrence model for estimation of long-term earthquake probabilities." US Geological Survey Open-File Report **99**(522): 22.

Eurocode 8, E. S. E.-. (2004). Eurocode 8. Design of structures for earthquake resistance - Part 1: General rules, seismic actions and rules for buildings.

Frankel, A. D., et al. (1996). National seismic-hazard maps: documentation June 1996, US Geological Survey Reston, VA.

Frankel, A. D., et al. (2002). "Documentation for the 2002 update of the national seismic hazard maps." US Geological Survey Open-File Report **2**(420): 33.

Hagiwara, Y. (1974). "Probability of earthquake occurrence as obtained from a Weibull distribution analysis of crustal strain." Tectonophysics **23**(3): 313-318.

Hebden, J. S. and S. Stein (2009). "Time-dependent seismic hazard maps for the New Madrid seismic zone and Charleston, South Carolina, areas." Seismological Research Letters **80**(1): 12-20.

Jalalalhosseini, S. M., et al. (2018). "Time-dependent seismic hazard analysis for the Greater Tehran and surrounding areas." Journal of Seismology **22**(1): 187-215.

Kagan, Y. and L. Knopoff (1978). "Statistical study of the occurrence of shallow earthquakes." Geophysical Journal International **55**(1): 67-86.

Kennedy, R. P. (2011). "Performance-goal based (risk informed) approach for establishing the SSE site specific response spectrum for future nuclear power plants." Nuclear Engineering and Design **241**(3): 648-656.

Lawson, A. C. and H. F. Reid (1910). The California Earthquake of April 18, 1906: Report of the State Earthquake Investigation Commission, Carnegie institution of Washington.

Luco, N., et al. (2007). "Risk-targeted versus current seismic design maps for the conterminous United States."

Martin, C., et al. (2002). Revision du zonage sismique de la France: Etude probabiliste. Technical report. Rapport de Phase 3, GEO-TER, France, report GTR/MATE/0701-150, Affaire no. 1601, in French.

Martins, L., et al. (2015). Investigation of structural fragility for risk-targeted hazard assessment. 12th international conference on applications of statistics and probability in civil engineering, ICASP12, Vancouver, Canada.

Matthews, M. V., et al. (2002). "A Brownian model for recurrent earthquakes." Bulletin of the Seismological Society of America **92**(6): 2233-2250.

Mousavi, M. and M. Salehi (2018). "Temporal distribution of earthquakes using renewal process in the Dasht-e-Bayaz region." Journal of Seismology **22**(1): 153-159.

Nishenko, S. P. and R. Buland (1987). "A generic recurrence interval distribution for earthquake forecasting." Bulletin of the Seismological Society of America **77**(4): 1382-1399.

Pace, B., et al. (2006). "Layered seismogenic source model and probabilistic seismic-hazard analyses in central Italy." Bulletin of the Seismological Society of America **96**(1): 107-132.

Panthi, A., et al. (2011). "Time-predictable model applicability for earthquake occurrence in northeast India and vicinity." Natural Hazards and Earth System Sciences **11**(3): 993-1002.

Papaioannou, C. A. and B. Papazachos (2000). "Time-independent and time-dependent seismic hazard in Greece based on seismogenic sources." Bulletin of the Seismological Society of America **90**(1): 22-33.

Papazachos, B. and C. Papazachou (2003). The earthquakes of Greece. Ziti publications.

Petersen, M. D., et al. (2007). "Time-independent and time-dependent seismic hazard assessment for the State of California: Uniform California Earthquake Rupture Forecast Model 1.0." Seismological Research Letters **78**(1): 99-109.

Silva, V., et al. (2016). "Exploring risk-targeted hazard maps for Europe." Earthquake Spectra **32**(2): 1165-1186.

Ulrich, T., et al. (2014). "Fragility curves for risk-targeted seismic design maps." Bulletin of Earthquake Engineering **12**(4): 1479-1491.

Utsu, T. (1984). "Estimation of parameters for recurrence models of earthquakes." Bulletin of the Earthquake Research Institute **59**: 53-55.

Zafarani, H. and S. Ghafoori (2013). "Probabilistic assessment of strong earthquake recurrence in the Iranian plateau." Journal of Earthquake Engineering **17**(3): 449-467.

# 3 Influence of time-dependent seismic hazard on structural design

## ABSTRACT

Recent works on seismic hazard have introduced the concept of time-dependent seismic hazard and different models have been proposed to predict the inter-arrival time between consecutive events. Currently, the reliability assessment of structures and relevant design rules proposed by the codes are based on the Poisson recursive model, for which the frequency of the occurrence of seismic events does not change over time. This study presents results on the impact of Time-Dependent Seismic Hazard on structural design, by evaluating the strength required by the structure (seismic capacity) for different time intervals elapsing from the last event. "Seismic capacity" is understood here as the capacity required to provide a fixed reliability level, measured by the failure rate. Two different seismic scenarios have been investigated and results concerning different site-to-source distance, capacity dispersion of the structure and different recurrence properties of the time-dependent source have been discussed. Finally, the impact of recursive properties of time-dependent model is analysed and discussed. The results obtained from the analyses highlight a significant influence of time-dependent hazard properties on the structural capacity required to attain a target reliability, and give evidence to the different roles played by the parameters considered in the analysis. Within the set of the considered case studies, the ratios between seismic capacities evaluated by the time-dependent and non-time-dependent model span the range [0,1.32] in the first scenario, where inter-arrival time varies from 0 to two times the mean return period. The second scenario involves multiple sources and observed ratios were in the range [0.84,1.23], extreme values are relevant to inter-arrival times equal to 139y and 371y, respectively.

## 3.1 Introduction

Risk assessment involves analysis of the level of earthquake hazard, building vulnerability and exposure; it is based on a prediction of possible hazardous events, in terms of recurrence in time, and is oriented to the estimation of potential consequences, considering different sources of uncertainty. In this section, attention is focused on the probability of construction failure (consequence) due to an earthquake (event).

Generally, this type of risk analysis is developed within the context of the PEER framework (Porter 2003; FIB 2012) where the seismic hazard assessment is typically based on a constant rate of occurrence in time (constant-rate seismic hazard model), described by the Poisson recursive model, and involves potential sources with different locations and different intensities.

However, it is observed that small and medium magnitude events generally show different occurrence properties with respect to large magnitude events. The former generally occur as independent events, for which the recursive Poisson model is adequate, while the occurrence of the latter events is notably influenced by the previous history of the source activity. In this case, earthquakes tend to show a periodic trend and the fault activity generates earthquakes with similar magnitudes, also denoted as characteristic earthquakes (Schwarz et al. 1984; Wesnousky et al. 1994, Kramer 1996; Tondi and Cello 2003).

From a theoretical point of view, approaches considering the recursive properties of strong events and models providing a time-dependent prediction of the interarrival time passing between two events, as noted by Anagnos and Kiremidjian (1988), dates back to earlier than the 70s (e.g., Vere-Jones, and Davies 1966, Kameda and Ozaki 1979, Hagiwara 1974). In addition, in their work, an overview of different models and a proposal for their classification is presented.

Only recently have time-dependent models found practical applications thanks to the improvements in fault mechanism knowledge in some earthquake prone areas. Some studies, mainly limited to an assessment of the seismic hazard, have been proposed (Petersen et al. 2007; Akinci et al. 2009; Chan 2013; Jalalalhosseini 2017; Mousavi 2018).

Previous studies were generally oriented towards evaluating and comparing seismic intensities obtained by different models and their evolution in time while this research intends to analyse

the problem from a structural engineering perspective, evaluating the potential impacts of time-dependent models on the structural dimensioning and, more generally, on the design process.

Regarding this, it is useful to recall that the final objective of structural engineering consists of bounding the probability of failure of constructions during their lifetime and some target values are proposed in the codes of practice, such as Eurocode 0 or in Fajfar (2018), CEN (2006) and ASCE/SEI 7-16 (2017). This objective is generally obtained by simplified procedures avoiding a full probabilistic analysis and many recent works have been oriented to improving these methods in order to control the effective probability of failures (Fiorini et al. 2014; Franchin et al. 2018; Gkimprixis et al. 2019).

This study presents a general methodology to relate a time-dependent prediction of the earthquake occurrence to the structural properties (time-dependent capacity) necessary to ensure a target level of safety, expressed by the failure rate, including uncertainties regarding earthquake intensity, source location, propagation and structural properties. Results obtained by using the time-dependent Brownian Passage Time model (BPT) (Mattheus et al. 2002) are compared with results obtained with the time independent Poisson model (POI).

Numerical results concerning two different scenarios are presented, in order to provide general information about the potential impact of the time-dependent hazard on the required strength for constructions. In the former scenario a single point-source is considered and the constant structural capacity associated with the Poisson recursive model is compared to the time-dependent capacity associated with a time-dependent hazard rate, analysing the influence of the different site-to-source distances and the dispersion of the structural response. Effects produced by near fault phenomena are not considered in this study.

From a qualitative point of view, these results can be considered as an upper bound of the impact of the time-dependent hazard on design because realistic scenarios generally involve both low-medium intensity sources, spread across a large area, without periodic recursive properties, and high intensity sources, concentrated on faults and with periodic recursive properties. In order to provide some quantitative results about realistic situations, a scenario involving both a time-independent area-source and a time-dependent line-source are considered and the results relevant to different locations of the time-dependent source are discussed, as well as the influence of the structural system response and the recursive properties of the time-dependent source.

## 3.2 Methodology

### *3.2.1 Time-dependent seismic capacity*

The earthquake is here considered as an event $E$ whose occurrence in time is described by $f_T(t)$ providing the Probability Density Function (PDF) of the time elapsing from the last event (inter-arrival time). The origin $t=0$ of the time axis is placed at the instant of occurrence of the last event. Different models have been proposed to describe the probabilistic distribution of the inter-arrival time in literature and a review can be found in (Anagnos and Kiremidjian 1988). These models are generally based on the mean value of the inter-arrival time $T_R$ and on one or more parameters describing the expected dispersion of the inter-arrival time.

Starting from $f_T(t)$ and the relevant Cumulative Density Function (CDF) $F_T(t)$, it is possible to evaluate the hazard rate function $r_T(t)$ using the expression

$$r_T(t)=\frac{f_T(t)}{1-F_T(t)} \qquad (3.1)$$

which provides the instantaneous probability of occurrence at the time *t*, given that no event had occurred previously, and describes the hazard variation in time.

The probability of occurrence of one event within a time interval $\Delta t$ (e.g. construction lifetime) starting at $t$, given that the event had not occurred before, can be obtained by integrating the ratio $f_T(t+\tau)/(1-F_T(t))$. In the case of time intervals notably shorter than $T_R$, the likelihood connected to multiple events can be neglected and the occurrence of only one event can be considered as representative of the total probability of failure (Takahashi et al. 2004).

The system reliability depends both on the hazard rate and the properties of the response system. The latter can be collected in a vector $\vartheta \in \Theta$ of parameters describing dynamic properties and capacity limits, and the system reliability can be expressed by the failure rate $p_f(t,\vartheta) \cong r_T(t) P_f(\vartheta)$, expressing the instantaneous probability of failure at time $t$. It is obtained by combining the hazard rate function with the probability of failure $P_f(\vartheta)=P\left[\, failure \middle| \vartheta, E \right]$ conditional to the occurrence of the event $E$. In this study it is assumed that structural properties do not vary in time.

Structural design requires that the failure rate be lower than a threshold $p_f^*$ suggested by the codes and this study focuses on the evaluation of the structural seismic capacity necessary to

strictly satisfy the safety requirement $p_f(t,\vartheta) \le p_f^*$ for different values of time elapsing from the last event.

The failure rate of the structural system, given the event occurred, depends on system properties $\vartheta$, thus it is possible to associate to each instant *t* a relevant sub-set of system properties $\vartheta^* \in \Theta^* \subset \Theta$ necessary to strictly satisfy the target condition, i.e. $\vartheta^* : p_f\left(t,\vartheta^*\right) \cong r(t) P_f\left(\vartheta^*\right) = r(t) P_f^* = p_f^*$.

As a final result, the relationship $t \leftrightarrow \vartheta^*$ between the time elapsing from the last event and the minimum capacity required to achieve a fixed safety level can be discussed in order to analyse the impact of time-dependent hazard on structural design.

It should be noted that the recurrence models $f_T(t)$ proposed in literature are generally continuous and start from a probability density equal to 0 at the initial instant (in many models the function slope is also 0 at the initial instant), therefore previous equality can be evaluated only for $t > \bar{t}$ with $\bar{t} : r(\bar{t}) = p_f^*$, i.e. when the instantaneous probability of occurrence of the event becomes larger than the acceptable failure rate, otherwise inequality $p_f(t,\vartheta) \le p_f^*$ also holds for $P_f^* = 1$. This is not a marginal point because $P_f^* = 1$ means that no seismic capacity is required for an elapsed time shorter than $\bar{t}$ and models proposed in the literature sometimes provide quite large values for time $\bar{t}$.

### *3.2.2 Conditional probability of failure*

The probability of failure conditional to event occurrence $P_f$ depends on uncertainties regarding the event characteristics and the system properties.

Concerning uncertainties about the event, the following will consider a seismic scenario where magnitude $M$ may vary within a specified interval and sources are located at a point as well as spread on line or surface and combined sources are considered. The random values of $M$ are described by a PDF $f_M(m)$ defined on the magnitude interval $\Omega_M$. The site-to-source distance is also a random variable $X$ and it is described by a PDF $f_X(x)$ and the range of distance values is denoted as $\Omega_X$. The ground motion properties at the site are described by a scalar value providing the motion intensity $I$ and the relevant time evolution. The distribution of the intensity probability, given magnitude and distance of the event, is denoted by $f_I\left(i|M,X\right)$ and it can be determined on the basis of Ground Motion Prediction Equations (GMPEs), also

referred as Ground Motion Models (GMMs). Generally, GMPEs are in the form $\ln(I) = g(M,X) + \varepsilon_E(0,\sigma_E) + \varepsilon_A(0,\sigma_A)$, where $\varepsilon_E$ and $\varepsilon_A$ represent the inter-event and the intra-event residual respectively (Stafford et al. 2008). Both inter-event and intra-event residuals, are assumed to be a Gaussian random variable with 0-mean and variances $\sigma_E$ and $\sigma_A$, and their parameters are related to soil conditions and faulting type.

In conclusion, the PDF of the intensity can be obtained by combining different sources and possible magnitudes, as follow

$$f_I(i) = \int_{\Omega_X}\int_{\Omega_M} f_I(i|m,x)\, f_M(m)\, f_X(x)\, dm dx \qquad (3.2)$$

The recursive properties of ground motion with intensity $i$ at the site of interest are often described by the so called seismic hazard $H_I(i;t) = r_T(t)\left(1 - F_I(i)\right)$ providing the instantaneous probability of occurrence of a ground motion with intensity larger than $i$.The response properties of the structural system are described by using parameters providing the fragility curve $P_f^* = 1$ commonly described by using a log-normal CDF (Kennedy and Short 1994; Cornell et al. 2002), whose characteristic parameters are $\hat{c}$ and $\beta$ collected in the vector $\vartheta = [\hat{c}, \beta]$. In particular parameter $\hat{c}$ is the intensity measure producing 50% of failure (median) and $\beta$ is the logarithmic standard deviation describing the dispersion of results due to both record-to-record variability and uncertainties about the system response.

Therefore, the conditional probability of failure can be obtained by the convolution integral

$$P_f(\vartheta) = \int_{R^+} F_C(i,\vartheta)\, f_I(i)\, di \qquad (3.3)$$

where $R^+$ represent the set of positive real numbers.

## 3.3 Flow chart methodology

The relationship between system properties $\vartheta$ and failure rate $p_f$ is nonlinear and parameters $\vartheta^*$ at the target $p_f^*$ at a given inter-arrival time have been obtained by the following iterative procedure, illustrated in Figure 3.1. In the methodology, the hazard at the considered site is initially evaluated, considering the recursive and intensity properties of the source, as well as the propagation of the ground motion, from source to site. Separately, the structure capacity is obtained by computing the fragility curve, for the tentative value of the parameter vector $\vartheta$. Finally, the failure rate obtained is compared with the target value and $\vartheta$ is updated if required.

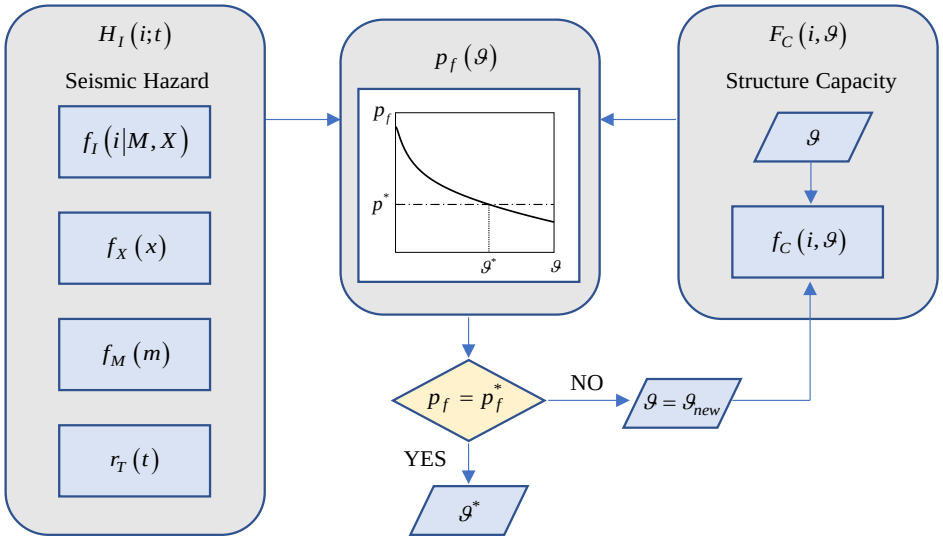



***Figure 3-1****. Flow chart illustration of the employed methodology.*

## 3.4 Point source

In this section a point-source has been analysed, comparing the results obtained from a POI recursive model with those from a BPT model. In addition, a parametric analysis has been carried out considering, different source-site distances and different values of logarithmic standard deviation for the structure's response.

### *3.4.1 Overall scenario and parametric analysis*

The characteristic earthquake refers to the Paganica Fault located in central Italy. The Paganica Fault is a normal fault, striking north-west and dipping to the south-west, belonging to the active fault systems of the central Apennines and responsible of the 2009-L'Aquila earthquake Mw=6.3 (Chiaraluce et al., 2009). The surface expression is represented by several fault segments which have ruptured differently in time and space during the 2009-L'Aquila earthquake (Galli et al. 2011). Several models have been proposed recently for the related seismogenic source and in this article we use the model of 3D seismogenic sources proposed by Pace et al. (2006), which is based on an interdisciplinary analysis integrating structural, geological (surface and subsurface), morphotectonic, paleoseismological, seismological, and rheological data. The relevant properties of the model of the fault have been chosen according to Polidoro et al. (2013).

The return period of the event $E$ is $T_R = 750\,yr$ and is characterised by a magnitude distribution following a truncated Gaussian distribution centred at $\mu = 6.3$ and spanning the range $\Omega_M = [m_m, m_M] = [5.8, 6.8]$ with a standard deviation $\sigma = 0.1667$. The magnitude PDF, scaled coherently with the interval length, assumes the form

$$f_M(m) = \frac{n(m;\mu,\sigma)}{N(m_M;\mu,\sigma) - N(m_m;\mu,\sigma)} \qquad (3.4)$$

where $n(m;\mu,\sigma)$ is the PDF of a normal distribution with mean $\mu$ and standard deviation $\sigma$, $N(m;\mu,\sigma)$ is the corresponding CDF.

The results presented compare outcomes from the POI model, providing a constant hazard rate, with results coming from a time-dependent recursive model. The POI model is defined by

$$f_T(t) = \frac{1}{T_R} e^{-t/T_R} \qquad F_T(t) = 1 - e^{-t/T_R} \text{ (3.5a, b)}$$

based on one parameter only, $T_R = 750yr$. In this case, the hazard rate defined in Equation 1, obtained starting from Equation 3.5, does not change over time, it is constant and assumes the value of $r(t) = r_0 = 1/T_R = 0.00133yr^{-1}$.

The distribution of the inter-arrival time of the time-dependent hazard was determined utilising the BPT model (Mattheus 2002), based on rebound theory (Reid 1910) and often used in the description of characteristic earthquake recurrence (e.g. Working Group of California Earthquake Probabilities 1999; Takahashi et al. 2004; Polidoro et al. 2013). Its expression is

$$f_T(t) = \sqrt{\frac{T_R}{2\pi\alpha^2 t^3}} \cdot e^{-\frac{(t-T_R)^2}{2T_R\alpha^2 t}} \qquad (3.6)$$

which represents a renewal model depending on two parameters, the mean inter-arrival time $T_R = 750yr$ and parameter $\alpha$ ruling the aperiodicity, intended as the possible deviation from the reference return period $T_R$. Some Authors (Matthew et al. 2002, Jalalalhoseini et al. 2017, Garcia-Aristizabal et al. 2012), conducted studies on the impact of $\alpha$ on hazard rate. It is noteworthy that low values of $\alpha$ leads to nearly symmetrical densities with a pronounced peak near the mean value and the curve degenerates into a Dirac's function when $\alpha$ tends to 0 (periodic event).

The hazard rate varies over time and it has been evaluated by Equation 6 assuming $\alpha = 0.43$, according to the study on the seismic scenario considered in Pace (2006). Figure 3.2 reports trends of the interarrival time probability density functions and hazard rates of the two models; in particular, the time-dependent model provides the same hazard rate as the POI model at the time $t_{1.0} = 422yr$, evaluated starting from the occurrence of the last event. It can be observed in

Figure 3.2a that POI model provides a not realistic description of the inter-arrival time, the PDF $f_T(t)$ is initially high for short time intervals and it regularly decreases, despite source activities generally show a periodic behaviour. Consequently, a structural design based on POI model is more demanding with respect the design based on BPT model when time intervals passed from the last event are short and it is less demanding for long time intervals. This can also be deduced from Figure 3.2b where the two hazard rates are compared, it is evident that the constant value of POI is larger than time-dependent hazard rate of the BPT model for the time intervals lower than $t_{1.0}$, and the opposite occurs for longer time intervals. For $t = t_{1.0}$ both models, POI and BPT model, furnish the same value of failure rate.

The seismic intensity has been measured using the Peak Ground Acceleration (PGA), obtained from the GMPE proposed by Sabetta and Pugliese (1996), where the logarithm of the seismic intensity $I$ produced by an event with intensity $m$, at a distance $x$, can be evaluated by

$$\log_{10}(I) = a + bm - c\log_{10}\sqrt{x^2 + h^2} + e_1 S_1 + e_2 S_2 + \varepsilon \qquad (3.7)$$

where $\varepsilon = \varepsilon_E = 0$ is a 0-mean Gaussian random variable that represent the inter-event residual, $S_1$ and $S_2$ are parameters depending on characteristic of soil, $h$ is the fictitious depth, and $a, b, c, e_1, e_2$ are the model constants. Respect to expression previously introduced for the GMPE, in the Equation (3.7) the intra-event residual is not taken into account ($\varepsilon_A = 0$). In this study, parameters of the Equation (3.7) are evaluated from the work of Sabetta and Pugliese (1987). For the parameters, the following values are assumed: $a = -1.562$, $b = 0.306$, $c = 1$, $S_1 = 1$, $e_1 = 0.169$, $S_2 = 0$ and $h = 5.8$, while the standard deviation of $\varepsilon$ equal to 0.173. It is worth noting that in the case of point source, the site to source distance $x = \bar{x}$ is fixed, thus the integral of Equation 2 reduces to

$$f_I(i) = \int_{\Omega_M} f_I\left(i \middle| m, \bar{x}\right) f_M(m)\, dm \qquad (3.8)$$

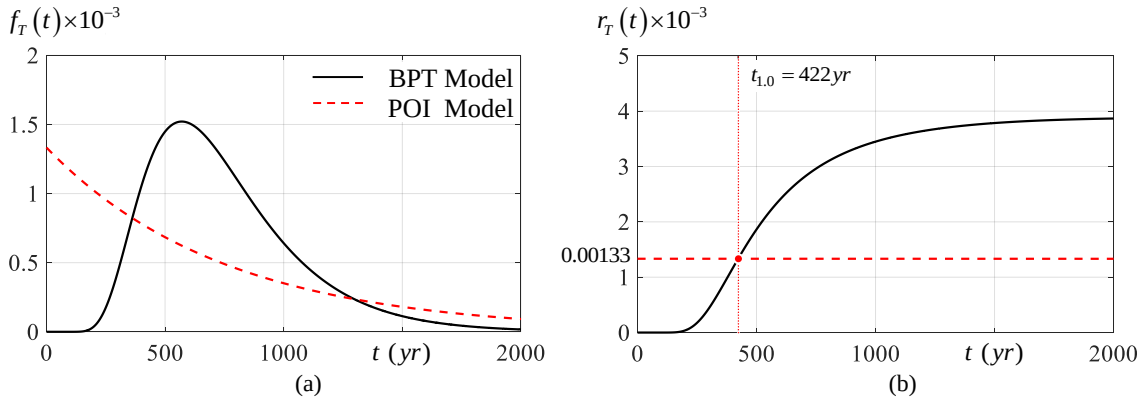



***Figure 3-2****. (a) Probability density functions of interarrival time and (b) hazard rates for the POI model and the BPT model.*

### *3.4.2 Target failure rate and capacity properties*

In this section the seismic capacity required to ensure a target failure rate equal to $p_f^*$ has been evaluated for different time intervals elapsing from the last event.

The seismic capacity has been described by the parameter $\hat{c}$, introduced in Section 2, and the target value of the failure rate has been assumed equal to $p_f^* = 6.667 \cdot 10^{-5}$. This was obtained by combining the conditional probability of failure $P_f = 0.05$ (ASCE/SEI 7-16, Table 1.3-2) considering a Category III for the building risk, with the reference value of the hazard rate $r_0 = 0.00133 yr^{-1}$ provided by the POI recursive model for the seismic scenario considered.

The following results report the numerical values referring to three specific times $t_{0.5} = 333yr$, $t_{1.0} = 422yr$ and $t_{2.0} = 661yr$. In particular, at the intermediate value 422*yr*, the hazard rates for both the POI model and BPT model are the same and are equal to $r_0$, while at the first value 333*yr* the hazard rate of the time-dependent model is equal to $0.5r_0$, and at the third value 661*yr* the time-dependent hazard rate is equal to $2.0r_0$. Finally, the variation in the required capacity for $t$ larger than the limit value $\bar{t}$, such that $r_T(\bar{t}) = p_f^*$, is graphically reported and discussed.

Two parametric analyses were developed separately in order to study the results for source-to-site distances $x$ varying from 5 km to 20 km, and capacity dispersion $\beta$ varying from 0.40 to 0.80 (FEMA P-750; ASCE/SEI 7-16; Kennedy 2011).The extended range of $\beta$ values is defined considering that the epistemic uncertainties in the structural model which could have a significant influence on the collapse capacity producing a high capacity dispersion (Dolsek 2009).

In the former analysis the site-to-source distance varies and the capacity dispersions are fixed, whereas in the latter parameter $\beta$ varies while the results are evaluated at a fixed distance.

### *3.4.3 The impact of site-to-source distance on the required capacity*

The effect of the distance on the required capacity has been evaluated by changing site-to-source distance $\bar{x}$ = 5 km, 10 km and 20 km, whilst parameter $\beta = 0.60$ remained the same. As mentioned above, the effects produced by near fault phenomena are not considered. In addition, the influence of the time period from the last event has been considered and compared using the three different times previously introduced.

Figure 3.3 illustrates a comparison between seismic hazard and required capacity, given event occurrence. The red dashed line reports the Complementary Cumulative Density Function (CCDF) $G_I(i) = 1 - F_I(i)$ of seismic intensity $I$ and describes the probability of exceedance of $I$. The black continuous line reports the probability of failure for systems with different capacities, expressed by the response parameter $\hat{c}$. As expected, the two curves intersect approximately when the median value of the predicted seismic hazard coincides with the capacity parameter $\hat{c}$, while the probability of failure is larger than the probability of intensity exceedance for rare events as a consequence of the capacity dispersion.

In order to discuss the trend of $P_f$ it is useful to start from the case of deterministic capacity ( $\beta$ = 0) In this case, the integral of Equation 3.3 weights only the portion of intensity PDF where intensity is larger than capacity, and the conditional $P_f$ is overlapped to the Hazard Function $G(i)$. When a dispersion $\beta$ is added, values of $i$ that are lower or higher than $\hat{c}$ are differently weighted by the PDF of $I$, based on its slope. Consequently, $P_f$ is larger than $G(i)$ when the slope of PDF of $I$ is negative in the neighbourhood of $\hat{c}$ and it is lower when the slope is positive.

Figure 3.3 also shows a comparison of the results concerning different site-to-source distances. It is evident that the capacity corresponding to the target conditional probability of failure $P_f^* = 0.05$ decreases when the site-to-source distance increases. More precisely, the capacity corresponding to $P_f^*$ is $\hat{c} = 1.485\text{g}$ for the smallest considered distance, equal to 5 km, and it decreases by 34% and 63% passing from 5 km to the larger values 10 km and 20 km, respectively.

Figure 3.4 shows the analysis results considering different hazard rates, assuming parameters $\beta = 0.60$ and $\bar{x} = 5;10;20\,\text{km}$. In particular, the first row reports the hazard curve $H_I(i)$ for three specific instants, i.e. 333*yr*, 422*yr*, and 661*yr*, while the second row reports the related $p_f(\hat{c})$ curves. Table 3.1 provides the data related to the required capacity and compares the results obtained for the 3.3 different site-to-source distances. As mentioned previously, the reference time is 422*yr*, in this case the hazard rates provided by the POI and BPT models are the same. In addition, variations from the reference case have been evaluated by means of the expression $\Delta\hat{c} = \left(\hat{c} - \hat{c}_{ref}\right)/\hat{c}_{ref}$ .

According to the data reported in Table 1, by elapsing time from the last event, required capacity, $\hat{c}$ increases. However, the variation of the required capacity is notably smaller than the variation in event occurrence. In this regard, it is useful to recall that the hazard rate at $t_{0.5} = 333yr$ is half the reference value $r_0$ but the capacity reduction is approximately 23.28% for all the distances considered. On the other hand, the hazard rate at $t_{2.0}$=661*yr* is twice the reference value but the required increment in the capacity parameter is limited to 25.84% across the distance values. Therefore, a notable variation in the hazard rate does not translate into a similarly notable variation in the capacity required for the structure. Based on the percentages of the differences provided in the last column of Table 3.1, it can be observed that the change in the required capacity $\hat{c}$ by elapsing time is the same for different values of $r$. At this regard, it can be observed that the integral of Equation 3.3 can be posed in the form

$$P_f = \int \Phi\left(\frac{\eta - \ln(\hat{c})}{\beta}\right)\varphi\left(\frac{\eta - \ln(\hat{i})}{\sigma}\right)d\eta \qquad (3.9)$$

where $\Phi$ and $\varphi$ are the CDF and the PDF of the normal distribution, respectively, and $\eta = \ln(i)$ . In the case considered in this section $\beta$ and $\sigma$ are two constants, therefore the integral does not vary if the difference between $\ln(\hat{c})$ and $\ln(\hat{i})$ does not vary or, in other words, the ratio $\hat{c}/\hat{i}$ does not vary.

As a first consequence, once the target $P_f^*$ is chosen, the ratio $\hat{c}/\hat{i}$ becomes fixed. Furthermore, if $\hat{i}$ varies by a factor $\lambda$, also the relevant $\hat{c}$ must change by the same factor. It can be observed that $i$ varies by factors 0.66 and 0.37 passing from 5km to 10km and 20 km, (according to Equation 3.7), and the same factors influence the variation of $\hat{c}$ . As an additional result, it can be observed that the relative variations of $\hat{c}$ , reported in the last column, concern ratios between

quantities varied by the same factor and they cannot vary with the distance, so that the same results hold for different location.

For a deeper insight into the trend in time of the capacity required for the target failure rate, the values of $\hat{c}$ obtained in the range $[0yr, 1500yr]$ have been reported in Figure 3.5, for the three distances considered. Based on the curves shown in Figure 5, it can be said that when the site is close to the source ($\bar{x}$ = 5 km) elapsed time affects the required structure performance considerably more than the case in which the site is far from the source ($\bar{x}$ = 20 km in this study). As expected, for each distance to source, the capacity increases as the time elapsing since the last event increases and derive from the shape of hazard rate reported in Figure 3.2. The trend in time shows that capacity requirements based on the POI model are lower than the corresponding requirements evaluated by the BPT model for inter-arrival times larger than the balance point $t_{1,0} = 422yr$, but the differences are not so large and they do not significantly vary for $t$ larger than $T_R$. A notably different trend is observed in the range of short time intervals between POI model and BTP model, where it can be observed that the BPT model rapidly vary and drop to zero at time $\bar{t} = 215yr$. Based on this result, a null or very low seismic capacity is required for buildings erected after the last earthquake in a time window approximately equal to $T_R/2$. This conclusion is partially mitigated in the case of multiple and independent sources, as discussed in the following section. The capacity required for intervals shorter than the limit value $\bar{t}$ is equal to 0, as discussed in the previous section, and derives from decay properties of the inter-arrival time distribution function $f_T$.

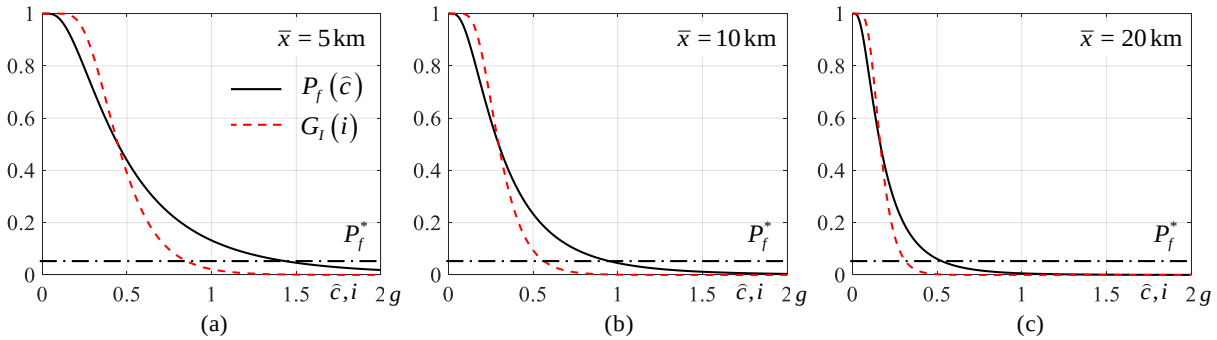



***Figure 3-3****.* $G_I(i)$ *and* $P_f(\hat{c})$ *curves for* $\beta = 0.60$ *and (a)* $\bar{x}$ *= 5 km, (b)* $\bar{x}$ *= 10 km and (c)* $\bar{x}$ *= 20 km.*

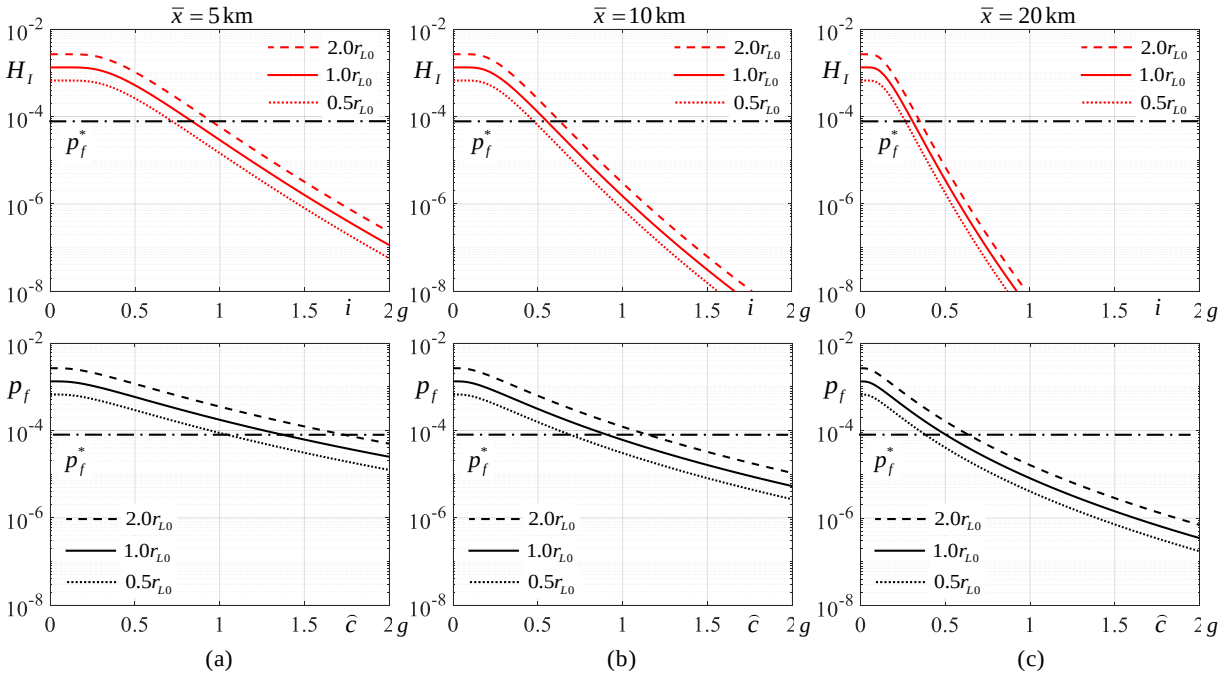



***Figure 3-4****. Analysis result for different hazard rates considering parameters* $\beta = 0.60$ *and (a)* $\bar{x} = 5\,\text{km}$*, (b)* $\bar{x} = 10\,\text{km}$*, (c)* $\bar{x} = 20\,\text{km}$*; first row reports* $H_I(i)$ *curves for point source; second row reports* $p_f(\hat{c})$ *curves.*

***Table 3-1****. Required capacity for the target failure rate considering* $\beta = 0.60$*, different distances and different times.*

| *Distance* $\bar{x}$ *(km)* | *Median value of demand (g)* | *t (yr)* | *Capacity* $\hat{c}$ *(g)* | $\Delta\hat{c}$ *(%)* |
|---|---|---|---|---|
| *5* | *0.4474* | *333* | *1.139* | *-23.28* |
| | | *422* | *1.485* | *0.00* |
| | | *661* | *1.869* | *25.84* |
| *10* | *0.2964* | *333* | *0.755* | *-23.28* |
| | | *422* | *0.984* | *0.00* |
| | | *661* | *1.238* | *25.84* |
| *20* | *0.1645* | *333* | *0.419* | *-23.28* |
| | | *422* | *0.546* | *0.00* |
| | | *661* | *0.687* | *25.84* |

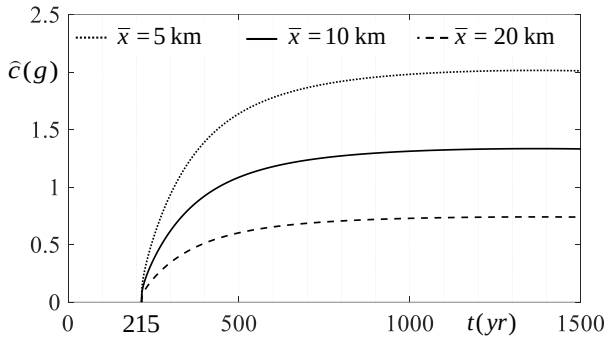



***Figure 3-5**. The change in required capacity by elapsing time considering $\beta = 0.60$ and variable $\bar{x}$.*

### *3.4.4 The impact of $\beta$ on the required capacity*

The influence of parameter $\beta$, describing the dispersion of the seismic intensity producing failure, has been analysed in this section considering a fixed site-to-source distance $\bar{x}$ = 10 km and three different values of $\beta$, i.e. $\beta = 0.40$, $\beta = 0.60$, and $\beta = 0.80$.

Figure 3.6 provides a comparison between seismic intensity and required capacity, concerning structural systems with different dispersion $\beta$, given earthquake occurrence. The dashed red line reports the CCDF $G_I(i)$ of seismic intensity $I$ and the continuous black line reports the conditional probability of failure for systems with different capacity parameters $\hat{c}$.

The capacity corresponding to the target conditional probability of failure $P_f^* = 0.05$ notably increases as the capacity dispersion increases. More precisely, the capacity necessary for $P_f^*$ is $\hat{c} = 0.765g$ for the smallest considered value of $\beta = 0.40$, while it increases about 29%, and 70% passing from $\beta = 0.40$ to the larger values of $\beta = 0.60$ and $\beta = 0.80$, respectively. Figure 3.7 shows the analysis results considering different hazard rates assuming parameters $\beta = 0.40; 0.60; 0.80$ and epicentral distance $\bar{x} = 10$ km; in particular, the first row reports the hazard curve $H_I(i)$ for three specific instants, i.e. 333yr, 422yr, and 661yr, while the second row reports the related $p_f(\hat{c})$ curves.

Table 3.2 provides the data related to the required capacity and compares the results obtained for the 3 different values of $\beta$, and the expression $\Delta\hat{c} = \left(\hat{c} - \hat{c}_{ref}\right)/\hat{c}_{ref}$ has been used to evaluate the change in the required capacity with respect to reference case $t_{1,0} = 422yr$. Also, in this case, the variation in the required capacity is notably smaller than the variation in event recurrence.

Contrary to the previous case, it can be observed that the percentage variation in the required capacity $\hat{c}$ due to different elapsed time is different for different values of $r$. Based on the

previous discussion about the results reported in Table 3.1 and the integral ruling of the target $P_f^*$, it can be concluded that the results of this section are influenced by the parameter $\beta$ and the ratio $\hat{c}/\hat{i}$ at the target $P_f^*$ is no longer a constant. Indeed, according to the discussion of Figure 3.3, in this case $\hat{c}$ is greater than $\hat{i}$ and increments in $\beta$ lead to increments in $\hat{c}$.

Finally the trend in time of the capacity required for the target failure probability has been discussed and the values of $\hat{c}$ obtained in the range $[0yr, 1500yr]$ have been reported in Figure 3.8 for the three values of $\beta$ considered. As expected, the changes in required capacity by elapsing time are almost the same, as shown in Figure 5.

The curves in Figure 3.8 show that the required capacity increases as time elapses from the last event and that the rate of this increment is not a constant but decreases with elapsing time.

The qualitative trends of the curves reported in Figure 3.8 are similar to the trends of curves reported in Figure 3.5, and follow from the trend of time-dependent hazard rate. In this case, it can be observed that the relationship between required capacity and dispersion is nonlinear, so that differences between the cases $\beta = 0.40$ and $\beta = 0.60$ are smaller with respect to differences between the cases $\beta = 0.60$ and $\beta = 0.80$. Previous conclusions about the different trends for $t$ smaller or larger that the balance value $t_{1,0}$ hold in this case too.

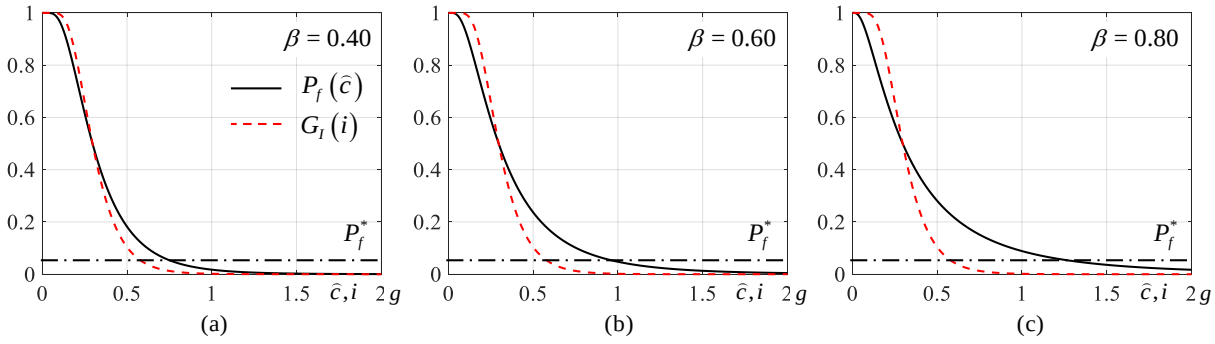



***Figure 3-6.*** *$G_I(i)$ and $P_f(\hat{c})$ curves for $x = 10\text{km}$ and (a) $\beta = 0.40$, (b) $\beta = 0.60$ and (c) $\beta = 0.80$.*

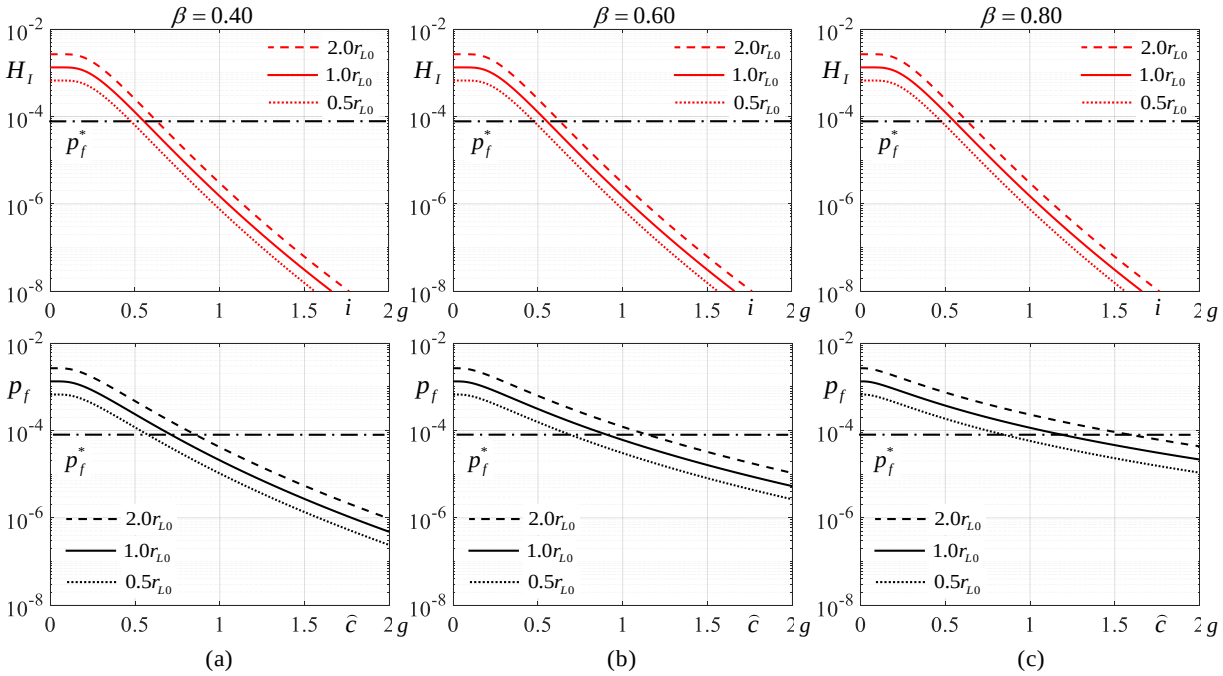



***Figure 3-7****. Analysis result for different hazard rates considering the epicentral distance* $\bar{x} = 10\text{km}$ *and (a)* $\beta = 0.40$ *, (b)* $\beta = 0.60$ *, (c)* $\beta = 0.80$ *; first row reports* $H_I(i)$ *curves for point source; second row reports* $p_f(\hat{c})$ *curves.*

***Table 3-2****. Table 2 Required capacity for the target failure rate,* $\bar{x} = 10$ *km, different* $\beta$ *and different times.*

| *Parameter* $\beta$ | *Median value of demand (g)* | *t (yr)* | *Capacity* $\hat{c}$ *(g)* | $\Delta\hat{c}$ *(%)* |
|---|---|---|---|---|
| *0.40* | *0.2964* | *333* | *0.620* | *-18.89* |
| | | *422* | *0.765* | *0.00* |
| | | *661* | *0.917* | *19.91* |
| *0.60* | *0.2964* | *333* | *0.7548* | *-23.28* |
| | | *422* | *0.9838* | *0.00* |
| | | *661* | *1.2381* | *25.84* |
| *0.80* | *0.2964* | *333* | *0.9406* | *-27.92* |
| | | *422* | *1.3049* | *0.00* |
| | | *661* | *1.7334* | *32.84* |

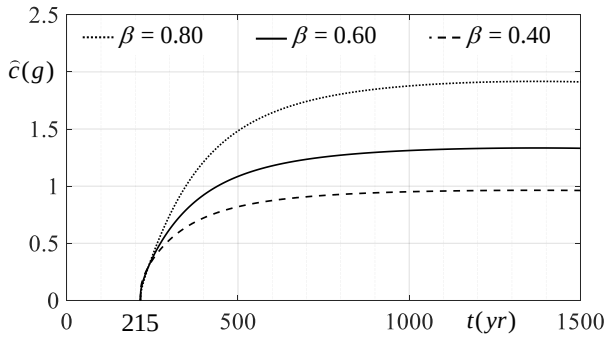



***Figure 3-8****. The change in required capacity by elapsing time considering* $\bar{x} = 10\,\text{km}$ *and variable* $\beta$*.*

## 3.5 Combined sources

In this section the more realistic case of multiple sources has been considered and the analysis combines two sources with different properties: an area-source producing medium-low intensity earthquakes whose recurrence has been described by the no-memory POI model and a line-source producing high intensity earthquakes with a periodic recurrence described by the BPT model. The results consider different time periods elapsing from the last event and the influence of three variable parameters have been discussed: the location of the time-dependent source, the dispersion of the structural response due to model uncertainty and record to record variability, and the recursive properties of the time-dependent source.

### *3.5.1 Overall scenario and parametric analysis*

A time-independent reference scenario, is initially introduced. It consists of an area source and the probability of event occurrence is homogeneously distributed across a circular surface with a radius $x_M = 50\,\text{km}$, centred at the considered site. The relevant PDF of the site-to-source distance has been described by the expression

$$f_{XA}(x) = \frac{2x}{x_M^2} \qquad \Omega_{XA} = [0, x_M] \tag{3.10}$$

Furthermore, the reference source is able to generate events $E$ with magnitude $m$ in the range $\Omega_M$, and the PDF of possible magnitudes is provided by the Gutenberg-Richter law

$$f_M(m) = \frac{\gamma e^{-\gamma m}}{e^{-\gamma m_m} - e^{-\gamma m_M}} \qquad \Omega_M = [m_m, m_M] \tag{3.11}$$

where $m_m = 5$ and $m_M = 8$ are the lower and upper magnitude limits, respectively, and the function slope is described by parameter $\gamma = 1 \times \ln(10)$ (Gulia et al. 2016). The time recurrence is provided by the POI model (Equation 3.4), defined by the characteristic parameter $r_0 = 0.31 yr^{-1}$ (Scozzese et al. 2020), which describes the constant hazard rate and coincides with the mean annual frequency of the event, so the mean inter-arrival time is $T_R = 1/0.31 = 3.22 yr$.

This reference scenario has been compared with a different scenario characterized by the same magnitude range and overall mean annual frequency, but, in the alternative scenario, the magnitude interval has been split into two sub-intervals, i.e. $\Omega_M = \Omega_{MA} \cup \Omega_{ML}$ as illustrated in Figure 3.9. The former range $\Omega_{MA} = [m_m, m_1]$ concerns low-medium intensity earthquakes generated by a time independent area-source and the latter range $\Omega_{ML} = [m_1, m_M]$ concerns high intensity earthquakes generated by a time-dependent line-sources (characteristic earthquakes, Ellsworth et al. 1999). The magnitude that bounds the two ranges has been fixed at $m_1 = 7$. The mean annual frequencies of occurrence of events belonging to the two intervals $\Omega_{ML}$ and $\Omega_{MA}$ have been defined according to the previous distribution (Equation 3.11), considering that the frequency of events with a magnitude in the sub-domain $\Omega_{ML}$ is $r_{0L} = r_0 (1 - F_M(m_1)) = 0.0028 yr^{-1}$ (mean inter-arrival time $T_{RL} = 1/0.0028 = 358 yr$), where $F_M(m)$ is the CDF of previous magnitude distribution defined in the total range $\Omega_M$. The resulting mean annual frequency of events in the sub-domain $\Omega_{MA}$ is $r_{0A} = r_0 F_M(m_1) = 0.3072 yr^{-1}$ (mean inter-arrival time $T_{RA} = 1/0.3072 = 3.25 yr$).

The PDF of the magnitudes must be defined separately for the two sources. For what concern the area-source, it has been assumed that magnitude values are distributed according to the previous Gutenberg-Richter law (reference case), and can be simply obtained by scaling the previous overall distribution (Equation 3.11)

$$f_{MA}(m) = \frac{f_M(m)}{1 - F_M(m_1)} \qquad \Omega_{MA} = [m_m, m_1] \qquad (3.12)$$

For what concern the line source, it has been assumed that the magnitudes of characteristic earthquakes are distributed according to a Gaussian distribution (Equation 3.4) centred at the intermediate magnitude $\mu = 7.5$, characterised by a standard deviation $\sigma = 0.1667$ (Polidoro et al. 2013), and scaled coherently with the interval length $\Omega_M = [m_1, m_M]$, as in Equation 3.4.

Magnitude distribution of area source and line source considering three different times is reported in Figure 3.9.

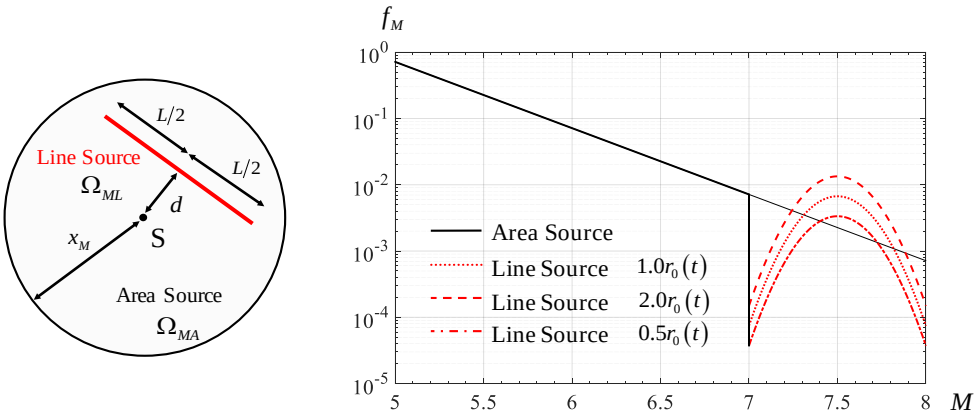



***Figure 3-9**. Magnitude distribution of area source and line source at three different times.*

Regarding recurrence in time of events in the combined source (Line + Area), the recurrence coming from the area-source has been described using the POI model (Equation 3.5) with a constant hazard rate is $r_{0A}$, while the recurrence of events coming from the line-source has been described using the BPT model (Equation 3.6) with mean inter-arrival time $T_{RL} = 1/r_{0L} = 358yr$. In the latter case, the PDF of the inter-arrival time $t$ also depends on periodicity parameter $\alpha$, which influences the shape of the curve.

The site-to-source distance $x$ of earthquakes with magnitudes in $\Omega_{MA}$ is homogeneously distributed in the previous aforementioned area, so the previous distribution of Equation 10 has been used, while the site-to-source distance $x$ of earthquakes with magnitudes in $\Omega_{ML}$ is homogeneously distributed along the line and the following distribution has been used

$$f_{XL}(x) = \frac{2}{L}\frac{x}{\sqrt{x^2 - d^2}} \qquad \Omega_{XL} = \left[d, \sqrt{d^2 + (L/2)^2}\right] \qquad (3.13)$$

where $L = 30$ km is the length of the line-source and $d$ is the distance from the line to the site. In the following it has been assumed that the length of the line source is fixed while a parametric analysis has been carried out considering different positions, corresponding to different values of $d$. The GMPE $f_I(i|m,x)$, previously introduced for the point source case, has been adopted here too.

Finally, the dispersion of the structural response has been described by $\beta$, considering different values in the following analyses, and the capacity parameter $\hat{c}$ is calculated for each set of

parameters, in order to obtain a failure rate $p_f^* = 2.0 \cdot 10^{-4}$, chosen in accordance with general trends of codes (Fajifar 2018).

Summing up, the following parametric analysis considers three independent parameters: the distance $d$, in order to evaluate the influence of the location of the time-dependent source on the required capacity $\hat{c}$, the periodicity parameter $\alpha$ in order to evaluate the influence of the recurrence properties of high intensity earthquakes, and the dispersion $\beta$, in order to evaluate the influence of the structural system properties.

Concerning the distances, three values have been investigated $d = \bar{d}$, $d = 0.5\bar{d}$ and $d = 2.0\bar{d}$, where $\bar{d} = 15\,\text{km}$. The reference distance $\bar{d} = 15\,\text{km}$ represents a particular location of the line source producing approximately the same hazard as the reference scenario (area source with POI recursive model) at $p_f^* = 2.0 \cdot 10^{-4}$.

Regarding parameter $\beta$, values $\beta = 0.40$, $\beta = 0.60$, and $\beta = 0.80$ have been considered (as in the previous case of point source), in accordance with the variability proposed in FEMA P-750, ASCE/SEI 7-16.

Finally, for periodicity parameter $\alpha$, the three values $\alpha = 0.40$, $\alpha = 0.50$ and $\alpha = 0.60$ have been investigated.

### *3.5.2 Impact of the distance of the time-dependent line-source*

The inter-arrival time of the line source has been defined according to Equation 6, using parameters $\alpha = 0.5$ and $T_R = 358yr$. The first group of analyses carried out considered an elapsed time $t_{1.0} = 185yr$ such that the constant hazard rates of the POI model coincide with the hazard rate of the time-dependent BPT model, i.e. $r_L(t_{1.0}) = r_{0L}$.

In the first part, the effects of the line source distance on the required capacity (in this discussion on result, the term "capacity" will be used to briefly denote parameter $\hat{c}$ corresponding to the intensity associated with a probability of failure equal to 0.5) has been evaluated by changing distance $d$, whilst parameter $\beta = 0.60$ and periodicity parameter $\alpha = 0.50$ are fixed. The effect of the distance on the required capacity has been evaluated by comparing the results obtained for $d = \bar{d} = 15\,\text{km}$ with the results obtained for the distances $d = 0.5\bar{d} = 7.5\,\text{km}$ and $d = 2.0\bar{d} = 30\,\text{km}$

.

In the second part, the effect of time-dependent hazard rate has been investigated considering different values of time elapsing $t_i$ from the last event. In particular three specific times $t_{0.5} = 139yr$, $t_{1.0} = 185yr$, and $t_{2.0} = 371yr$ have been considered, the hazard rate of the time-dependent line source is equal to the reference value $r_{0L}$ at $t_{1.0}$, while it is $0.5r_{0L}$ at $t_{0.5}$ and it is $2.0r_{0L}$ at $t_{2.0}$. Figure 3.10 shows the hazard rate considered and the particular specific times adopted in the following analyses.

The first row in Figure 3.11 reports the seismic hazard $H_I(i)$ describing the probability of exceedance of the intensity $i$, for the three different locations of the line-source, $d = 7.5\,\text{km}$ (Figure 3.11a), $d = 15\,\text{km}$ (Figure 3.11b), and $d = 30\,\text{km}$ (Figure 3.11c). The figure reports both the total hazard $H_I(i)$ derived from the combined sources and the single contributions coming from the area source $H_I^A(i)$ and line source $H_I^L(i)$. It can be observed that, for exceedance probabilities higher than $10^{-2}$, usually used for checks at serviceability limit states, the hazard is essentially due to the area source while the contribution due to the line source is negligible. At lower values of exceedance probabilities, usually interesting for checks at ultimate limit states ($2 \cdot 10^{-3}$), the hazard is mainly due to the line source. Developed study concerns a target failure rate equal to $2 \cdot 10^{-4}$ and results are strongly influenced by rare events related to high intensities with low probability of exceedance. This range of intensities is notably influenced by the time-dependent source and this makes the required capacity varying with inter-arrival time. By increasing the site to source distance, the hazard decreases, the contribution due to the line source becomes negligible for a distance $d = 30\,\text{km}$ and the hazard is almost completely due to the area source (Figure 3.11c).

The second row in Figure 3.11 compares the combined hazard with the failure rate related to different capacity $\hat{c}$. In general, it can be observed that the required capacity is close to the hazard intensity when high probability of failure is considered. Differently the distance between the two curves becomes ever larger when considering decreasing probabilities of failure. As expected, the location of the line-source producing rare and strong earthquakes notably influences the capacity required for low values of the probability of failure.

Figure 3.12 compares the results obtained for different time intervals elapsing from the last event, for the three line-source locations. The upper row reports the hazard curves for the three-time intervals and the lower row reports the curves relating required capacities with expected probability of failure. A significative influence of the time-dependent hazard has been observed

in the case $d = 7.5\,\text{km}$, where the contribution of the line source is prevalent respect the to area source in the definition of the overall hazard. Moving away from the line source, the time-dependent effects are less relevant up to be negligible for $d = 30\,\text{km}$ where the area source contribution is predominant. Also, the curves reporting failure rate vs capacity follow the same trends of the hazard curve. Furthermore, the differences among hazard curves and capacity curves reduce when increasing the distance of the time-dependent line-source, and the differences relevant to required capacity become almost negligible in the case $d = 30\,\text{km}$.The result provides a useful information about the dimension of the area where structural design should be carried out considering the previous seismic history.

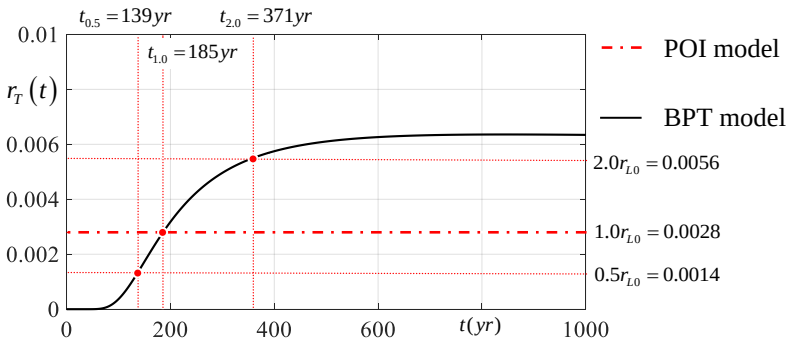



***Figure 3-10****. Hazard rate for BPT model and POI model with* $T_R = 358yr$ *considering* $\alpha = 0.50$.

Table 3.3 provides numerical values of the required capacity for line sources located at different distances from the site, considering the three different hazard rates, and relevant instants, as previously discussed. Reported capacities refer to the target failure rate ( $p_f = 2\times10^{-4}$) and the table also reports the percentage variations $\Delta\hat{c} = \left(\hat{c} - \hat{c}_{ref}\right)/\hat{c}_{ref}$ of required capacity with respect to the reference case ($d = 15\,\text{km}$, $t = t_{1.0}$). At the reference time, required capacity spans 1.366g to 2.105g, passing from the farthest to the nearest location of the line-source, and, in the latter case, it attains the maximum value 2.590g at the time $t_{2.0} = 371yr$.

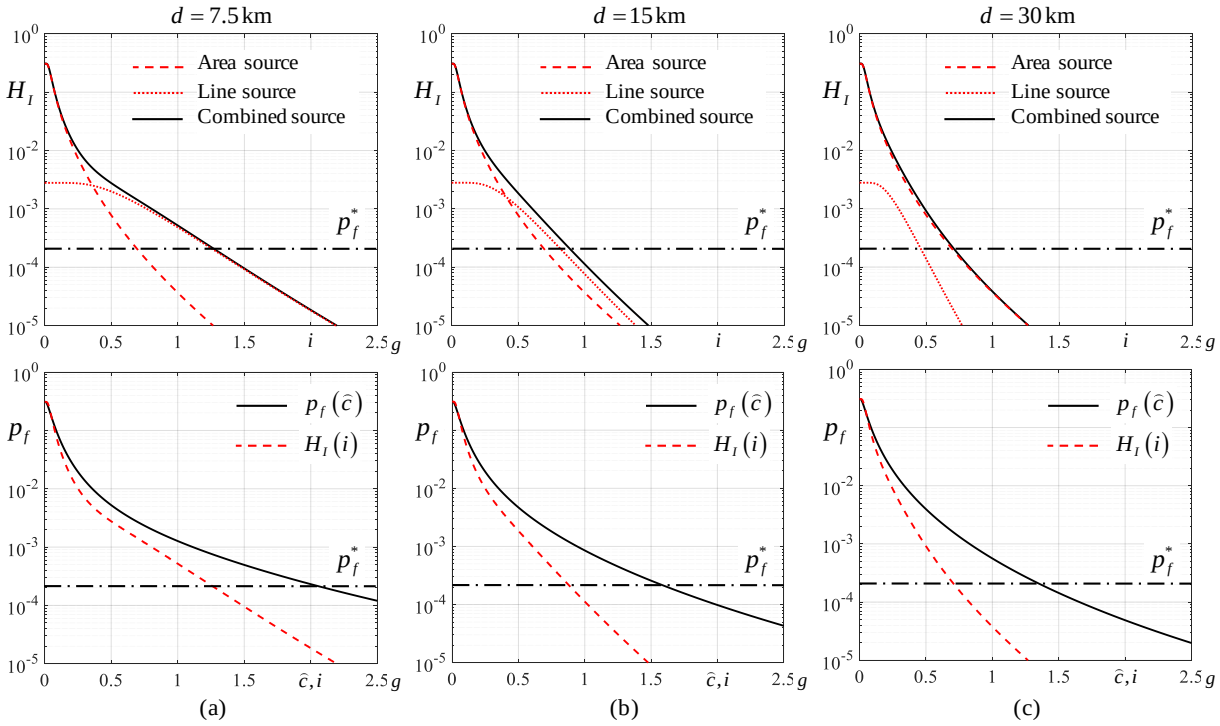



***Figure 3-11**. Analysis results considering parameters $\alpha$=0.50 , $\beta$ = 0.60 and (a) $d$ = 7.5 km , (b) $d$ = 15 km , (c) $d$ = 30 km ; first row reports $H_I(i)$ curves for area, line and combined source; second row reports $H_I(i)$ and $p_f(\hat{c})$ curves.*

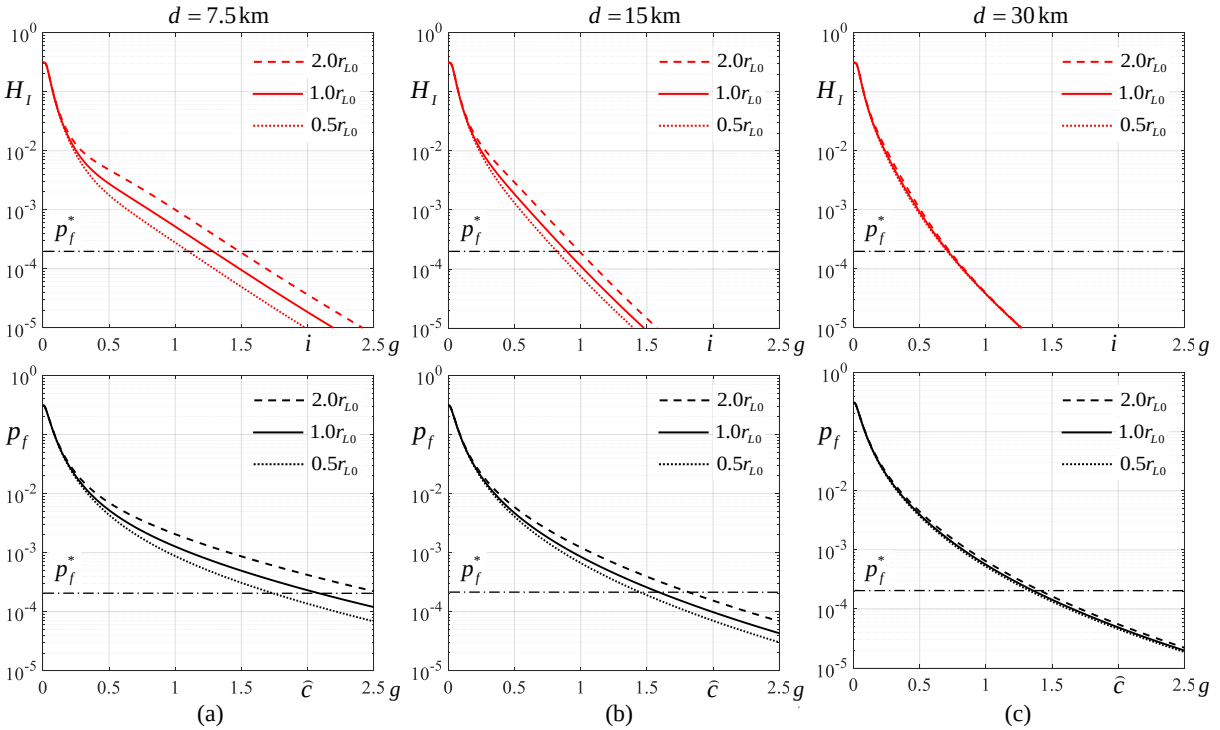



***Figure 3-12**. Analysis result of combined source for different hazard rates considering $\alpha$=0.50 , $\beta$ = 0.60 and (a) $d$ = 7.5 km , (b) $d$ = 15 km , (c) $d$ = 30 km ; first row reports $H_I(i)$ curves and second row $p_f(\hat{c})$ curves.*

***Table 3-3****. Required capacity for the target failure rate considering* $\beta = 0.60$ *and* $\alpha = 0.50$ *for different site to source distance* $d$ *and different times elapsing since the last event.*

| | *Capacity* $\hat{c}$ *(g)* | | | $\Delta\hat{c}$ *(%)* | |
|---|---|---|---|---|---|
| *d (km)* | $0.5r_{L0}$ $r_T(139yr)$ | $1.0r_{0L}$ $r_T(188yr)$ | $2.0r_{0L}$ $r_T(371yr)$ | $0.5r_{L0}$ $r_T(139yr)$ | $2.0r_{0L}$ $r_T(371yr)$ |
| *7.5* | *1.755* | *2.105* | *2.590* | *-16.638* | *23.018* |
| *15* | *1.478* | *1.623* | *1.851* | *-8.919* | *14.104* |
| *30* | *1.337* | *1.366* | *1.418* | *-2.121* | *3.866* |

### *3.5.3 Impact of system capacity dispersion (*$\beta$*)*

A second group of analyses have been carried out considering three different values of the structural response dispersion $\beta$, $\beta = 0.40$, $\beta = 0.60$, and $\beta = 0.80$. Distance $d = 15\,\text{km}$ and recursive parameter $\alpha = 0.50$ do not vary and the previous three elapsed times $t_{0.5} = 139yr$, $t_{1.0} = 185yr$, and $t_{2.0} = 371yr$ have been considered.

Figure 3.13 reports the probability of failure $p_f(\bar{c})$ of systems with different capacities, considering different values of $\beta$, in addition to the graph of the seismic hazard $H_I(i)$. In general, it should be noted that by increasing $\beta$, the difference between hazard and capacity increases (roughly doubling from the lower to the higher $\beta$ value).

Although the values of capacity obtained considering $\beta = 0.40$ are close to hazard values for probabilities greater than $10^{-2}$, there is a considerable difference between hazard and capacity for all probabilities when values 0.60 or 0.80 are considered. In addition, by decreasing failure probability, the difference between hazard and failure probability increases considerably for all the values of $\beta$ especially for the value of $\beta = 0.80$. In this case, the slope of the curve describing the failure rate is quite small and this leads to a noticeable increment in the required capacity $\hat{c}$.

Figure 3.14 shows the influences of $\beta$ variation for the three different hazard rates ($0.5r_{0L}$, $1.0r_{0L}$ and $2.0r_{0L}$). It can be noted that both hazard and required capacity increase when increasing the elapsing time. In this case the influence of the time-dependent hazard are limited as the line source does not play an important role in relation to the overall hazard.

The values of required capacity $\hat{c}$ for $\beta = 0.40$, $\beta = 0.60$ and $\beta = 0.80$ at the three mentioned times $t$ have been presented in Table 4. According to the outcomes provided in this table, the values

of capacity corresponding to $\beta = 0.40$ , $\beta = 0.60$ and $\beta = 0.80$ are 1.193g, 1.623g, and 2.379g, respectively. In general, the parameter $\beta$ strongly influences the required capacity. However, it can be observed that the percentage variations over time are similar for the considered cases, because $\beta$ influences a similar manner the contributions to the failure rate due to the two sources.

### *3.5.4 Impact of recursive properties ( $\alpha$ )*

The third important parameter which can affect the required capacity in time-dependent case is $\alpha$ . As mentioned before, $\alpha$ is the aperiodicity parameter describing the recursive properties and influencing the trend of the hazard rate $r_T(t)$ over time. In this group of analyses, three values of $\alpha$ are considered: $\alpha = 0.40$ , $\alpha = 0.50$ and $\alpha = 0.60$ (Akinchi et al. 2009). Values larger than 0.60 have not been investigated because they provide a negative slope and a discussion about its physical meaning is still open.

The variation in time of required capacity has been analysed by reporting results relevant to the same instants considered in the previous sections, i.e. $t_{0.5} = 139yr$ , $t_{1.0} = 185yr$ , and $t_{2.0} = 371yr$ . This makes it possible to compare outcomes concerning variation of $\alpha$ with previous results but it must be recalled that the previous correspondence between instants and hazard rates holds for the reference case only $\alpha = 0.50$ , and it is now lost for the other values of $\alpha$ .

Figure 3.15 shows how the hazard rate varies in time for the three different values of $\alpha$ . The values of elapsed times considered in the comparison are highlighted. It can be observed that the three curves intersect approximately at 250*yr*, so, for elapsing times lower than 250*yr* the hazard rate provided by $\alpha = 0.40$ is lower than the reference hazard rate, and higher for $\alpha = 0.60$ . Conversely, the hazard rate for $\alpha = 0.40$ is higher than the reference case, and lower for $\alpha = 0.60$ when elapsed time is larger than 250*yr*. It is worth noting that the case $\alpha = 0.60$ cannot attain a hazard rate value twice the reference hazard rate $r_{0L} = 0.0028$ .

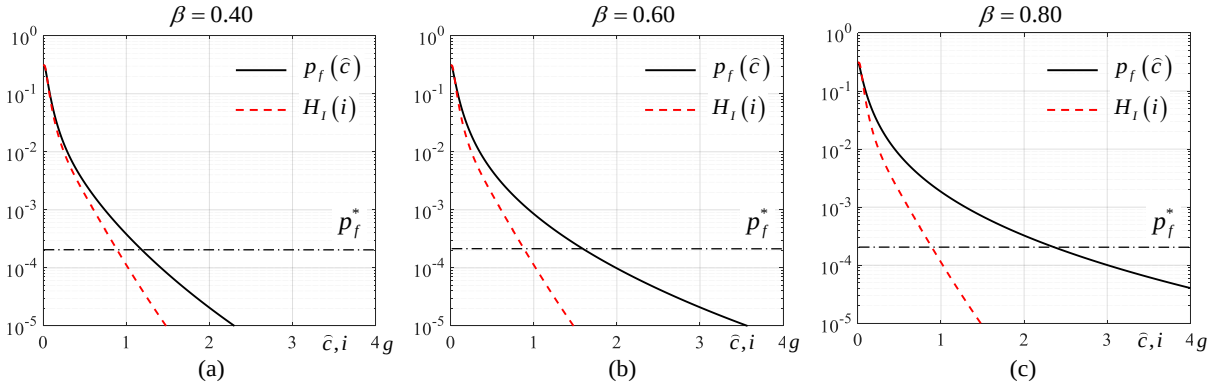



***Figure 3-13***. *Analysis results of combined source in terms of* $H_I(i)$ *and* $p_f(\hat{c})$ *for different hazard rates, considering* $\alpha = 0.50$, $d = 15\text{km}$ *and (a)* $\beta = 0.40$, *(b)* $\beta = 0.60$, *(c)* $\beta = 0.80$.

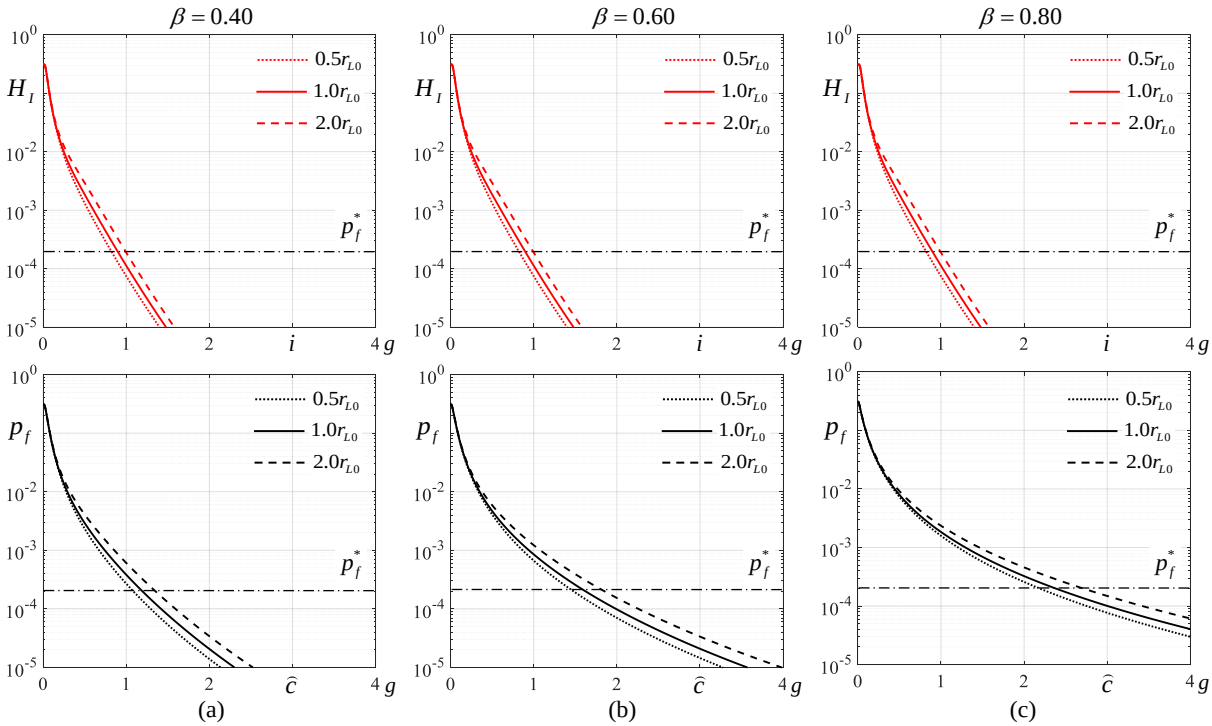



***Figure 3-14***. *Analysis results of combined source for different hazard rates considering* $\alpha = 0.50$, $d = 15\text{km}$ *and (a)* $\beta = 0.40$, *(b)* $\beta = 0.60$, *(c)* $\beta = 0.80$; *first row reports* $H_I(i)$ *curves and second row* $p_f(\hat{c})$ *curves.*

***Table 3-4***. *Required capacity for the target failure rate considering* $d = 15\text{km}$ *and* $\alpha = 0.50$ *for different values of* $\beta$ *and different times elapsing since the last event.*

| | *Capacity* $\hat{c}$ *(g)* | | | $\Delta\hat{c}$ *(%)* | |
|---|---|---|---|---|---|
| $\beta$ | $0.5r_{L0}$ $r_T(139yr)$ | $1.0r_{0L}$ $r_T(185yr)$ | $2.0r_{0L}$ $r_T(371yr)$ | $0.5r_{L0}$ $r_T(139yr)$ | $2.0r_{0L}$ $r_T(371yr)$ |
| *0.40* | *1.084* | *1.193* | *1.350* | *-9.168* | *13.149* |
| *0.60* | *1.478* | *1.623* | *1.851* | *-8.919* | *14.104* |
| *0.80* | *2.186* | *2.379* | *2.709* | *-8.095* | *13.888* |

A comparison between the hazard rate curves and the failure rate curves for different values of $\alpha$, considering different elapsed time $t$, is reported in Figure 3.16, and numerical results are reported in Table 3.5. For the minimum value of elapsed time $t_{0.5}=139yr$ the required capacity goes from $\hat{c}$ = 1.380 g to $\hat{c}$ = 1.563g, passing from $\alpha=0.40$ to $\alpha=0.60$. In the case $t_{1.0}=185yr$ the variation of capacity is limited, while for the case $t_{2.0}=371yr$ the required capacity decreases from $\hat{c}$ =1.932g to $\hat{c}$ = 1.794g, varying $\alpha$ from 0.40 to 0.60. It can be observed that the variation in capacity across different $\alpha$ values is more pronounced for $t_{0.5}$ and $t_{2.0}$ than $t_{1.0}$; this arise from the effects that parameter $\alpha$ produces in the hazard rate curve (Figure 15).

Finally, Table 3.5 also reports the numerical values of variation $\Delta\hat{c}=\left(\hat{c}-\hat{c}_{ref}\right)/\hat{c}_{ref}$. It can be observed that these variations are quite limited in both cases ($\alpha=0.60$ and $\alpha=0.40$) and, lower than variations observed in previous sections. The structural design is not very sensitive to this parameter, at least for the inter-arrival times considered in this section. Based on Figure 3.15, larger variations are expected for large time intervals (greater than 600yr) in the case of $\alpha=0.40$

.

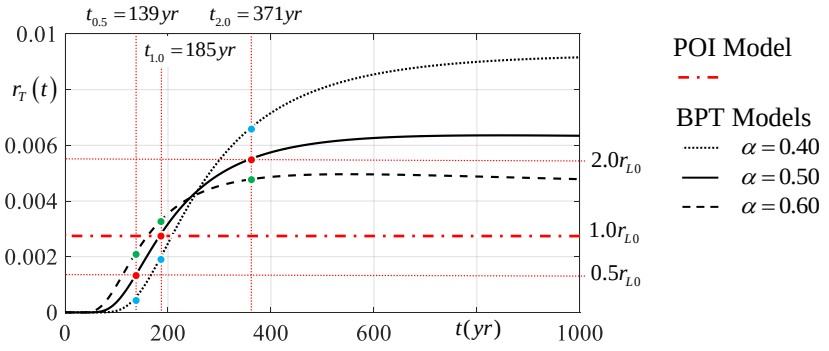



***Figure 3-15****. Hazard rate for BPT model and POI model with* $T_R=358yr$ *considering three values of* $\alpha$.

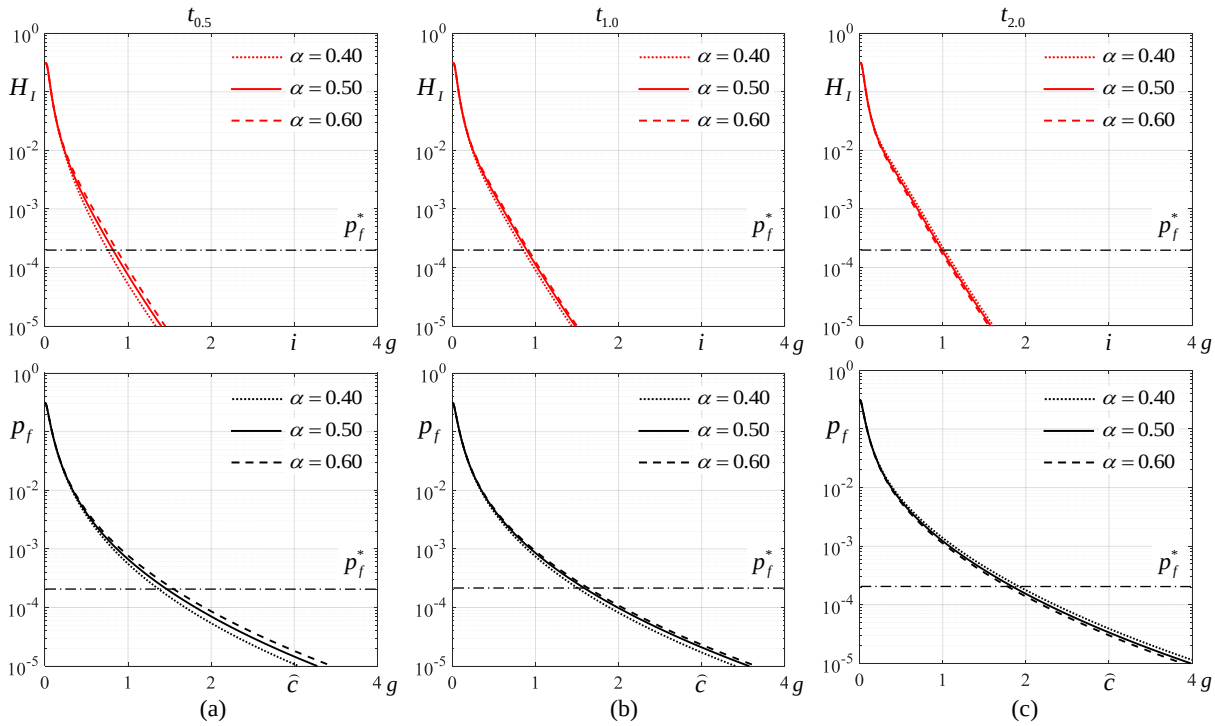



***Figure 3-16****. Analysis results of combined source for different hazard rates considering* $\bar{d}$ *=15 km,* $\beta = 0.60$ *and (a)* $\alpha = 0.40$ *, (b)* $\alpha = 0.50$ *and (c)* $\alpha = 0.60$ *; first row reports hazard curves* $H_I(i)$ *and second row* $p_f(\hat{c})$ *curves.*

***Table 3-5****. Required capacity considering* $\bar{d}$ *=15 km and* $\beta = 0.60$ *for different of* $\alpha$ *and different times elapsing since the last event.*

| | *Capacity* $\hat{c}$ *(g)* | | | $\Delta\hat{c}$ *(%)* | |
|---|---|---|---|---|---|
| $r_{L0}(t)$ | $\alpha$ *= 0.40* | $\alpha$ *= 0.50* | $\alpha$ *= 0.60* | $\alpha$ *= 0.40* | $\alpha$ *= 0.60* |
| $t_{0.5}$ *= 139yr* | *1.380* | *1.478* | *1.563* | *-6.607* | *5.739* |
| $t_{1.0}$ *= 185yr* | *1.540* | *1.623* | *1.669* | *-5.076* | *2.859* |
| $t_{2.0}$ *= 371yr* | *1.932* | *1.851* | *1.794* | *4.366* | *-3.092* |

## 3.6 Conclusions

The impact of Time-Dependent Seismic Hazard on design capacity has been assessed in this study by evaluating the strength required by the structure (seismic capacity) for different time intervals elapsing from the last event. “Seismic capacity” is understood here as the capacity required to provide a fixed reliability level, measured by the failure rate. Two different seismic scenarios have been investigated. In the first case, a point-source has been considered and results concerning different three site-to-source distances $\bar{x}$, and capacity dispersion $\beta$ have been discussed. The case study presented concerns the Paganica fault, located in Central Italy. In the second case, a combined source consisting of an area-source producing medium-low intensity earthquakes, whose recurrence was described by the no-memory POI model, and a

line-source producing high intensity earthquakes with a periodic recurrence described by the time-dependent BPT model has been considered. Results concerning three different site to line source distances $d$, capacity dispersions $\beta$ and aperiodicity parameters $\alpha$, describing the recursive properties of the hazard rate over time, have been investigated.

Based on the results, the following conclusions can be drawn.

- The ratio between the required capacity based on time-dependent model and the required capacity based on non-time-dependent model span a large interval and varies from values notably smaller than 1 for short inter-arrival times to value notably larger than 1 for long time intervals. Variations span the range [0, 1.32] in the case study with point source, lowest and highest values concern inter-arrival times equal to 0 and $2T_R$, respectively Variations in the range [0.84, 1.23] have been observed in the parametric analysis involving combined sources. Extreme values concern the case where the time dependent line source is located at a distance $d = 7.5km$ and the extreme values of inter-arrival times $139y$ and $371y$, are considered, in conjunction with the intermediate values of dispersion $\beta = 0.6$ and aperiodicity parameter $\alpha = 0.5$ Results reveals that the time-recurrence properties of seismic sources exhibiting a periodic behaviour cannot be neglected in the structural design, also in the case of combined sources.
- Required capacity rapidly varies for time interval shorter than the balance time $t_{1.0}$, at which time-dependent and non-time-dependent hazards provide the same hazard rate, whilst it weakly varies for time intervals longer than the return period; Furthermore, high variations in hazard rates do not translate into proportional variations in the capacity required to ensure the target failure rate. Concerning structural design, special attention should be payed to source with time elapsing from last event shorter then return period and further investigation are required about the reliability of time-dependent models because notable reduction of capacity may be involved.
- In the case of point-source, (BPT model, $T_R = 750yr$), required capacity drops to 0 at a critical instant, $\bar{t} = 215yr$. This is due to the decay properties of the inter-arrival time distribution function, widely used in previous studies. This means that no seismic capacity is required to the structure in a period long $t_{1.0}$, after the last event.
- The required capacity is notably influenced by the capacity dispersion and it increases more and more as the elapsed time increases. In particular, the detrimental effect of the

dispersion cannot be neglected for $\beta>0.4$. So that, the reduction of uncertainties in the structural model predicting the seismic response may play an important role in limiting the time-dependent failure rate and this is of special importance in the analysis of existing constructions.

- The impact of time-dependent hazard on structural design is mitigated in the more realistic case study where both time-dependent sources and non-time-dependent sources are present. The time-dependent behaviour can be neglected in the design when the time-dependent source is sufficiently far (30 km in the case study).
- The periodicity parameter provides a moderate influence of structural design (lower than 10% in the parametric analysis) but it is interesting to observe that a strong periodicity (small $\alpha$) reduces the required capacity for short time intervals and a weak periodicity (large $\alpha$) increases the required capacity. The opposite when large time-interarrival times are considered.
- Numerical results reported in the research only concern some case studies where a limited set of realistic scenarios (source recursive properties, propagation laws and structural properties) have been considered. They show that time-dependent hazard may notably influence the structural design and provide some information about the most important parameters. However, a deeper investigation is required to provide more precise suggestions about structural design rules for constructions in seismic prone area

## References

Anagnos T., Kiremidjian, A.S., (1988). A Review of Earthquake Occurrence Models for Seismic Hazard Analysis", Probabilistic Engineering Mech., 3 (1), 3–11.

Akinci A., Galadini F., Pantosti D., Petersen M., Malagnini L., Perkins D., (2009). Effect of Time Dependence on Probabilistic Seismic-Hazard Maps and Deaggregation for the Central Apennines, Italy, Bulletin of the Seismological Society of America, 99( 2A), 585–610

ASCE/SEI 7-16 (2017). Minimum design loads and associated criteria for buildings and other structures. American Society of Civil Engineers, Reston

CEN (2006). EN 1990:2006 Eurocode - Basis of structural design. European Committee for Standardization, Brussels.

Chan C.-H., Wu Y.-M., Cheng C.-T., Lin P.-S., Wu, Y.-C., (2013). Time-dependent probabilistic seismic hazard assessment and its application to Hualien city, Taiwan, Nat. Hazards Earth Syst. Sci., 13, 1143–1158, doi:10.5194/nhess-13-1143-2013.

Chiaraluce L., Valoroso L., Piccinini D., Di Stefano R., De Gori P., (2011). The anatomy of the 2009 L'Aquila normal fault system (central Italy) imaged by high resolution foreshock and aftershock locations. Journal of Geophysical Research, 116, B12311. https://doi. org/10.1029/2011JB008352

Cornell C.A., Jalayer F., Hamburger R.O., Foutch D.A., (2002). Probabilistic basis for 2000 SAC Federal Emergency Management Agency steel moment frame guidelines. J Struct Eng 128(4):526–533

Dolsek M., (2009). Incremental dynamic analysis with consideration of modeling uncertainties. Earthquake Engng Struct. Dyn. 2009; 38:805–825

Ellsworth W.L., Matthews M.V., Nadeau R.M., Nishenko S.P., Reasenberg P.A., Simpson R.W., (1999). A physically-based earthquake recurrence model for estimation of long-term earthquake probabilities, USGS, Open-File Report 99–522.

Fajfar P., (2018). Analysis of seismic provisions for buildings: past, present, and future. Bulletin of Earthquake Engineering, 16, 2567–2608.

FEMA (2009a). NEHRP recommended seismic provisions for new buildings and other structures (FEMA P750). Federal Emergency Management Agency

FIB (2012). Probabilistic performance-based seismic design, Bulletin 68, International federation of structural concrete, Lausanne, CH

Fiorini E., Bazzurro P., Silva V., (2014). Preliminary Results of risk targeted design maps for Italy, second European conference on earthquake engineering and seismology, Istanbul.

Franchin P., Petrini F., Mollaioli F., (2018). Improved risk-targeted performance-based seismic design of reinforced concrete frame structures. Earthq Eng Struct Dyn 47(1):49–67

Galli P., Giaccio B., Messina P., Peronace E., (2011). Paleoseismology of the L'Aquila faults (Central Italy, 2009 Mw 6.3 earthquake). Clues on active fault linkage. Geophys. J. Int. 187, 1119–1134.

Garcia-Aristizabal A., Marzocchi W., Fujita E., (2011). A Brownian model for recurrent volcanic eruptions: an application to Miyakejima volcano (Japan). Bulletin of Volcanology, 74(2), 545–558. doi:10.1007/s00445-011-0542-4

Gkimprixis A., Tubaldi E., Douglas J., (2019). Comparison of methods to develop risk-targeted seismic design maps, Bulletin of Earthquake Engineering, doi:10.1007/s10518-019-00629-w.

Gulia L., Tormann T., Wiemer S., Herrmann M., Seif S., (2016). Short-term probabilistic earthquake risk assessment considering time-dependent values. Geophysical Research Letters, 43(3), 1100–1108. doi:10.1002/2015gl066686.

Hagiwara, T. (1974), Distribution of seismic intensity of the great earthquakes in 1854 and 1707, Rep. Coord. Comm. Earthq. Prediction 12, 143-145.

Jalalalhosseini S. M., Zafarani H., Zare M., (2017). Time-dependent seismic hazard analysis for the Greater Tehran and surrounding areas, J Seismol, 22(1), DOI: 10.1007/s10950-017-9699-4.

Kameda, H. and Y. Ozaki (1979). A renewal process model for use in seismic risk analysis, Mem. Fac. Eng., Kyoto Univ. XLI, 11- 35.

Kennedy R.P. , Short S.A., (1994). Basis for seismic provisions of DOE-STD-1020. Rep. No. UCRLCR-111478, Lawrence Livermore National Laboratory, Livermore, Calif., and Rep. No. BNL-52418, Brookhaven National Laboratory, Upton, N.Y

Kennedy, R.P., (2011). Performance-goal based (risk informed) approach for establishing the SSE site specific response spectrum for future nuclear power plants. Nuclear Engineering and Design, 241(3), 648–656. doi:10.1016/j.nucengdes.2010.08.001

Kramer S.L., (1996). Geotechnical Earthquake Engineering. Prentice Hall: Upper Saddle River, NJ, 1996.

Matthews M V., Ellsworth W. L., Reasenberg P. A., (2002). A Brownian model for recurrent earthquakes, Bulletin of the Seismological Society of America, 92(6), 2233–2250.

Mousavi M., Salhei M., (2018). Temporal distribution of earthquakes using renewal process in the Dasht-el-Bayaz region. J. Seismol, 22, 153-159.

Pace B., Peruzza L., Lavecchia G., Boncio P., (2006). Layered seismogenic source model and probabilistic seismic-hazard analyses in central Italy, Bulletin of the Seismological Society of America, 96(1), 107-132.

Petersen M.D., Cao T., Campbell K.W., Frankel A.D., (2007). Time-independent and Time-dependent Seismic Hazard Assessment for the State of California: Uniform California Earthquake Rupture Forecast Model 1.0, Seismological Research Letters 78(1), 99-109.

Polidoro B., Iervolino I., Chioccarelli E., (2013). Models and issues in history-dependent mainshock hazard. Proc. of 11th Conf. on Structural Safety and Reliability, ICOSSAR '13, New York, US, June 16-20 2013.

Porter K.A., (2003). An overview of PEER's performance-based earthquake engineering methodology. In: Proc. 9th international conference on applications of statistics and probability in civil engineering

Reid H.F., (1910). The Mechanics of the Earthquake, The California Earthquake of April 18, 1906, Report of the State Investigation Commission, Vol.2, Carnegie Institution of Washington, Washington, D.C.

Sabetta, F. and A. Pugliese (1987). Attenuation of peak horizontal acceleration and velocity from Italian strong-motion records, Bull Seism. Soc. Am. 77, 1491-1513.

Sabetta F., Pugliese A., (1996). Estimation of response spectra and simulation of non-stationarity earthquake ground motion, Bulletin of the Seismological Society of America 86: 337 352.

Schwarz D.P., Coppersmith K.J., (1984). Fault behavior and characteristic earthquakes: examples from the Wasatch and San Andreas fault zones. Journal of Geophysical Research 89(B7), 5681–5697.

Scozzese F., Tubaldi E., Dall'Asta A., (2020). Assessment of the effectiveness of Multiple-Stripe Analysis by using a stochastic earthquake input model. Bull Earthquake Eng. https://doi.org/10.1007/s10518-020-00815-1.

Stafford P.J., Strasser F.O., Bommer J.J., (2008). An evaluation of the applicability of the NGA models to ground-motion prediction in the Euro-Mediterranean region. Bull Earthquake Eng 6, 149–177. https://doi.org/10.1007/s10518-007-9053-2

Takahashi Y., Der Kiureghian A., Ang A. H-S., (2004). Life-cycle cost analysis based on a renewal model of earthquake occurrences, Earthquake Engng Struct. Dyn., 33, 859-880. DOI: 10.1002/eqe.383.

Tondi E., Cello G., (2003). Spatiotemporal Evolution of the Central Apennines Fault System (Italy). Journal of Geodynamics, 36, 113-128.

Vere-Jones, D. and Davies, R. B. (1966). Statistical Survey of Earthquakes in the Main Seismic Region of New Zealand, Part 2: Time Series Analysis, NZ J. Geol. Geoph., 9, 251-284.

Wesnousky S.G., (1994).The Gutenberg–Richter or characteristic earthquake distribution, which is it? Bulletin of the Seismological Society of America; 84(6):1940–1959.

Working Group on California Earthquake Probabilities. Earthquake probabilities in San Francisco Bay region: 2000 to 2030. A summary of findings. U.S. Geological Survey Report 99-517, U.S. Geological Survey. 1999

# 4 Structural Capacity in the Case of Time-dependent Seismic Hazard Analysis Using Different GMPEs in Different Periods

## Abstract

Time-dependent hazard models have been used by researchers to evaluate the probability of earthquake occurrence by elapsing time. Time-dependent response of structures, on the other hand, has attracted the structural engineers' concern recently; however, not enough studies have been conducted on this important subject matter to date.

In this study, the structural capacity in the case of time-dependent seismic hazard is evaluated and discussed. To aim the study objective, a combined source consisting of an area source (generating medium to low-intensity earthquakes by a POI model) and a line source (generating high-intensity earthquakes by a time-dependent BPT model) is considered. The GMPE proposed by Ambraseys et al. was used in order to consider the effect of site soil conditions, period and time elapsed since the last event on both demand and capacity values. Furthermore, the obtained results are compared by the results of analysis performed using Sabetta and Pugliese to figure out the influence of GMPE on the outcomes.

Results showed that the difference between demand and capacity for sites located on soil is higher than those located on rock. Moreover, in very short or very long periods, elapsing time does not affect the results significantly. More importantly, the changes of capacity are not the same as those of demand are slightly higher. Finally, The GMPE affects the results remarkably which shows the importance of appropriate GMPE selection for the region of the study.

## 4.1 Introduction

Estimating the level for ground motion with a specified probability of exceedance could be done through probabilistic seismic hazard analysis (PSHA) [5]. The rate of exceedance is obtained in PSHA by considering parameters namely magnitude, site-to-source distance and ground motion propagation determined by ground motion prediction equations (GMPE). Although the influence of the local site on PSHA is seldomly considered in GMPE (e. g., (Ambraseys, Simpson et al. 1996)), it is not taken into account in the PSHA formulation directly (Bazzurro and Cornell 2004). Soil classification of the area where buildings are located on, however, has a crucial impact on both hazard and structural response (Bazzurro and Cornell 2004, Shahbazi, Mansouri et al. 2018, Nabilah, Koh et al. 2019, Ezzelarab, Hassoup et al. 2021).

In PSHA, it is generally assumed that the soil classification of the site is the same as the one in the site considered for generating GMPE. This assumption could lead to inaccuracy of the results caused by ignoring soil condition of the site (Bazzurro and Cornell 2004). It is worth explaining that the main reason for the aforementioned inaccuracy is due to soil nonlinearity in the soil response (Bazzurro and Cornell 2004). As an example, the values obtained by two different GMPEs, (Sadigh, Chang et al. 1997) and (Boore, Joyner et al. 1997) are the same while different assumptions of soil classification are considered for each of which. As noted previously, a few GMPEs are developed in which site soil type is considered (e.g., (Ambraseys, Simpson et al. 1996) , and (Abrahamson and Silva 1997)). One method which could be used for considering soil effect is multiplying the hazard at the soil surface by the bedrock hazard (in the form either of uniform hazard spectra or of hazard curves) using a deterministic amplification function. This technique generates surface ground-motion levels whose exceedance rates are unknown, non-uniform, inconsistent across frequency, and generally nonconservative. Similarly, soil surface ground motions with unknown exceedance rates are in general obtained if one multiplies the bedrock hazard by the average National Earthquake Hazards Reduction Program (NEHRP) amplification factors (Schnabel, Seed et al. 1972). The influence of soil parameters on hazard and structures' response has been evaluated more precisely in a limited number of studies recently.

(Bazzurro and Cornell 2004) developed two novel techniques for addressing the above-mentioned shortcomings: (i) adding a term, representing the soil response, in the current attenuation laws for rock ground motion, and (ii) adding a correction factor, referring the uncertainty in the amplification of the site soil at an appropriate level of bedrock ground

shaking. The proposed methods exhibited accuracy for sandy and clayey soils. (Ezzelarab, Hassoup et al. 2021) investigated the effect of local soil on the seismic hazard level in a specific area located in Egypt. The different earthquake response spectrum derived for the different varying areas considered in their study, proves the high importance of soil condition on the results. It is also claimed that motion amplification in soil could be up to 4 times than that of rock at high period ranges (Lam, Tsang et al. 2015). Nabilah et al., as an example, concluded that amplification of clayey soils is higher than those are recommended by Eurocode 8 (Eurocode 8 2004, Nabilah, Koh et al. 2019).

As far as structures' response is concerned, soil effect could be also expressed as soil-structure interaction, which has attracted researchers' attention due to its significance (Rajeev and Tesfamariam 2012, Carbonari, Morici et al. 2016, Givens, Mylonakis et al. 2016, Montoya-Noguera and Lopez-Caballero 2018). It could be stated that structures constructed on flexible soil behave completely different in terms of deformation, stiffness and seismic parameters when compared to those located on rock (Zhao, Wang et al. 2017, Anvarsamarin, Rofooei et al. 2018, Shahbazi, Mansouri et al. 2018).

The influence of time-dependent hazard on structural response considering soil classification, period and GMPE selection has not been assessed precisely to date. The main objective of this report, therefore, is to find out the trends of changing hazard and capacity required to a target failure probability in the case of time-dependent hazard model. Two main effective parameters are considered: (a) different soil conditions, and (b) different Ground Motion Prediction Equations (GMPE).

Soil types (i. e., very soft or soft soil, stiff soil and rock), as has been proven in previous studies, changes both hazard and capacity significantly. However, the rate of change depends heavily on various parameters such as period, GMPE and time elapsed since the last event (earthquake). This research aims at realizing the impact of above-mentioned parameters on both demand and capacity for the first time.

The results of this research are going to be presented in two main parts. In the first section, the results of a time-dependent combined source in different soil conditions are presented and discussed. In the second part, the results of two GMPE on the time-dependent combined source are compared and evaluated.

## 4.2 Methodology

The methodology introduced in the first section, is also used in the present section for exploring the effect of soil classification, period and GMPEs on both demand and capacity.

Briefly explained, in the methodology, the hazard at the considered site is initially evaluated, considering the recursive and intensity properties of the source, as well as the propagation of the ground motion, from source to site. Separately, the structure capacity is obtained by computing the fragility curve. Finally, the failure rate obtained is compared with the target value.

Simply stated, PSHA and fragility curves are implemented to obtain capacity of a structure for a target probability failure. Otherwise mentioned, "seismic capacity" is defined as the capacity required to provide a fixed reliability level, measured by the failure rate. Details of the methodology are not represented here for the sake of shortness and readers are referred to chapter 3.

The source considered in this study is a combination of two sources: (a) an *area-source* producing medium-low intensity earthquakes whose recurrence has been described by the no-memory POI model and (b) a *line-source* producing high intensity earthquakes with a periodic recurrence described by the BPT model (Fig. 4.1). Only the line source was considered time-dependent since it has been proven that periodical recurrence of high-magnitude earthquakes follows the introduced time-dependent hazard (e.g., BPT) while occurrence probability of low- to medium- magnitude earthquakes could be accurately simulated through time-independent models like Poisson model (POI) (Schwartz and Coppersmith 1984, Wesnousky 1994, Kramer 1996, Tondi and Cello 2003).

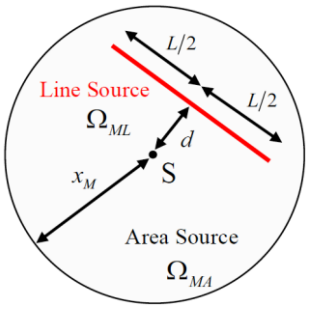



***Figure 4-1.*** *combined source considered in this study*

It is noteworthy that $d$ (site to line source distance) and $x_M$ (radius of area source) are selected so that the case study source reflects a realistic seismic source (respectively *15 km* and *50 km*).

Dispersion of structural capacity, $\beta$, was also considered 0.60 which is in accordance FEMA P-750, ASCE/SEI 7–16 (FEMA 2009a, ASCE 7-10 2013).

## 4.3 Different soil conditions

In order to assess the effects of soil types on both demand and capacity, a GMPE which considers soil type on the prediction of wave propagation should be used in the analysis process. In the present study, the GMPE proposed by (Ambraseys, Simpson et al. 1996). is used (Eq. 4.1) _ENREF_1 accordingly:

$$\log(y) = C_1' + C_2 M + C_4 \log(r) + C_A S_A + C_S S_S + \sigma P \qquad (4.1)$$

where $y$ is the parameter being predicted, in this case peak horizontal ground acceleration in ***g,*** $M_s$ is the surface wave magnitude, and:

$$r = \sqrt{d^2 + h_0^2} \qquad (4.2)$$

where $d$ is the shortest distance from the station to the surface projection of the fault rupture, in km, and $h_o$ is a constant to be determined with $C_1$, $C_2$, $C_3$ and $C_4$**.** The standard deviation of log(y) is *0,* and the constant $P$ takes a value of 0 for mean values and 1 for 84-percentile values of log(y).

The term $h_o$ in equation (4.2) accounts for the fact that the source of the peak motion is not necessarily the closest point on the surface projection of the fault, or from the epicenter, and it does not represent explicitly the effect of the depth on the acceleration.

As given in Eq. (4.1), this GMPE gives different values (here peak horizontal ground acceleration or PGA) in different soil conditions due to changing two parameters: $S_A$ and $S_S$ which vary as bellow for various soil types:

- Very soft or soft soil ($Vs \leq 180$ or $180 < Vs \leq 360$ m/s): $S_A=0$, $S_S=1$;
- Stiff soil ($360 < Vs \leq 750$ m/s): $S_A=1$, $S_S=0$;
- Rock ($Vs > 750$ m/s): $S_A=0$, $S_S=0$;

In the case of using spectrum ordinates, the parameters of the attenuation equation given in Eq. (1) changes by changing period as provided in Table 4.1:

***Table 4-1****. Changes of coefficients of Eq. (1) and (2).*

| *T* | *C'1* | *C2* | *h0* | *C4* | *CA* | *CS* | *σ* |
|---|---|---|---|---|---|---|---|
| *0.1* | *-0.84* | *0.219* | *4.5* | *-0.954* | *0.078* | *0.027* | *0.27* |
| *0.2* | *-1.21* | *0.284* | *4.2* | *-0.922* | *0.135* | *0.142* | *0.27* |
| *0.4* | *-1.94* | *0.377* | *3.6* | *-0.888* | *0.139* | *0.172* | *0.31* |
| *0.6* | *-2.49* | *0.438* | *2.5* | *-0.881* | *0.124* | *0.212* | *0.32* |
| *0.8* | *-2.86* | *0.485* | *3.7* | *-0.925* | *0.127* | *0.218* | *0.32* |
| *2* | *-3.79* | *0.503* | *3.2* | *-0.728* | *0.101* | *0.182* | *0.32* |

Fig. 4.2 compares the changes of parameters of Ambraseys GMPE when period alters. Based on the curves illustrated in this figure, parameters $C'_1$ and $h_0$ change considerably by changing T while other parameters remain almost the same.

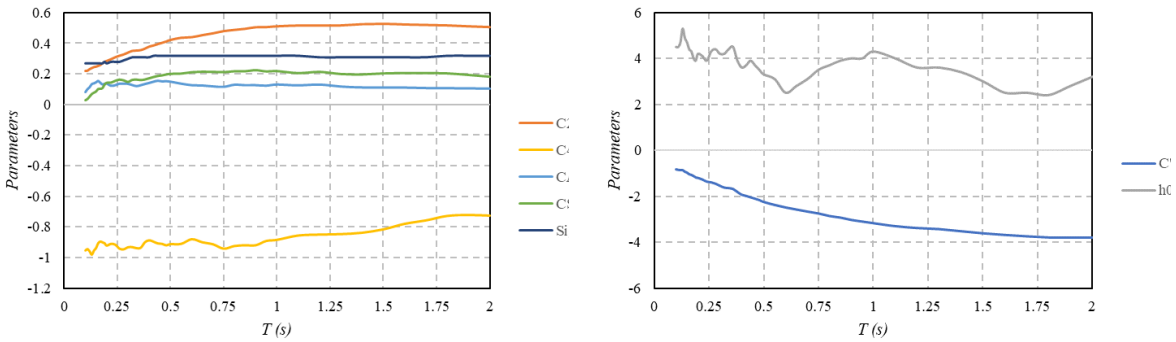



***Figure 4-2****. effect of T(s) on the parameters used in Eq. (1)*

As mentioned before, in the Ambraseys GMPE $S_A$ and $S_S$ vary for different soil types while other parameters remain constant. As far as different values of period are concerned, changes of other parameters are as explained previously and presented in Table 4.1. For the sake of shortness, six periods (which reflect the changes in results perfectly) were chosen to be used for obtaining demand and capacity required for a target failure probability: 0.1, 0.2, 0.4, 0.6, 0.8 and 2 s.

In order to consider the changes of demand and capacity by elapsing time, the results are compared in three different times: 139 yr, 185yr and 371 yr (corresponding to $0.5r_{L0}$, $1.0r_{L0}$ and $2.0_{rL}$, respectively).

### *4.3.1 Very soft or soft soil*

Figure 4.3 compares the results corresponding to different periods for a site located on a very soft (or soft) soil; $H_I(i)$ curves for line, area and combined source are depicted in the left columns while the right column presents the curves reporting $H_I(i)$ and $p_f(\bar{c})$ for the combined source.

Considering the $H_I(i)$ curves in the left column, it could be said that: (a) for short periods (0.1 and 0.2) the hazard is affected mostly by the area source rather than the line source, (b) on the contrary, for long periods (>0.8 s), the results of combined source is almost the same as line source and area source could be approximately ignored, and (c) for moderate periods (0.6s<T<0.8s) the contribution of line and area source in the results of the combined source is considerable.

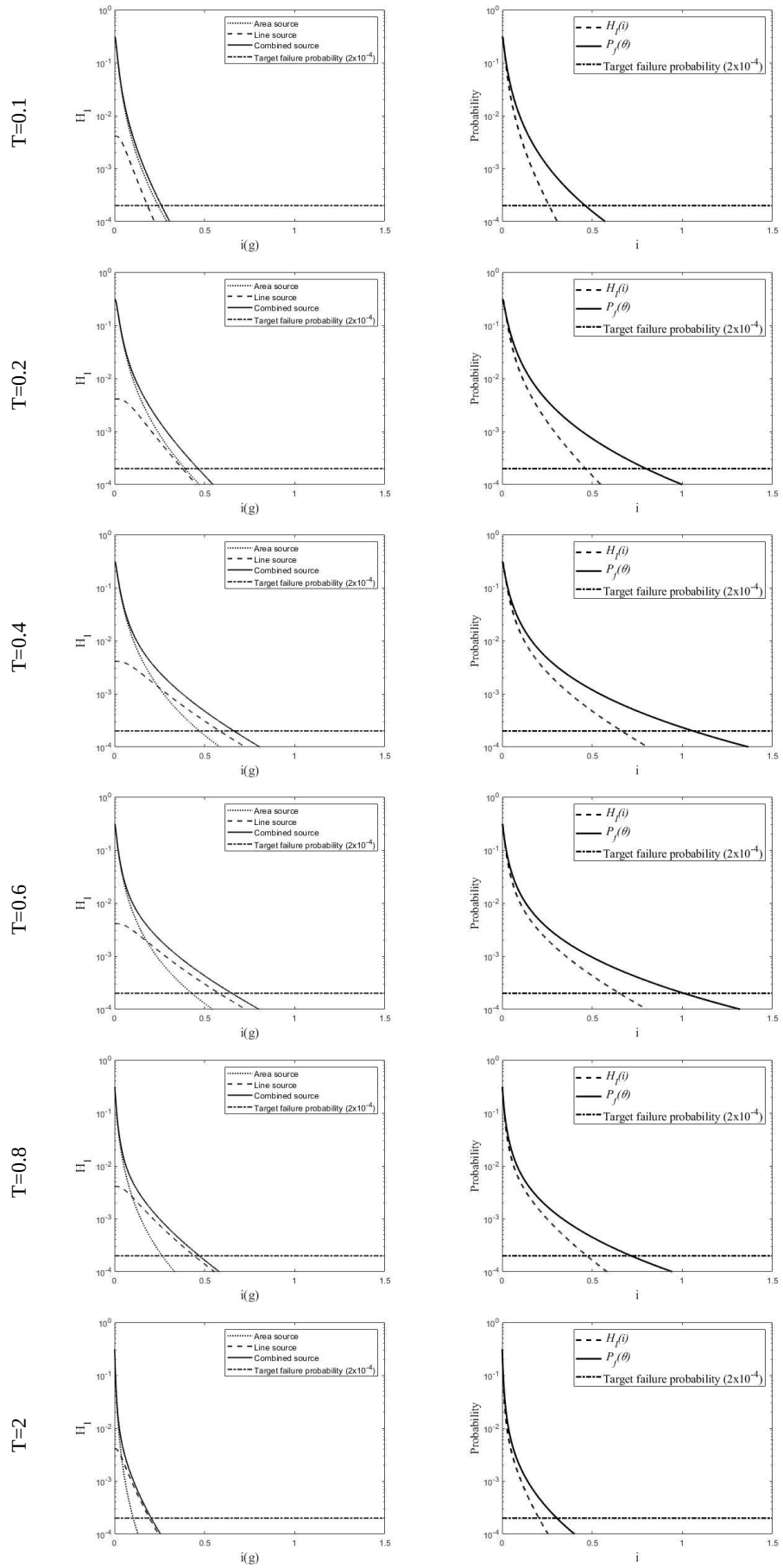



***Figure 4-3****. very soft and soft soil at* $1.0r_{L0}$ *: left column:* $H_I(i)$ *curves for area, line and combined source; right column:* $H_I(i)$ *and* $p_f(\bar{c})$ *curves*

Regarding the curves demonstrated in the right column, it could be noted that capacity values in all the considered periods are higher than the demand values for a target failure probability. The difference between capacity and demand, however, is not constant and changes by altering period. Otherwise noted, for short and long periods capacity and demand difference is less than that of moderate periods.

In order to investigate the effect of time elapsed since the last event, demand and capacity values for a target failure probability (0.0002) for three different times (139yr, 185yr and 371yr) are obtained and illustrated in Fig. 4.4. The left diagram shows the demand for different periods and times whereas the right diagram compares the capacity values. As expected, both demand and capacity increases by elapsing time. However, it should be taken into consideration that the difference of demand and capacity in different times is not constant; in short periods (up to 0.2s) demand and capacity do not change considerably by elapsing time while for higher periods (up to approximately 0.75s) elapsing time affects the result remarkably. For periods higher than 0.75s, the difference between the results of various times commence to decrease.

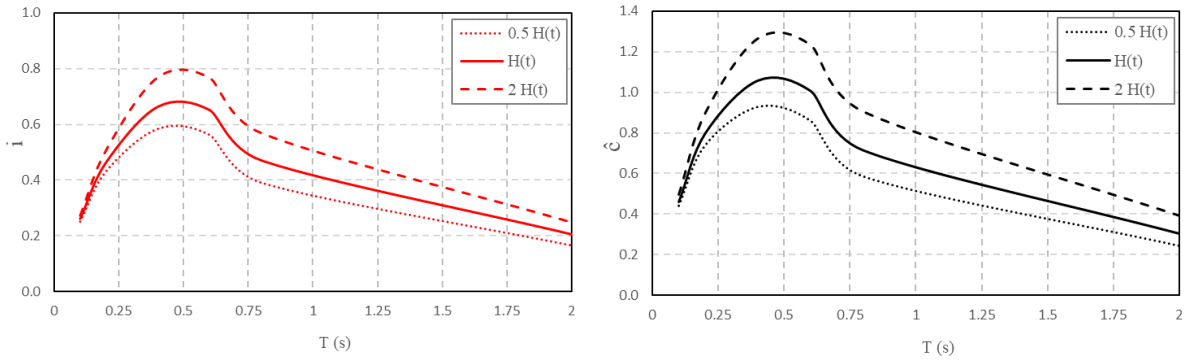



***Figure 4-4.*** *Analysis result of combined source for different hazard rates for a site located in soft or very soft soil; left: demand curves and right: capacity curves*

The other significant point which should be pointed out is that the changes of capacity in different time elapses since the last event is slightly higher than those of demand values (see Fig. 4.5).

A quantitative comparison between demand and capacity values for different periods and times elapsed since the last event is provided in Table 4.2. As was noted before, the changes of capacity values are more than those of demand values in two considered times ($0.5r_{L0}$ *and* $2.0r_{L0}$). It should be also noted that the changes of results for 371yr elapsed since the last event is higher than those of 139yr elapsed since the last event. This difference in the results is due to the changes of time-dependent hazard model (BPT) used in this study.

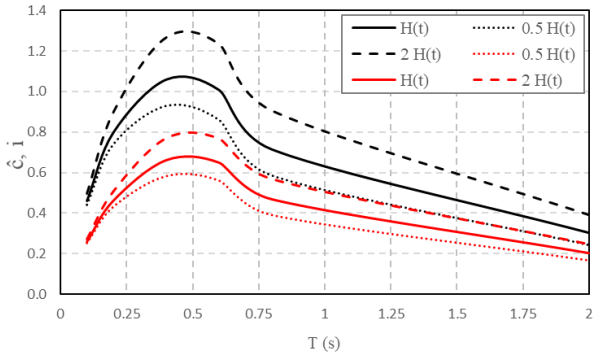



***Figure 4-5.*** *Analysis result of combined source for different hazard rates for a site located in soft or very soft soil*

***Table 4-2.*** *demand and capacity values for different T(s) for a site located on very soft or soft soil.*

| | ***0.5 H(t)*** | | | | ***H(t)*** | | ***2 H(t)*** | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| ***T (s)*** | ***Demand (g)*** | ***Δi (%)*** | ***Capacity (g)*** | ***Δĉ (%)*** | ***Demand (g)*** | ***Capacity (g)*** | ***Demand (g)*** | ***Δi (%)*** | ***Capacity (g)*** | ***Δĉ (%)*** |
| *0.1* | *0.252* | *-3.08* | *0.441* | *-4.34* | *0.260* | *0.461* | *0.274* | *5.38* | *0.495* | *7.38* |
| *0.2* | *0.432* | *-6.70* | *0.739* | *-7.39* | *0.463* | *0.798* | *0.508* | *9.72* | *0.897* | *12.41* |
| *0.4* | *0.584* | *-11.92* | *0.928* | *-12.37* | *0.663* | *1.059* | *0.769* | *15.99* | *1.266* | *19.55* |
| *0.6* | *0.562* | *-13.67* | *0.859* | *-14.78* | *0.651* | *1.008* | *0.767* | *17.82* | *1.234* | *22.42* |
| *0.8* | *0.391* | *-16.81* | *0.585* | *-18.18* | *0.470* | *0.715* | *0.569* | *21.06* | *0.905* | *26.57* |
| *2* | *0.166* | *-18.63* | *0.243* | *-20.07* | *0.204* | *0.304* | *0.249* | *22.06* | *0.390* | *28.29* |

### *4.3.2 Stiff soil*

The left column of Fig. 4.6 depicts hazard curves, $H_I(i)$, for area, line and combined source for different periods; the combined hazard and capacity curves are also compared in the right column.

Like the results obtained in the previous section, the hazard of the combined source is affected mostly by area and line source in short and long periods, respectively. For the moderate periods, however, the hazard due to combined source is influenced by both time-independent area and time-dependent source. It should be also noted that the hazard due to each case (the line, area and combined source) increases by increasing period, then it decreases for periods longer than 0.4s.

In terms of capacity curves, it should be stated that the changes of capacity are not the same as hazard and it is higher than hazard values in all periods. More importantly, the difference between hazard and capacity in moderate periods is larger than short and long periods.

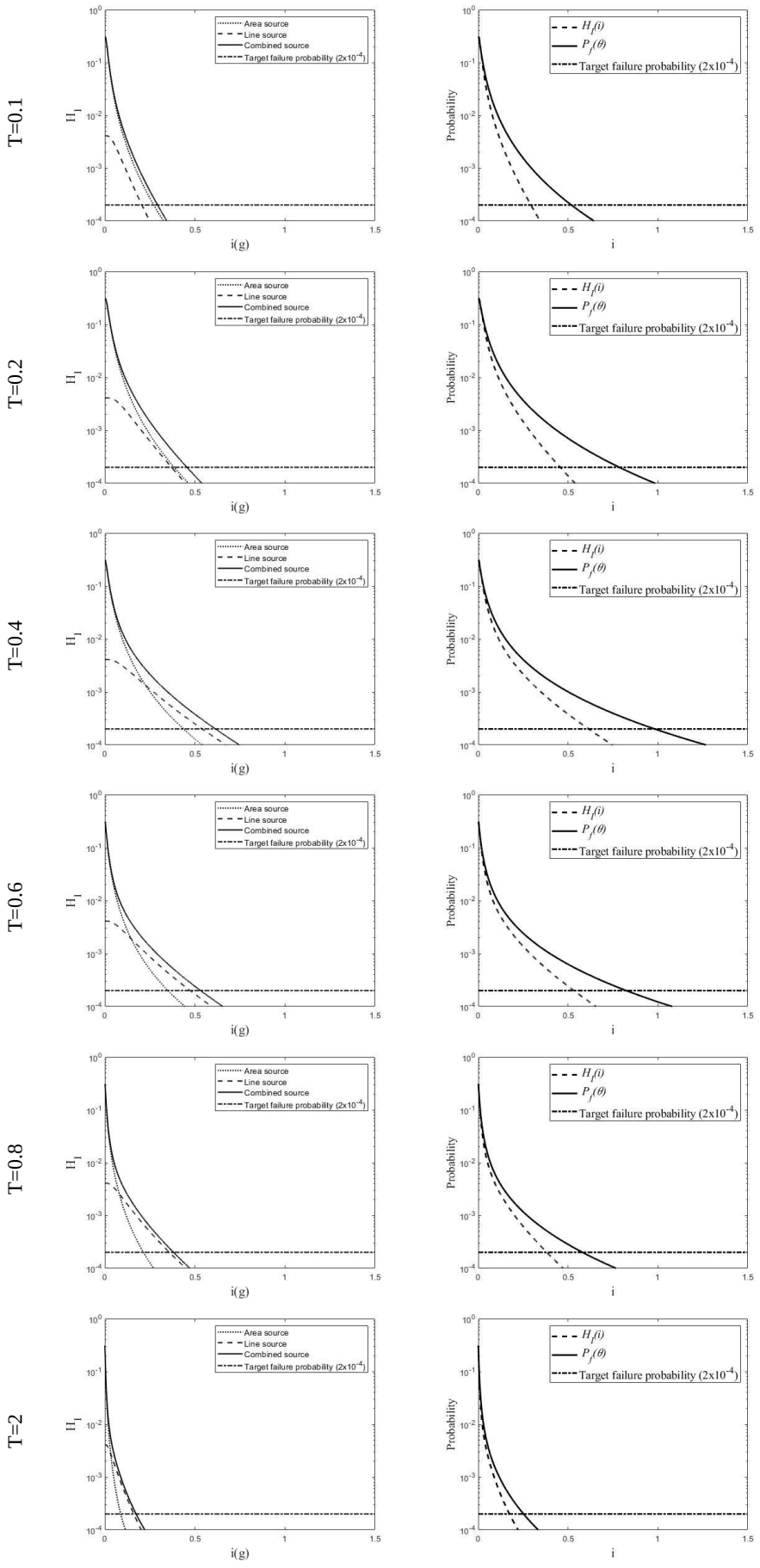



***Figure 4-6.*** *stiff soil at* $1.0r_{L0}$ *: left column:* $H_I(i)$ *curves for area, line and combined source; right column:* $H_I(i)$ *and* $p_f(\hat{c})$ *curves*

Fig. 4.7 compares demand and capacity values for different periods in three different times elapsed since the last event (139yr, 185yr and 371yr). As could be observed in both diagrams, for very short periods (up to 0.2s) elapsing time does not have a notable effect on the results. For longer periods, both demand and capacity increase by elapsing time. It should be taken into account that the highest difference between demand or capacity values at different times belongs to T=0.4s.

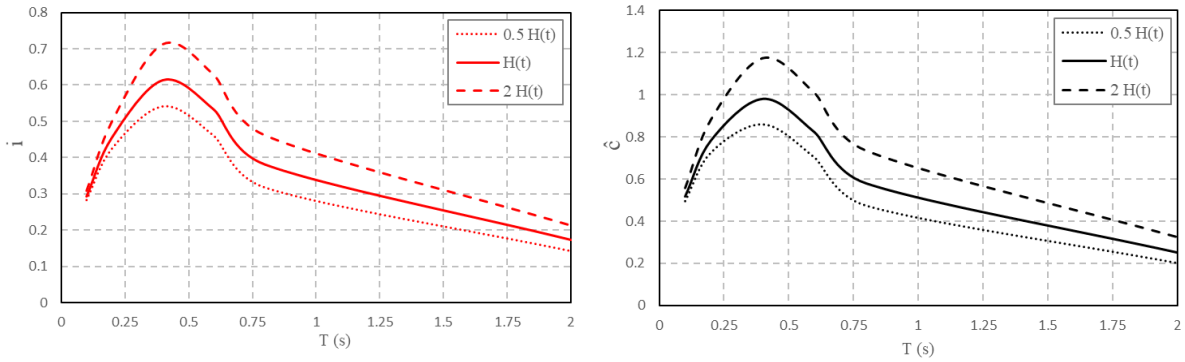



***Figure 4-7.*** *Analysis result of combined source for different hazard rates for a site located in stiff soil; left: demand curves and right: capacity curves*

In order to make a comparison between demand and capacity changes for three different considered times, all the curves are shown together in Fig. 4.8. It is obviously observed that capacity values are more dependent on time rather than demand values. In other words, the difference between capacity values at different times is larger than those of demand values.

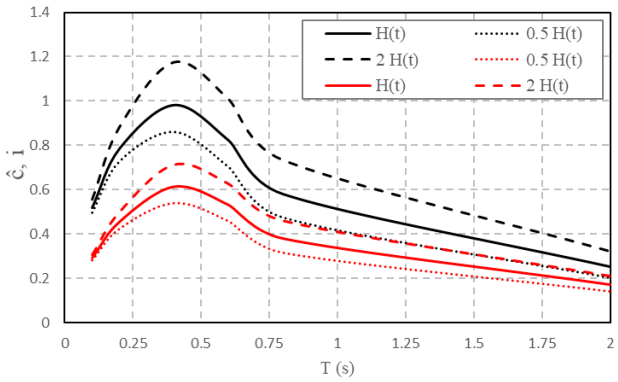



***Figure 4-8.*** *Analysis result of combined source for different hazard rates for a site located in stiff soil*

The changes of demand and capacity required for a target failure probability are provided quantitatively in Table 4.4 for different periods and three considered times.

***Table 4-3.*** *demand and capacity values for different T(s) for a site located on stiff soil*

| | ***0.5 H(t)*** | | | | ***H(t)*** | | ***2 H(t)*** | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| ***T (s)*** | ***Demand (g)*** | ***Δi (%)*** | ***Capacity (g)*** | ***Δĉ (%)*** | ***Demand (g)*** | ***Capacity (g)*** | ***Demand (g)*** | ***Δi (%)*** | ***Capacity (g)*** | ***Δĉ (%)*** |
| *0.1* | *0.283* | *-3.41* | *0.496* | *-4.25* | *0.293* | *0.518* | *0.308* | *5.12* | *0.557* | *7.53* |
| *0.2* | *0.425* | *-6.80* | *0.727* | *-7.39* | *0.456* | *0.785* | *0.5* | *9.65* | *0.883* | *12.48* |
| *0.4* | *0.541* | *-11.89* | *0.86* | *-12.42* | *0.614* | *0.982* | *0.713* | *16.12* | *1.174* | *19.55* |
| *0.6* | *0.459* | *-13.72* | *0.702* | *-14.70* | *0.532* | *0.823* | *0.627* | *17.86* | *1.008* | *22.48* |
| *0.8* | *0.317* | *-16.80* | *0.474* | *-18.28* | *0.381* | *0.58* | *0.461* | *21.00* | *0.734* | *26.55* |
| *2* | *0.142* | *-17.92* | *0.202* | *-19.84* | *0.173* | *0.252* | *0.212* | *22.54* | *0.324* | *28.57* |

### *4.3.3 Rock*

Hazard curves, $H_I(i)$, for three considered sources for different periods and the combined hazard and capacity curves are demonstrated in the left and right columns of Fig. 4.9, respectively.

The trend of demand and capacity changes for a site located on rock is similar to those observed for other soil conditions. The most significant point which should be mentioned is that the hazard due to the combined source is close to the hazard due to each of either line or area source. Regarding capacity curve, it could be said that while there is a notable difference between demand and capacity for short and moderate periods, they are almost the same in long periods (T=2.0 s).

A comparison between demand and capacity curves at different times is made in Fig. 4.10. Like the results obtained for other soil types, both demand and capacity increases by elapsing time and the maximum increase of the mentioned values is in the period T=0.4s. It is also worth noting that due to the BPT hazard model, the difference of the values corresponding 185yr and 371yr since the last event is higher than those of 139r and 185yr.

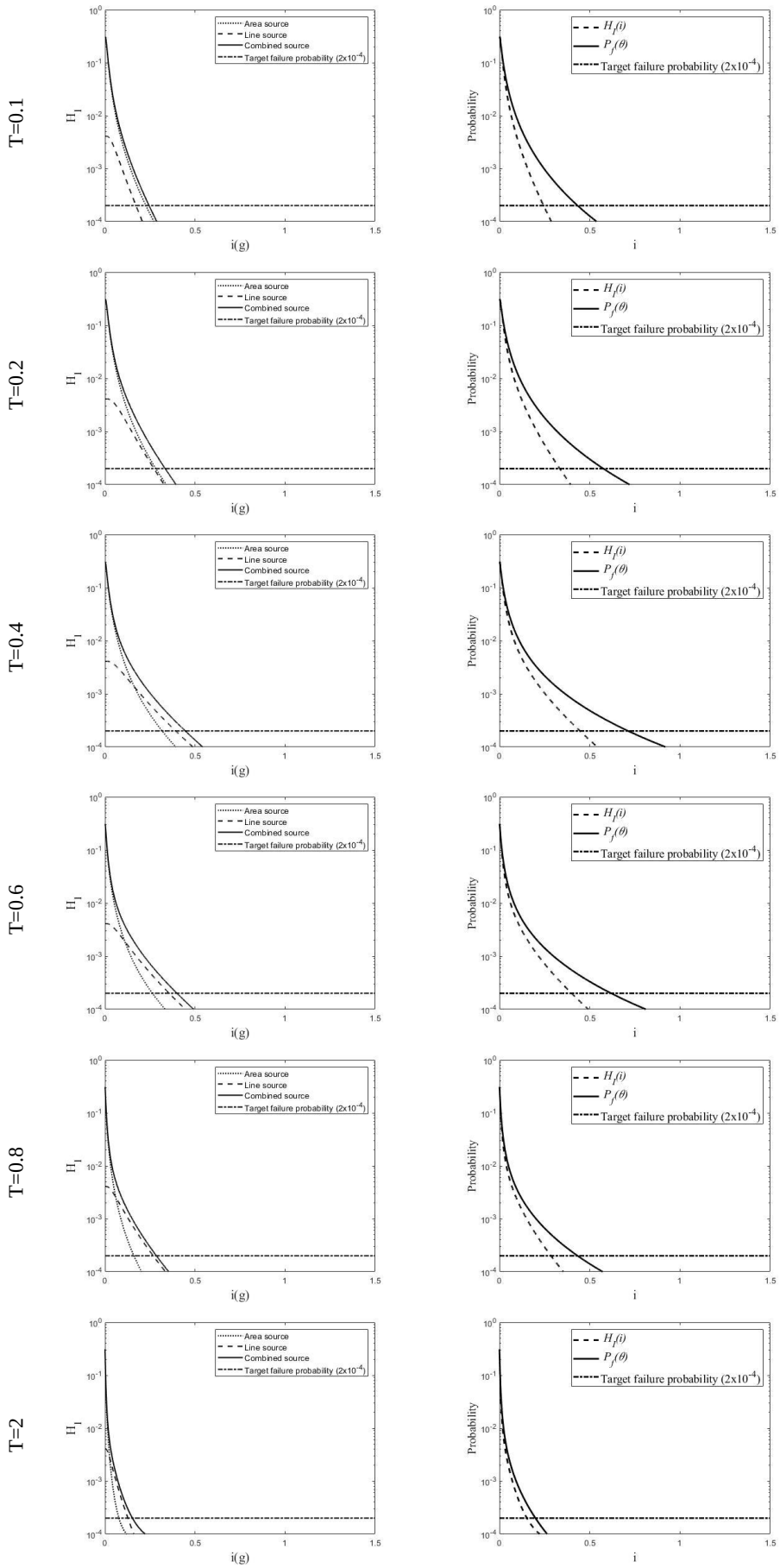



***Figure 4-9.*** *Rock at* $1.0r_{L0}$ *: left column:* $H_I(i)$ *curves for area, line and combined source; right column:* $H_I(i)$ *and* $p_f(\bar{c})$ *curves*

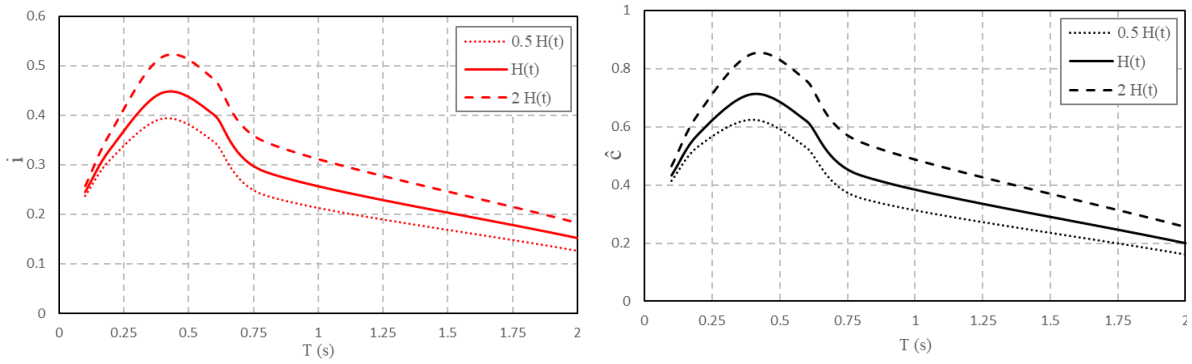



***Figure 4-10.*** *analysis result of combined source for different hazard rates for a site located on rock; left: demand curves and right: capacity curves*

In order to compare the changes of demand and capacity values, they are depicted in Fig. 4.11. As also provided quantitatively in Table 4, the changes of capacity values are higher than those of demand values, which means that the changes of demand do not translate to the changes of capacity exactly.

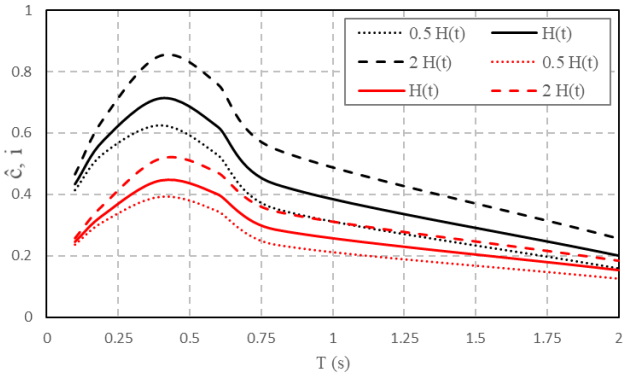



***Figure 4-11.*** *analysis result of combined source for different hazard rates for a site located on rock*

***Table 4-4.*** *demand and capacity values for different T(s) for a site located on rock*

| | ***0.5 H(t)*** | | | | ***H(t)*** | | ***2 H(t)*** | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| ***T (s)*** | ***Demand (g)*** | ***Δi (%)*** | ***Capacity (g)*** | ***Δĉ (%)*** | ***Demand (g)*** | ***Capacity (g)*** | ***Demand (g)*** | ***Δi (%)*** | ***Capacity (g)*** | ***Δĉ (%)*** |
| *0.1* | *0.237* | *-3.27* | *0.414* | *-4.39* | *0.245* | *0.433* | *0.257* | *4.90* | *0.466* | *7.62* |
| *0.2* | *0.312* | *-6.59* | *0.533* | *-7.47* | *0.334* | *0.576* | *0.367* | *9.88* | *0.647* | *12.33* |
| *0.4* | *0.393* | *-11.88* | *0.624* | *-12.48* | *0.446* | *0.713* | *0.518* | *16.14* | *0.852* | *19.50* |
| *0.6* | *0.345* | *-13.75* | *0.527* | *-14.86* | *0.4* | *0.619* | *0.471* | *17.75* | *0.758* | *22.46* |
| *0.8* | *0.237* | *-16.84* | *0.354* | *-18.24* | *0.285* | *0.433* | *0.345* | *21.05* | *0.548* | *26.56* |
| *2* | *0.126* | *-17.11* | *0.16* | *-20.00* | *0.152* | *0.2* | *0.183* | *20.39* | *0.257* | *28.50* |

Sections 4.1 – 4.3 compared demand and capacity values at different times and periods. The effect of soil conditions on the results is compared more precisely in Fig. 4.12 and Table 4.5.

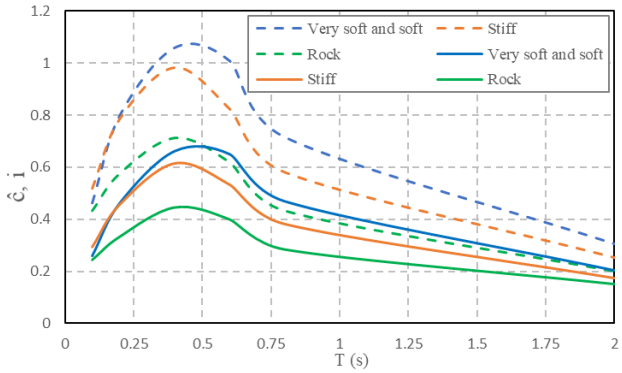



***Figure 4-12.*** *analysis result of combined source for different soil conditions and different periods at* $1.0r_{L0}$ *(solid curves: demand, dashed curves: capacity)*

The first significant point which could be understood from Fig. 4.12 is that the capacity values are significantly higher than demand values in all the three soil types. Moreover, the results obtained for a site located in very soft or soft soil is higher than those obtained for stiff or rock. Otherwise noted, the looser the soil is, the higher the demand and capacity will be. Additionally, demand (or capacity) values could be considered the same for a site located on very soft, soft or stiff soil up to period 0.25s. The above-mentioned comparisons could be quantitatively observed in the data provided in Table 4.5.

***Table 4-5.*** *demand and capacity values for different soil types and T(s) at* $1.0r_{L0}$

| | ***Very soft and soft*** | | ***Stiff*** | | ***Rock*** | |
|---|---|---|---|---|---|---|
| ***T (s)*** | ***Demand (g)*** | ***Capacity (g)*** | ***Demand (g)*** | ***Capacity (g)*** | ***Demand (g)*** | ***Capacity (g)*** |
| *0.1* | *0.26* | *0.461* | *0.293* | *0.518* | *0.245* | *0.433* |
| *0.2* | *0.463* | *0.798* | *0.456* | *0.785* | *0.334* | *0.576* |
| *0.4* | *0.663* | *1.059* | *0.614* | *0.982* | *0.446* | *0.713* |
| *0.6* | *0.651* | *1.008* | *0.532* | *0.823* | *0.4* | *0.619* |
| *0.8* | *0.47* | *0.715* | *0.381* | *0.58* | *0.285* | *0.433* |
| *2* | *0.204* | *0.304* | *0.173* | *0.252* | *0.152* | *0.2* |

## 4.4 Different GMPE

Two GMPE models are considered in this research for analyzing demand and capacity values: (a) GMPE proposed by Ambrasyes et al., which was explained in section 4.2 (Ambraseys, Simpson et al. 1996), and (b) GMPE proposed by Sabetta and Pugliese which is given in Eq. (4.3) (Sabetta and Pugliese 1996):

$$\log_{10}(I) = a + bm - c\log_{10}\sqrt{x^2 + h^2} + e_1 S_1 + e_2 S_2 + \varepsilon \qquad (4.3)$$

where $\varepsilon$ is a 0-mean Gaussian random variable, $S_1$ and $S_2$ are parameters depending on characteristic of soil, $h$ is the fictitious depth, and $a, b, c, e_1, e_2$ are the model constants. In this study, according to the work of (Polidoro, Iervolino et al. 2013) regarding the soil type, the following values are assumed for the parameters: $a = -1.562$, $a = -1.562$, $b = 0.306$, $c = 1$, $S_1 = 1$, $e_1 = 0.169$, $S_2 = 0$ and $h = 5.8$ (Sabetta and Pugliese, 1996). The standard deviation of $\varepsilon$ is assumed equal to 0.173.

hazard, $H_I(i)$, due to area, line and combined sources obtained using above-mentioned GMPEs are depicted in the left column of Fig. 4.13. Corresponding capacity curves are also illustrated in the right column.

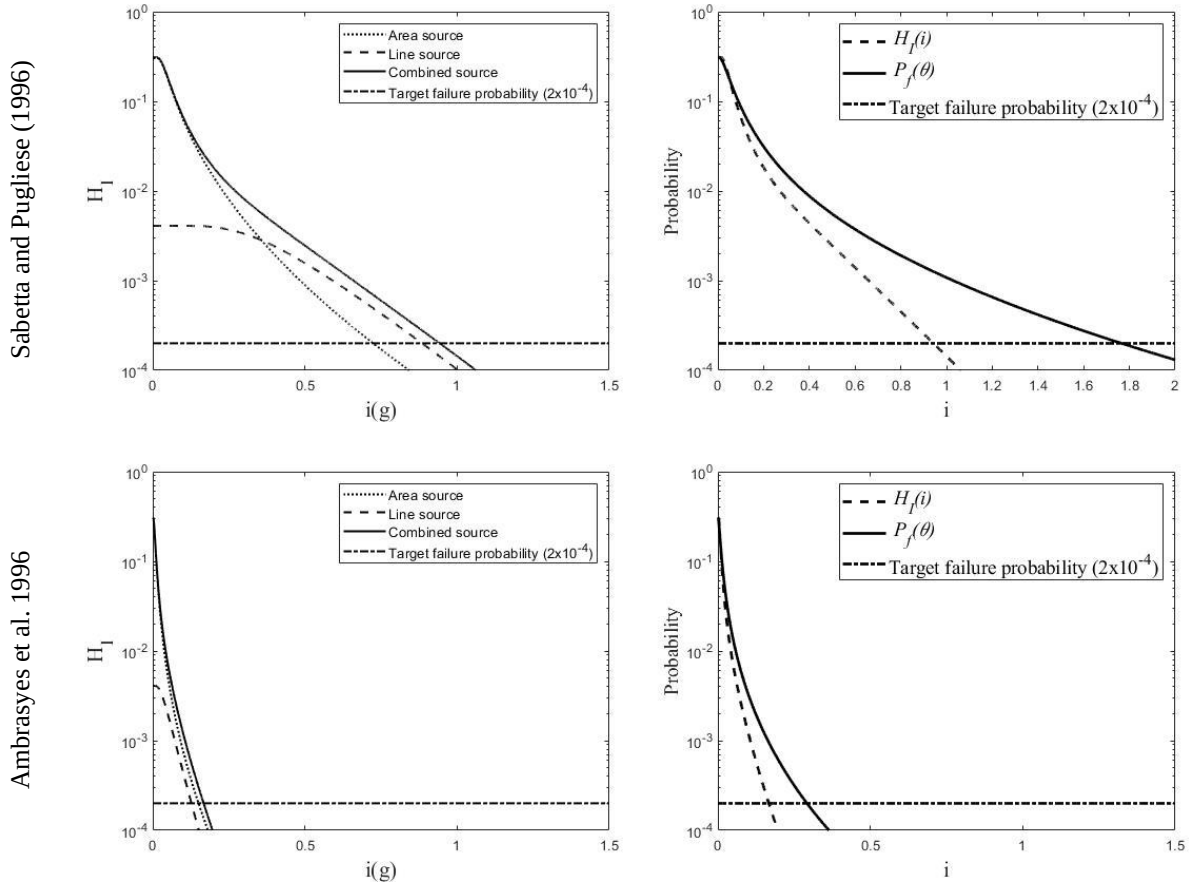



***Figure 4-13.*** *left column:* $H_I(i)$ *curves for area, line and combined source; right column:* $H_I(i)$ *and* $p_f(\bar{c})$ *curves at* $1.0r_{L0}$

Taking hazard curves shown in Fig. 4.13 into consideration, it could be observed that regardless of the GMPE used, the hazard due to the combined source is affected mostly by area and line sources for high and low probabilities, respectively. However, the results obtained by Sabetta and Pugliese GMPE are considerably higher than those of Ambrasyes et al. GMPE. The

difference between capacity and demand in the results obtained using Sabetta and Pugliese is consequently higher than those of Ambraseys et al. Table 6 also provides the demand and capacity for the target failure probability (0.0002).

In order to evaluate demand and capacity values in the case of time-dependent models, the obtained results are illustrated in Fig. 4.14. Red curves show hazard due to combined source while black curves depict capacity curves.

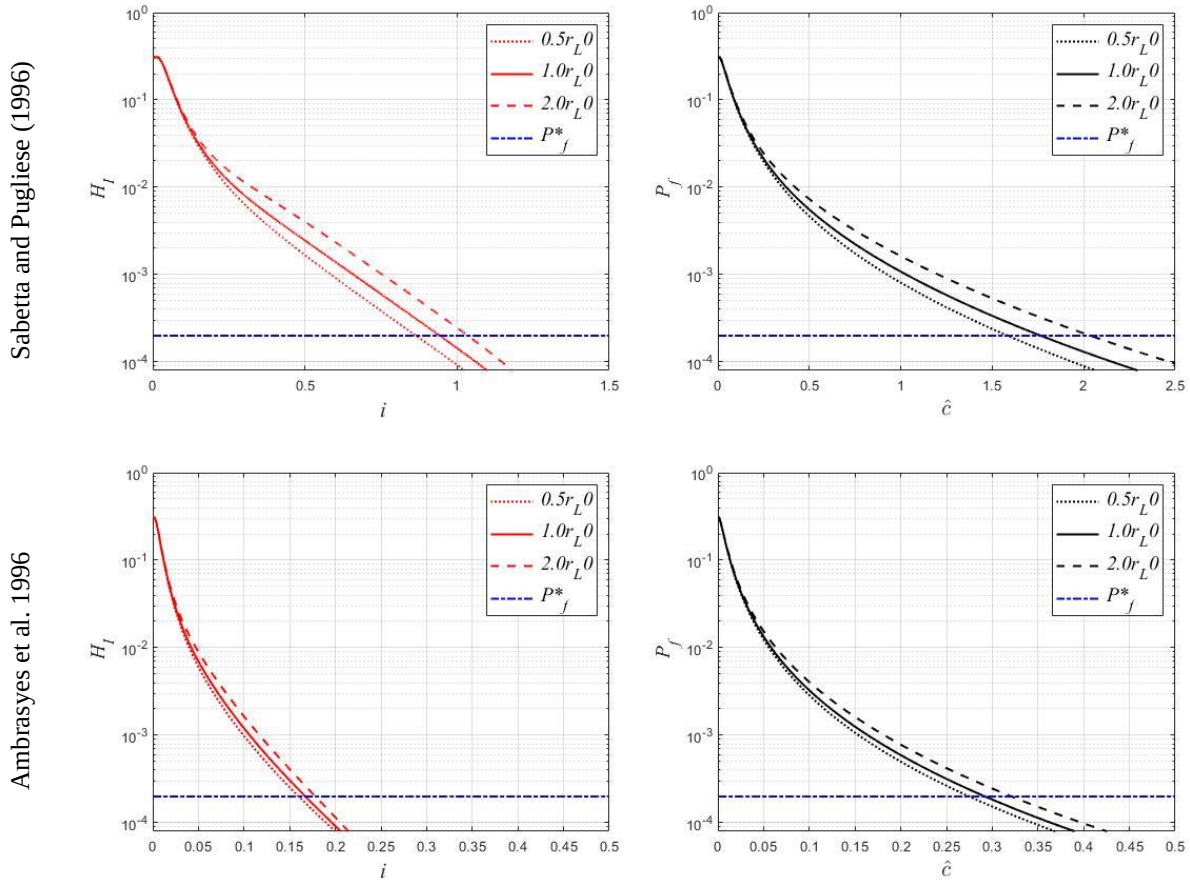



***Figure 4-14.*** *analysis result of combined source for different hazard rates using different GMPEs; left: demand curves and right: capacity curves*

As expected, independent of the GMPEs, demand and capacity increase by elapsing time. In the results obtained by the mentioned GMPEs, the capacity changes are slightly higher than those of demand (as also reported quantitatively in Table 4.6). It is noteworthy that the effect of elapsing time on both demand and capacity in the results obtained by Sabetta and Pugliese GMPE is more considerable in comparison to those obtained by Ambraseys et al. GMPE. These changes are provided in Table 4.6 quantitatively.

***Table 4-6.*** *demand and capacity values for different soil types and T(s) at* $1.0r_{L0}$

| | | **0.5 $r_0$ (139yr)** | **1.0 $r_0$ (185yr)** | **2 $r_0$ (371yr)** | **Δ (%)** | |
|---|---|---|---|---|---|---|
| | | | | | **0.5 $r_0$ (139yr)** | **2 $r_0$ (371yr)** |
| ***Ambraseys et al. 1996*** | *i (g)* | *0.16* | *0.167* | *0.178* | *-4.19* | *6.58* |
| | *ĉ (g)* | *0.275* | *0.292* | *0.321* | *-5.82* | *9.93* |
| ***Sabato & Pugliese et al. 1996*** | *i (g)* | *0.865* | *0.943* | *1.035* | *-8.27* | *9.75* |
| | *ĉ (g)* | *1.584* | *1.765* | *2.036* | *-10.25* | *15.35* |

## 4.5 Summary and conclusions

Seismic hazard and design capacity of structures in the case of time-dependent hazard model was investigated in this report. a combined source of an area source producing medium-low intensity earthquakes, whose recurrence was described by the no-memory POI model, and a line source producing high intensity earthquakes with a periodic recurrence described by the time-dependent BPT model were considered. The influence of GMPEs on the results was assessed by the application of two known GPMEs available in the literature: Sabetta and Pugliese 1996 and Amberaseys et al .1996. The effect of soil conditions was also evaluated by comparing the results obtained for three different soil categories: very soft and soft, stiff and rock. The main conclusions are:

- The hazard due to combined source is affected mostly by area and line source in short and long periods, respectively in all the considered soil types.
- The difference between demand and capacity for a site located on soil is larger than a site located on rock.
- Regardless of the soil type and GMPE, by elapsing time, both demand and capacity increase. However, their increase rate is not the same and the capacity changes are slightly larger than demand changes.
- For very short periods (less than 0.2s) or very long periods, elapsing time does not affect neither capacity nor demand.
- Elapsing time affects both demand and capacity corresponding to moderate periods (0.4s in this study) significantly.
- The GMPE used for the analysis affects the results remarkably and the most appropriate GMPE which is proposed for the region of the study should be used.

# References

Abrahamson, N. A. and W. J. Silva (1997). "Empirical response spectral attenuation relations for shallow crustal earthquakes." Seismological Research Letters **68**(1): 94-127.

Ambraseys, N. N., et al. (1996). "Prediction of horizontal response spectra in Europe." Earthquake Engineering & Structural Dynamics **25**(4): 371-400.

Anvarsamarin, A., et al. (2018). "Soil-structure interaction effect on fragility curve of 3D models of concrete moment-resisting buildings." Shock and Vibration **2018**.

ASCE 7-10, A. S. A. S. E. I. (2013). Minimum design loads and associated criteria for buildings and other structures

Bazzurro, P. and C. A. Cornell (2004). "Nonlinear soil-site effects in probabilistic seismic-hazard analysis." Bulletin of the Seismological Society of America **94**(6): 2110-2123.

Boore, D. M., et al. (1997). "Equations for estimating horizontal response spectra and peak acceleration from western North American earthquakes: A summary of recent work." Seismological Research Letters **68**(1): 128-153.

Carbonari, S., et al. (2016). "Analytical evaluation of impedances and kinematic response of inclined piles." Engineering Structures **117**: 384-396.

Eurocode 8, E. S. E.-. (2004). Eurocode 8. Design of structures for earthquake resistance - Part 1: General rules, seismic actions and rules for buildings.

Ezzelarab, M., et al. (2021). "Integration of local soil effect into the assessment of seismic hazard at the Kharga Oasis, Western Desert, Egypt." Scientific African **12**: e00747.

FEMA (2009a). NEHRP recommended seismic provisions for new buildings and other structures (FEMA P750). Federal Emergency Management Agency.

Givens, M. J., et al. (2016). "Modular analytical solutions for foundation damping in soil-structure interaction applications." Earthquake Spectra **32**(3): 1749-1768.

Kramer, S. (1996). "Geotechnical earthquake engineering, Prentice Hall Upper Saddle River." New Jersey: 07458.

Lam, N. T., et al. (2015). "Local intraplate earthquake considerations for Singapore." The IES Journal Part A: Civil & Structural Engineering **8**(1): 62-70.

Montoya-Noguera, S. and F. Lopez-Caballero (2018). "Effect of coupling excess pore pressure and soil deformation on nonlinear SSI in liquefiable soil deposits." Bulletin of Earthquake Engineering **16**(2): 681-705.

Nabilah, A. B., et al. (2019). "Effect of Flexible Soil in Seismic Hazard Assessment for Structural Design in Kuala Lumpur." International Journal of Geotechnical Earthquake Engineering (IJGEE) **10**(1): 30-42.

Polidoro, B., et al. (2013). Models and issues in history-dependent mainshock hazard. Proc. of the 11th Conference on Structural Safety and Reliability, ICOSSAR.

Rajeev, P. and S. Tesfamariam (2012). "Seismic fragilities of non-ductile reinforced concrete frames with consideration of soil structure interaction." Soil Dynamics and Earthquake Engineering **40**: 78-86.

Sabetta, F. and A. Pugliese (1996). "Estimation of response spectra and simulation of nonstationary earthquake ground motions." Bulletin of the Seismological Society of America **86**(2): 337-352.

Sadigh, K., et al. (1997). "Attenuation relationships for shallow crustal earthquakes based on California strong motion data." Seismological Research Letters **68**(1): 180-189.

Schnabel, P., et al. (1972). "Modification of seismograph records for effects of local soil conditions." Bulletin of the Seismological Society of America **62**(6): 1649-1664.

Schwartz, D. P. and K. J. Coppersmith (1984). "Fault behavior and characteristic earthquakes: Examples from the Wasatch and San Andreas fault zones." Journal of Geophysical Research: Solid Earth **89**(B7): 5681-5698.

Shahbazi, S., et al. (2018). "Effect of soil classification on seismic behavior of SMFs considering soil-structure interaction and near-field earthquakes." Shock and Vibration **2018**.

Tondi, E. and G. Cello (2003). "Spatiotemporal evolution of the Central Apennines fault system (Italy)." Journal of Geodynamics **36**(1-2): 113-128.

Wesnousky, S. G. (1994). "The Gutenberg-Richter or characteristic earthquake distribution, which is it?" Bulletin of the Seismological Society of America **84**(6): 1940-1959.

Zhao, X., et al. (2017). "Optimal design of viscoelastic dampers in frame structures considering soil-structure interaction effect." Shock and Vibration **2017**.

# 5 Fragility curves of buildings; a critical review and a machine learning-based study

## Abstract

Fragility curves are one of the substantial means required for seismic risk assessment of buildings in the framework of performance-based earthquake engineering (PBEE). Deriving fragility curves, however, needs an extensive numerical analysis which is time-consuming and sometimes inaccurate due to errors. In this study, hence, machine learning (ML)-based models are proposed for predicting fragility parameters of structures namely dispersion, β, and median, μ, based on a reduced set of information avoiding cumbersome computation. Firstly, a critical review on the analytical models proposed for deriving fragility curves is provided as well as ML-based models developed so far. Then, to achieve the research objective, a comprehensive database including 214 datasets from peer-reviewed international publications is collected. It is then divided into training (85%) and testing (15%) sub-datasets for the purpose of training the models and assessing the results, respectively. The most effective parameters on the target outputs are defined as input variables including construction material, building plan area, building height, damage state, buildings' period and soil classification. β and μ were estimated utilizing various ML-based techniques namely nonlinear regression, decision tree, random forest, K-Nearest Neighbours (KNN) and Artificial Neural Network (ANN). The actual values and the values predicted by the proposed models are compared. Moreover, the models' accuracy is assessed through performance metrics and Taylor diagram. The results demonstrated the high ability of the models for learning the relationship between inputs and outputs. According to the accuracy assessment, Decision tree was the most accurate model for predicting both β and μ. A sensitivity study was also conducted by changing input variables and estimation equations are provided accordingly.

## 5.1 Introduction

### *5.1.1 background*

Among various natural hazards, earthquakes are recognised as the greatest threat to human beings (Özel and Güneyisi 2011, Frankie, Gencturk et al. 2013). As a result, risk assessment is undoubtedly of high importance which could help humankinds in the sense of life and economics (Prieto, Journeay et al. 2018). In this context, Performance-based Earthquake Engineering (PBEE) and potential failure mode analysis (PFMA) have been developed over the past years (Kiani, Camp et al. 2019, Saouma and Hariri-Ardebili 2021). PBEE is known as a process with quantitative measures as the output which reflect the response of structures under seismic loads (Lange, Devaney et al. 2014). One of the main issues in seismic risk analysis is various sources of uncertainties (e.g., randomness in ground motion, structural modelling related uncertainties, damage state, etc.) which is addressed by considering fragility curves in the PBEE framework (Kiani, Camp et al. 2019, Hwang, Mangalathu et al. 2021).

### *5.1.2 Fragility curves definition*

Various definitions have been provided for clarifying the concept of fragility curves by researchers (Park, Towashiraporn et al. 2009, Rota, Penna et al. 2010, Cardone, Rossino et al. 2018, Del Gaudio, De Martino et al. 2019, Alwaeli, Mwafy et al. 2020, Dall'Asta, Dabiri et al. 2021). Simply stated, fragility is described as the probability of reaching or exceeding a specified damage measure (DM) under a given earthquake Intensity Measure (IM) for a structure (Park, Towashiraporn et al. 2009, Prieto, Journeay et al. 2018, Sandoli, Lignola et al. 2021). In the viewpoint of mathematic, fragility curves could be expressed as Eq. 5.1 (Mitropoulou and Papadrakakis 2011).

$$F_R(x) = P(\theta_{max} \geq y | IM = x) \qquad (5.1)$$

where $\theta_{max}$ is the maximum Engineering Damage Parameter (EDP), and $y$ and $x$ are the values of $\theta_{max}$ (the threshold defined the DM limit state) and $IM$, respectively. It is worth clarifying that that $\theta_{max}$ and IM are real-value positive random variables while y and x are particular values that can be assumed by samples of the random variable. The fragility curves could be drawn using lognormal distribution functions which are defined by two parameters: median (μ) and lognormal standard deviation (or dispersion, β). These parameters, therefore, are known as fragility curve's parameters and could be estimated by different techniques e.g., the maximum likelihood method (Shinozuka, Feng et al. 2000, Mitropoulou and Papadrakakis 2011, Kiani, Camp et al. 2019). As a result, the fragility curves could be expressed as Eq. 5.2:

$$F_R(IM) = \Phi\left[\frac{\ln(x) - \mu_{\ln(x)}}{\beta_{\ln(x)}}\right] \qquad (5.2)$$

where $\Phi(.)$ is the standardized normal distribution function, x is the PGA, $\mu_{ln(x)}$ is the logarithmic median (is represented μ hereafter) and $\beta_{ln(x)}$ is the logarithmic standard deviation (is represented β hereafter) defining the lognormal distribution (Mitropoulou and Papadrakakis 2011, Del Gaudio, De Martino et al. 2019).

As could be figured out from the above-mentioned explanations, the chosen IM could affect fragility curve of a building significantly. A macroseismic IM does not lead to accurate fragility curves because (i) it can cause interdependency between the vulnerability and the IM itself because it is obtained from observation of earthquake consequences on buildings, and (ii) it is a subjective parameter which means that it is influenced by the sensitivity and judgment of the surveyor (Del Gaudio, De Martino et al. 2019, Sandoli, Lignola et al. 2021). Hence, parameters namely peak ground acceleration (PGA), peak ground velocity (PGV) and peak ground displacement (PGD) have been applied instead of macro seismic IMs (Del Gaudio, De Martino et al. 2019). Among these parameters, PGA is more common than others because of its simplicity for deriving fragility curves of complex buildings (Pejovic and Jankovic 2016). It should be also stated that different parameters could be used as EDP such as interstory drift ratio (IDR), ultimate rotation, etc. The most usual EDP, however, is IDR since it could reflect the structures' damage state (Sandoli, Lignola et al. 2021) using different thresholds.

### *5.1.3 Review on the methods for deriving fragility curves*

IDA is one of the most widespread methods for obtaining fragility curves which reflects the relationship between the defined IM and EDP. IDA could be performed through: (a) defining a metric for both the seismic action (IM) and the construction response (EDP), (b) performing a nonlinear model to relate the two metrics, (c) collecting a sufficient number of samples of the seismic actions and the model parameters to include uncertainties (Park, Towashiraporn et al. 2009, Cardone, Rossino et al. 2018, Aljawhari, Gentile et al. 2021).

Distribution of the obtained curves through the above-mentioned steps, intersected by a vertical line passing from each damage state will be approximated by a lognormal distribution with median ($\mu$) and dispersion ($\beta$) which are used for deriving the fragility curve of the corresponding damage state. Fig. 5.1 displays the fragility curve derivation schematically.

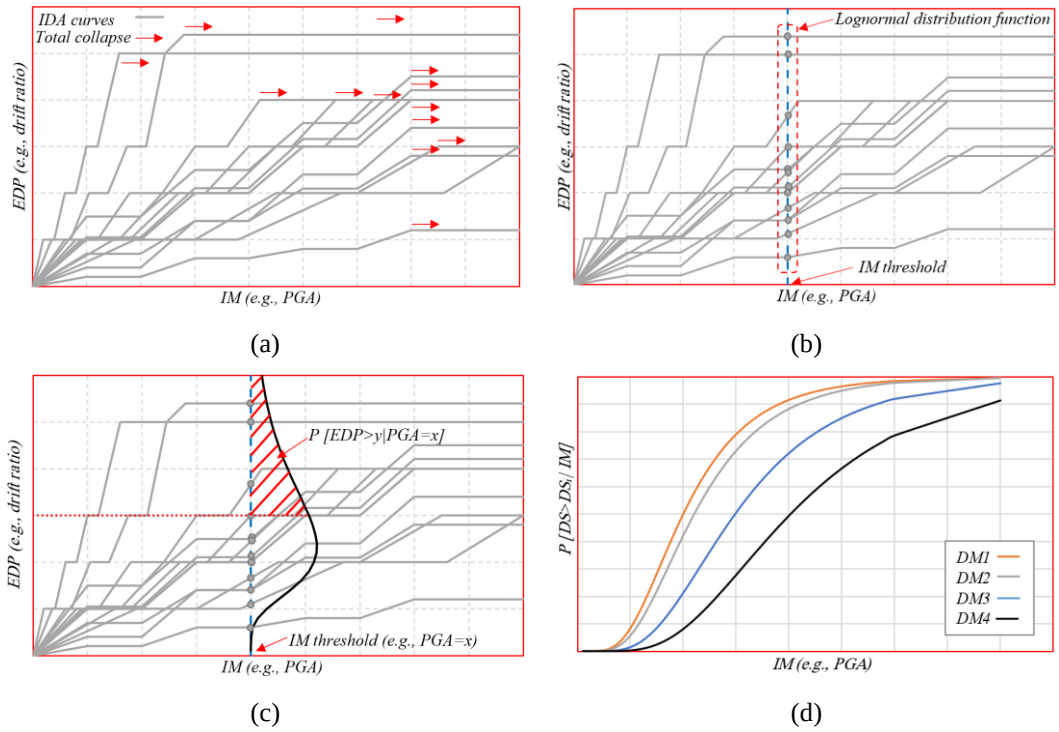



***Figure 5-1**. Schematically illustration of driving fragility curves; (a) incremental dynamic analysis, (b) distribution at a specified damage state, (c) probability of failure for the damage state, and (d) fragility curves.*

The method shown in Fig. 5.1 is recognized as the conventional method for generating fragility curves. Several novel approaches have been developed and presented in the last few decades as well. The fundamental concepts of these methods are summarized in Table 5.1.

***Table 5-1**. A summary of the methods developed for deriving fragility curves.*

| ***reference*** | ***aim*** | ***Methodology and outcomes*** |
|---|---|---|
| *Park et al. (Park, Towashiraporn et al. 2009)* | *Evaluating seismic fragility of low-rise unreinforced masonry buildings* | *A simplified spring model was presented for simulating the highly nonlinear dynamic behaviour of URM buildings.* |
| *Rota et al. (Rota, Penna et al. 2010)* | *Developing a methodology for driving fragility curves of masonry buildings* | *Mechanical properties were considered as random variables with a reasonable value. The input variables were extracted through Monte Carlo method from the distributions. The probability distribution of the damage states was defined by pushover analysis. Novelty: unlike other studies, results of nonlinear stochastic analyses of a prototype building is used while simplified models of buildings and approximate analysis are generally considered by researchers.* |
| *Cardone et al. (Cardone, Rossino et al. 2018)* | *Driving Fragility curves of RC buildings through a hybrid approach* | *In HAZUS, fragility curves are developed using inelastic static analysis, in this study, however, fragility parameters are obtained by a hybrid method. Median values are predicted by comprehensive loss assessment analysis and dispersion values are evaluated according to the results of accurate inelastic dynamic analysis.* |
| *Alwaeli et al. (Alwaeli, Mwafy et al. 2020)* | *Proposing a new methodology for driving fragility curves with less computational efforts and time* | *A novel record selection criterion and a fragility curve tolerance factor are provided for developing rigorous (refined) and less-demanding fragility relations for RC high-rise buildings.* |
| *Dona et al. (Donà, Carpanese et al. 2021)* | *Developing a mechanics-based fragility model for Italian residential URM buildings* | *A database including 500 building information was used to proposed a fragility model which was based on the classification of the buildings in terms of age and story numbers. The verification was made by simulating 2009 L'Aquila earthquake, proved the acceptable reliability of the model.* |
| *Sandoli et al. (Sandoli, Lignola et al. 2021)* | *Proposing a hybrid method for driving fragility curves of masonry buildings* | *The innovative method is a combination of: (i) an expert judgment by classifying buildings in the viewpoint of construction age, structural typology, seismic behaviour and damage of buildings caused by the earthquakes, and (ii) numerical analysis results.* |
| *Beilic et al. (Beilic, Casotto et al. 2017)* | *Developing out-of-plane fragility functions using probabilistic analysis and Monte Carlo simulation* | *The study novelty was that, unlike other studies which consider demand, this study used the uncertainties in the capacity namely aspect ratio, level of In-Plane damage, position of the infill walls and type of masonry (solid or hollow units).* |
| *Cardone et al. (Cardone, Perrone et al. 2019)* | *Generating collapse fragility curves for base-isolated RC buildings* | *Fragility curves of RC building with either low- or high-seismic resistance systems retrofitted by different isolation systems are obtained.* |

The main aim of the recently developed models is to ease the process of deriving fragility curves by (i) proposing building simulation assumptions (Park, Towashiraporn et al. 2009, Rota, Penna et al. 2010), (ii) proposing simpler methods by combining simpler techniques (Cardone, Rossino et al. 2018, Sandoli, Lignola et al. 2021), (iii) using the assets of building classification in terms of age, structural typology, story numbers and seismic behaviour (Donà, Carpanese et al. 2021, Sandoli, Lignola et al. 2021), (iv) presenting novel earthquake record selection (Alwaeli, Mwafy et al. 2020) and (v) considering the uncertainties of capacity rather than those of demand (Beilic, Casotto et al. 2017). Fragility curves of buildings with novel resisting systems (e.g., seismic isolators) have also been presented and discussed by researchers (Cardone, Perrone et al. 2019).

### *5.1.4 Review on the application of ML-based techniques for generating fragility curves*

Due to complicated and time-consuming analysis required for IDA and fragility assessment, there has been an increasing interest in the implementation of quicker and less-complicated models for deriving fragility curves. Artificial Intelligence (AI) and Machine Learning (ML) have been increasingly applied by researchers in various fields (Faramarzi, Javadi et al. 2013, Alani and Faramarzi 2014, Faramarzi, Alani et al. 2014) as well as structural and earthquake engineering (Salehi and Burgueño 2018, Sun, Burton et al. 2021, Dabiri, Kioumarsi et al. 2022, Dabiri, Rahimzadeh et al. 2022).

During the last years, a few attempts have been made to apply ML-based techniques in the process of emerging fragility curves. A summary of the relevant studies is given in Table 5.2. The main objective of the studies listed in Table 5.2 is to boost the process of generating fragility curves by either making the process quicker (Kiani, Camp et al. 2019) or reducing the uncertainty degrees of different parameters (Mitropoulou and Papadrakakis 2011, Jia and Wu 2021). As an example, maximum story drift was estimated through ML-based methods in the models proposed by Kiani et al (Kiani, Camp et al. 2019) and Hwang et al. (Hwang, Mangalathu et al. 2021). Jia and Wu (Jia and Wu 2021) have also developed a novel model for predicting a dimensionless parameter which was used for obtaining failure probability. The most remarkable conclusion of the studies is that the results of the prediction models developed based on ML or Neural Network (NN) are in line with the results of the conventional method with an acceptable level of accuracy. More specifically, their evaluations have clarified that tree-based approaches (e.g., Decision Tree, DT or Random Forest, RF) led to more accurate outcomes compared to other ML-based approaches (e.g., ANN) (Hwang, Mangalathu et al. 2021, Jia and Wu 2021).

The notable issue with the proposed models, on the other hand, is their limitations including (i) being capable only for a specific building. In other words, the prediction models are developed based on the results obtained for the buildings with specified number of stories or construction material, as also reported in the last column of Table 5-2. As an example, the prediction model developed by Kiani et al (2019) could be used for a 8-stody steel frame. Furthermore, (ii) Some effective parameters (e.g., construction material, soil type, building's location, etc.) are not considered in the prediction models. Accordingly, further studies are definitely required to provide a more generalized model for emerging fragility curves.

***Table 5-2**. Literature review on the application of ML-based models for obtaining fragility curves.*

| *reference* | *aim* | *methodology* | *conclusions* | ***Predicted parameters (output)*** | ***Limitations*** |
|---|---|---|---|---|---|
| *Mitropoulou and Papadrakakis (Mitropoulou and Papadrakakis 2011)* | *Developing fragility curves based on neural network IDA predictions using Harmony search Optimization algorithm.* | *(a) Analysing an 8-story regular plan and a 5-story irregular plan RC buildings in OpenSees, (b) considering Arias Intensity ($I_A$), Characteristic intensity ($I_C$ ) and Cumulative Absolute Velocity (CAV) IMs reflecting respectively the amplitude, the frequency content and the duration of a strong ground motion were considered as inputs while spectral acceleration in different DSs defined as outputs (4 nodes), (c) considering SA and maximum inter-story drift ratio (MIDR) as IM and EDP, respectively, (d) predicting fundamental period spectral acceleration (seismic demand) by NN, (e) developing four limit state fragility curves, (f) examining the computational cost of the neurocomputing scheme.* | *The obtained results with NN were in line with the results of the conventional method.* | *Period spectral acceleration was predicted.* | *(i) The results were limited to two buildings, (ii) Arias Intensity, Characteristic intensity and Cumulative Absolute Velocity IMs were considered as inputs while other parameters were not taken into account.* |
| *Kiani et al. (Kiani, Camp et al. 2019)* | *(a) Application of Machine learning methods for deriving fragility curves, (b) evaluating influence of input uncertainties (i, e., GM variability), (C) investigating the influence of training samples' size on the results of classification techniques.* | *(a) Modelling an 8-story steel special moment resisting frame with the period of 2.3s located in Los Angeles, US in OpenSees, (b) considering spectral velocity as IM, (c) considering as EDP in two groups (MIDR>0.03 rad and MIDR<0.03 rad), (e) predicting structure response in terms of Displacement Spectrum Intensity Ratio (DSIR)* | *(i) RF, SVM and DT were the most accurate methods, (ii) lasso regression and Naïve Bayes were not affected by training sample size while QDA was the most sensitive method.* | *DSIR is predicted.* | *(i) The results were limited to an 8-story steel frame, (ii) only one limit state (MIDR=0.03) was considered), (iii) other input parameters including structural systems, damage measures, thresholds of structural responses, and sites were not considered.* |
| *Jia and Wu (Jia and Wu 2021)* | *Driving fragility curves of RC frame-shear wall structures using ensemble Neural network.* | *(a) Analysing a 4-story shear wall RC structure with a basic acceleration of 0.2 g located on Type II soil in China using SAP2000, (b) Considering MDIR and peak floor acceleration (PFA) as EDP and PGA as IM, (c) Defining four common damage states, (d) Defining 26 parameters reflecting GM and structure properties as inputs and an index (L) as output.* | *The proposed ensemble ANN model predicted more accurate values in comparison to back propagation (BP), cascade BP and Alman ANN.* | *a dimensionless parameter (L) was predicted which is used for calculating failure probability.* | *(i) The results were limited to a 4-story shear wall RC structure, (ii) parameters namely material, period, soil type, and location were not considered in the model.* |
| *Hwang et al. (Hwang, Mangalathu et al. 2021)* | *(a) Predicting seismic response and structural collapse of RC frames, (b) identifying different input variables on RC frames' seismic collapse.* | *(a) Analysing a 4-story and an 8-story special moment resisting frame located on stiff soil in California using OpenSees, (b) Considering Sa and maximum story drift as IM and EDP, respectively, (c) using modelling-related parameters and ground motion intensity measures as input for predicting MSD in the first prediction model, (d) using MSD in addition to all the input models in the first part, as input for predicting survival-failure vector.* | *(i) Influence of structural modelling uncertainties on seismic demand could be safely neglected at minor to moderate damage levels for low- to mid-rise RC frame buildings, (ii) extreme gradient boosting algorithm and tree-based techniques (dt and RF) led to more accurate results compared to other methods.* | *Maximum story drift and collapse status were predicted.* | *(i) The results were limited to 4- and 8- story RC frame buildings, (ii) parameters namely material, period, soil type, and location were not considered in the model.* |

## 5.2 Significance, novelty and methodology of the study

Obtaining fragility curves is an inevitable key step in seismic risk assessment in the performance-based earthquake engineering which accounts for the uncertainties of risk due to seismic events. Deriving fragility curves, however, is generally time-consuming due to a huge amount of analytical analysis performed for IDA. Moreover, it needs powerful operating systems, particularly in the case of studying tall and complex buildings.

Although a few innovative models have been developed so far (Table 5.2) (Mitropoulou and Papadrakakis 2011, Kiani, Camp et al. 2019, Hwang, Mangalathu et al. 2021, Jia and Wu 2021), they suffer from shortcomings which were mentioned in the previous section. More importantly, the models do not estimate fragility curves directly. Otherwise noted, they could be used for predicting neither IM or EDP. This means that the time-consuming analysis still needs to be performed even if the prediction models are applied.

The main objective of this research, therefore, is to propose a prediction model for deriving fragility curves using regression- and ML-based techniques. The remarkable novelty of this study is proposing models which output fragility curves' parameters ($\beta$ and $\mu$) directly. In other words, the time-consuming IDA is eliminated and consequently the fragility curves could be obtained quickly by defining the inputs. Furthermore, the main parameters which are proved to have a strong effect on fragility curves are considered. More importantly, the proposed models are not limited to a specific building and could be utilized for a large class of reinforced concrete (RC), steel or masonry buildings.

Overall, the benefits of the proposed prediction models are: (i) predicting $\mu$ and $\beta$ directly, (ii) eliminating time-consuming IDA, (iii) considering the main effective parameters and (iv) being generalized and therefore applicable for a huge number of buildings.

To this aim, a comprehensive database is gathered and various ML-based methods including nonlinear regression (NLR), Decision Tree (DT), Random Forest (RF), K-nearest Neighbours (KNN) and Artificial Neural Network (ANN) are used to develop models for deriving fragility curves. The accuracy of the models is assessed through Taylor diagram and performance metrics and the most accurate model is introduced. Eventually, a parametric study is conducted to investigate the effect of input variables on the output parameters and prediction equations are provided accordingly.

## 5.3 Data collection

A database including 214 results of the fragility assessment of buildings was collected from peer-reviewed international publications on either empirical data or numerical analysis (Kirçil and Polat 2006, Park, Towashiraporn et al. 2009, Rota, Penna et al. 2010, Özel and Güneyisi 2011, Abo-El-Ezz, Nollet et al. 2013, Frankie, Gencturk et al. 2013, Saruddin and Nazri 2015, Pejovic and Jankovic 2016, Prieto, Journeay et al. 2018, Zucconi, Ferlito et al. 2018, Del Gaudio, De Martino et al. 2019, Kumar and Samanta 2020, Aljawhari, Gentile et al. 2021, Karafagka, Fotopoulou et al. 2021, Suzuki and Iervolino 2021). Otherwise noted, in the above-mentioned studies fragility curves of various building were obtained using numerical analysis or experimental investigation and their final outcome are used in this research. It is worth mentioning that the reliability of the collected database was improved by removing outliers (a value which differs from the rest of the data significantly) and incomplete datasets. It should be also explained that during the data collection, special attention was given to consider the parameters which affect fragility parameters significantly based on the literature results. As an example, Gaudio et al. (Del Gaudio, De Martino et al. 2019) have proved that buildings' height has a significant influence on its fragility while influence of its construction age could be neglected.

Accordingly, the parameters considered as inputs for the prediction models are: construction materials (i. e., RC, steel and masonry), buildings' plan area ($m^2$), buildings' height (m), lateral resisting system (i. e., shear wall, bearing masonry wall, bracing system or moment resisting frame-MRF), buildings' location (the country in which the building is located: Turkey, Malaya, Italy, Quebec, the US), damage state, buildings' period, soil classification (e. g., rock). In terms of period, it should be explained that the building's period is obtained either by the numerical analysis or the available equations in design codes. The output, on the other hand, are the fragility parameters: median value of the ground motion index at which the building reaches the defined damage state threshold (in this study PGA expressed by $\mu$) and dispersion of the natural logarithmic of ground motion index of damage level ($\beta$). One of the parameters which could reveal if the input variables are defined appropriately is Pearson correlation coefficient and is defined as the ratio of *x*, *y* covariance, *cov (x, y)*, to the production of their standard deviation ($\sigma_x\sigma_y$) as given in Eq. 5.3 (Nettleton 2014, Berman 2018):

$$\rho_{x,y} = \frac{\operatorname{cov}(X,Y)}{\sigma_x \sigma_y} = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2}\sqrt{\sum(y_i - \bar{y})^2}} \qquad (5.3)$$

Pearson coefficient of two parameters could be calculated for quantitative values only and is a value in the range of (0,1). $\rho \approx 1$ reflects almost perfect linear dependency of the parameters on each other while $\rho \approx 0$ stands for no linear dependency of the parameters. It should be explained that Pearson correlation represent only linear relationship between parameters, which means that $\rho \approx 0$ does not necessarily represent complete indecency of variables and they might have nonlinear dependency (Berman 2016, Profillidis and Botzoris 2019). Pearson correlation coefficients of the quantitative input variables, and $\beta$ and $\mu$ are reported in Table 5.3 and Table 5.4, respectively.

***Table 5-3***. *Pearson correlation coefficients between input variables and $\beta$.*

| | *Plan area ($m^2$)* | *Building height (m)* | *Building period (s)* | $\beta$ |
|---|---|---|---|---|
| *Plan area (m2)* | *1.00* | | | |
| *Building height (m)* | *0.57* | *1.00* | | |
| *Building period (s)* | *0.31* | *0.82* | *1.00* | |
| $\boldsymbol{\beta}$ | ***0.34*** | ***0.17*** | ***0.10*** | ***1.00*** |

***Table 5-4***. *Pearson correlation coefficients between input variables and $\mu$.*

| | *Plan area ($m^2$)* | *Building height (m)* | *Building period (s)* | $\mu$ |
|---|---|---|---|---|
| *Plan area (m2)* | *1.00* | | | |
| *Building height (m)* | *0.57* | *1.00* | | |
| *Building period (s)* | *0.31* | *0.82* | *1.00* | |
| $\boldsymbol{\mu}$ | ***0.53*** | ***0.34*** | ***0.27*** | ***1.00*** |

According to Tables 5.3 and 5.4, it could be claimed that plan area has the highest linear influence on both $\beta$ and $\mu$ compared to other quantitative input variables. It is also noteworthy that, building height has a considerable influence on building period, as given in Table 5-3 and 5-4. The equations proposed in design codes which obtain the building period based on its heigh proves their high relationship. Moreover, the coefficients reported in Table 5-3 and 5-4 clarifies that by increasing either building area or height, $\mu$ increases.

Distribution of the input variable and output parameter(s) influence the applicability and generalization of prediction models; the broader range of input values is, the more general the model will be. On that account, it is tried to include the datasets which lead to a high range of values for each input. Distribution of the input parameters and outputs is demonstrated in Fig. 5.2. The statistical characteristics including minimum, maximum, median, variance and standard deviation of the data are also given in Table 5.5.

***Table 5-5****. Statistical properties of the quantitative input and output parameters.*

| | *Minimum* | *Maximum* | *Mean* | *Median* | *Standard deviation* |
|---|---|---|---|---|---|
| *Plan area ($m^2$)* | *53.129* | *1296.00* | *596.61* | *276.75* | *486.22* |
| *Building height (m)* | *1.35* | *30.00* | *15.38* | *18.00* | *7.56* |
| *Building period (s)* | *.08* | *2.60* | *0.72* | *0.48* | *0.73* |
| *Dispersion (β)* | *0.16* | *0.84* | *0.47* | *0.45* | *0.16* |
| *Median (μ, log(PGA)* | *-3.25* | *1.43* | *-0.24* | *0.07* | *0.91* |

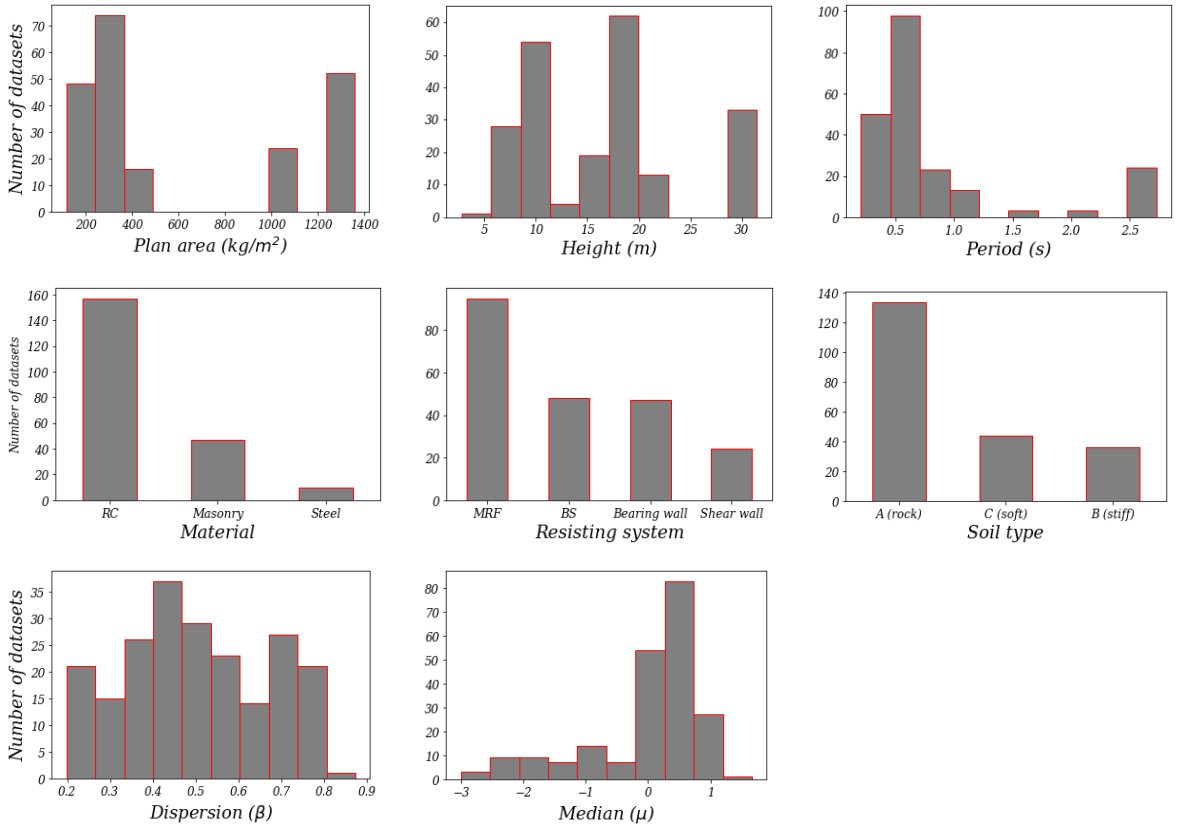



***Figure 5-2****. Distribution of the input and output parameters.*

The database was divided into two sub-databases namely training (85%) and testing (15%). The former is used for training the models the relationship between the inputs and outputs and the latter is used for assessing the accuracy of the predicted values. Otherwise mentioned, the testing datasets are not used for training purpose while they are utilized for validity of the prediction models. It is also noteworthy that no specific portions have been stated for dividing training and testing databases so far and the similar values are typically used for developing ML-based prediction models.

## 5.4 Prediction models

The prediction models are developed in two parts in this study. In the first part models are presented to predict dispersion while median is estimated in the second part. Various techniques including NLR, DT, RF, KNN and ANN were used for proposing prediction models. An extensive definition of the methods could be found in the literature; therefore, they are introduced here briefly for the sake of shortness.

### *5.4.1 Nonlinear regression (NLR)*

Simply stated, a regression-based model is a model which fits an equation to a set of data. The equation refers to a line which processes the least difference between the mean and each data. The obtained regression equation is known as linear and nonlinear when the fit line is a straight line and a curve, respectively. In the sense of mathematic, a regression model could be expressed as Eq. 5.4. When *f* is linear in $\theta$, *y* will be linear while when *f* is nonlinear in $\theta$, *y* will be a nonlinear regression model (Huang, Hsiao et al. 2010, Dabiri, Kioumarsi et al. 2022, Dabiri, Rahimzadeh et al. 2022).

$$y = f(x_i : \theta) + \varepsilon \qquad (5.4)$$

where *f* is the function showing the relationship between inputs and outputs, $x_i$ are the inputs, $\theta$ are the parameters and $\varepsilon$ is a random variable error with mean=0 and standard deviation=σ. NLR generally yields more accurate predicted values than linear regression since it can fit a much wider range of curves (nonlinear relationship) between its variables (Hasanuzzaman and Abd Rahim 2019, Wood and Cai 2021, Dabiri, Rahimzadeh et al. 2022). Moreover, NLR finds the most reliable fit by minimizing the sum of squares of the distance between the actual and model prediction values (RSS), as given in Eq. 5.5 (Sanft and Walter 2020):

$$\mathrm{RSS}\,(P_1, \dots, P_m) = \sum_{i=1}^{n} e_i^2, \qquad e_i = z_i - g(t_i, P_1, \dots, P_m) \qquad (5.5)$$

where $z_i$ is the prediction values, g are the actual values and $P_i$ are the parameters.

### *5.4.2 Decision Tree (DT)*

A decision tree is a method which uses a tree-shape graph for moving a dataset sample to the most accurate target output based on its characteristics. DT could be used for solving both classification (finite set of values) and regression (continues values) problems (Liu, McGree et al. 2015, Bellini 2019). The process of predicting a value is schematically depicted in Fig. 5.3. The most assets of DT technique are: (i) being simple for understanding and interpreting, (ii)

being independent of the nonlinear relation between parameters, and (iii) being flexible to be adopted to new scenarios (Kotu and Deshpande 2018, Shobha and Rangaswamy 2018).

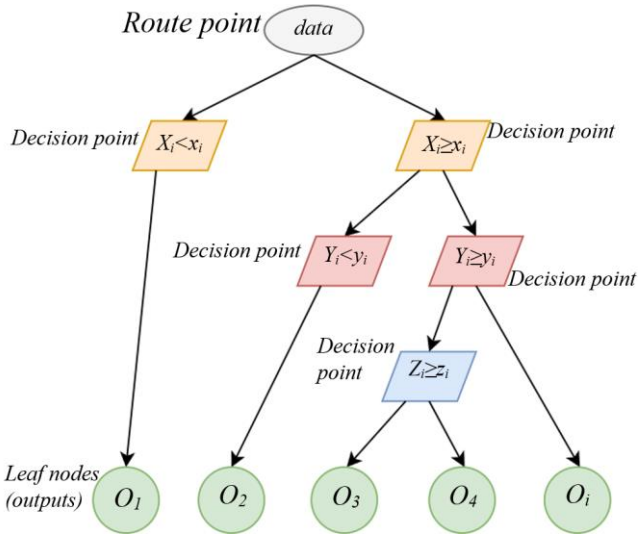



***Figure 5-3**. Schematic illustration of a DT prediction model.*

### *5.4.3 Random Forest (RF)*

In the simplest concept, RF consists of many decision trees and target output is predicted by considering either the average of the DTs' predicted values or the most voted value. To explain more precisely, RF is basically the combination of Bagging and Random selection of features by creating various decision trees. The most notable point about RF is that selecting training dataset for each tree is done through Bootstrap sampling and the feature which is chosen as the decision node is a random subset of the main dataset (Mao and Wang 2012, Paul and Bhatia 2020, Williams, Halloin et al. 2020). The above-mentioned explanations are simply demonstrated in Fig. 5.4.

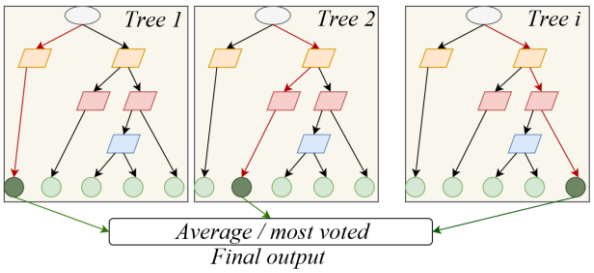
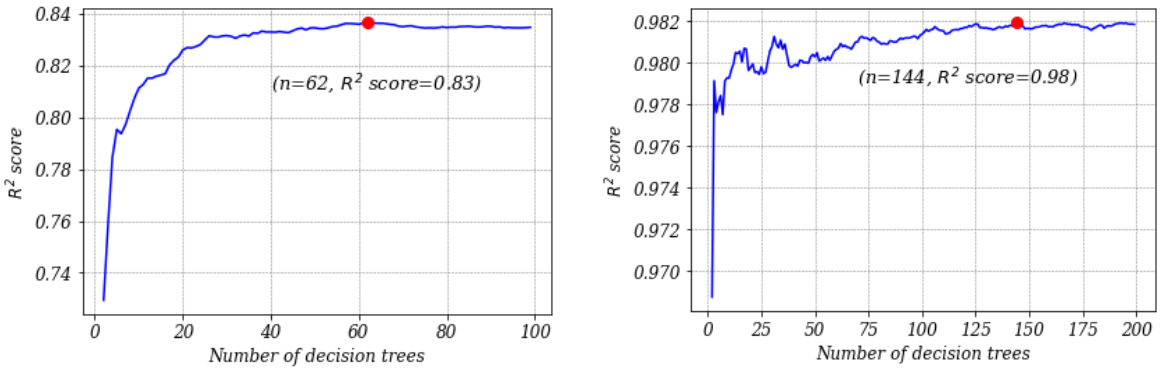



***Figure 5-4**. A simplified illustration of a RF model.*

In order to find out the most efficient number of trees in our RF model, $R^2$-score of various RFs with different numbers of trees was obtained (Fig. 5.5). As could be observed in Fig. 5.5,

62 and 144 trees led to the highest accuracy of the RF models developed for predicting $\beta$ and $\mu$, respectively.

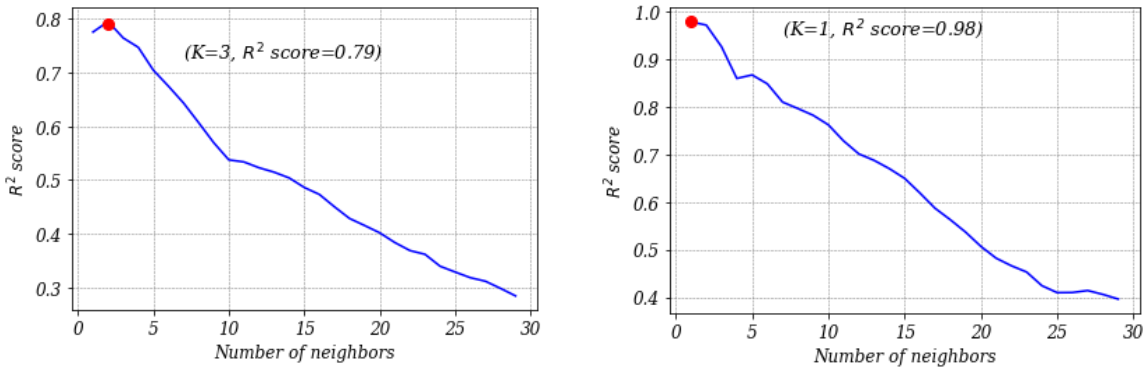



(a) (b)

***Figure 5-5****. The most efficient number of trees in the RF models developed for predicting (a) β and (b) μ.*

### *5.4.4 K-Nearest Neighbours*

K-nearest Neighbours (KNN) is recognized as a non-parametric prediction method which means that the prediction process is not influenced by the relationship between input and output parameters (Chanal, Steiner et al. 2021). Therefore, KNN is known as the simplest classification approach by data scientists. In this method, the data are plotted in a multi-dimensional space where the axis are the data's features. When a new data is added to the space according to its characteristics, the average of its "K" nearest neighbours defines the target output (Richman 2011, Subasi, Khateeb et al. 2020). As a result, the most effective parameter which could enhance the accuracy of predicted values is the number of nearest neighbours defined as "K". Although some researchers have claimed that square root of the total number of datasets could be the most accurate K (Subasi, Khateeb et al. 2020), it is generally obtained by trial-and-error process. The best K values obtained for predicting $\beta$ and $\mu$ are 3 and 1, respectively, as depicted in Fig. 5.6.




(a) (b)

***Figure 5-6**. Finding the most accurate K for the KNN models for predicting (a) β and (b) μ.*

### *5.4.5 Artificial Neural Network*

ANN is basically a combination of computation and mathematics which is inspired from the human brain by simulating the performance of the brain's nervous system (Richman 2011, Malekian and Chitsaz 2021). The architecture of an ANN model consists of (i) input layers: the number of the nodes in this layer is equal to the number of model inputs, (ii) hidden layer(s) which might be considered to enhance the model accuracy, and (iii) output layers: like input layer, the number of nodes is determined based on the number of target outputs. The nodes are connected to each other by weights which are updated in each iteration in order to reach an acceptable estimated output (Sadiq, Rodriguez et al. 2019, Malekian and Chitsaz 2021). Other optional components which might be considered in an ANN model are: (iv) bias values which could be defined for hidden and output layers and (v) activation functions which might be added to the weights in order to allow the ANN to account for nonlinear behaviour in the training dataset. Otherwise noted, an Ann model without an activation function can perform linearly with unreliable accuracy (Walczak and Cerpa 1999, Sadiq, Rodriguez et al. 2019). An ANN model could be formulized as below:

$$O_j = f\sum(w_{ij}I_i + b) \qquad (5.5)$$

Where $O_j$ is the model output, $w_{ij}$ is the associated weight which is updated in each epoch, $I_i$ is input data and $b$ is bias (Dabiri, Rahimzadeh et al. 2022).

Based on the above-mentioned definitions, number of hidden layers, number of nodes in each hidden layer and type of the activation function are the main factors which affect the performance of an ANN model significantly. These parameters are typically obtained through a trial-and-error process. The architecture of the ANN model developed for presenting $\beta$ and $\mu$

is displayed in Fig. 5.7. It should be noted that Rectified Linear activation function or ReLu function was defined as the activation function in the models.

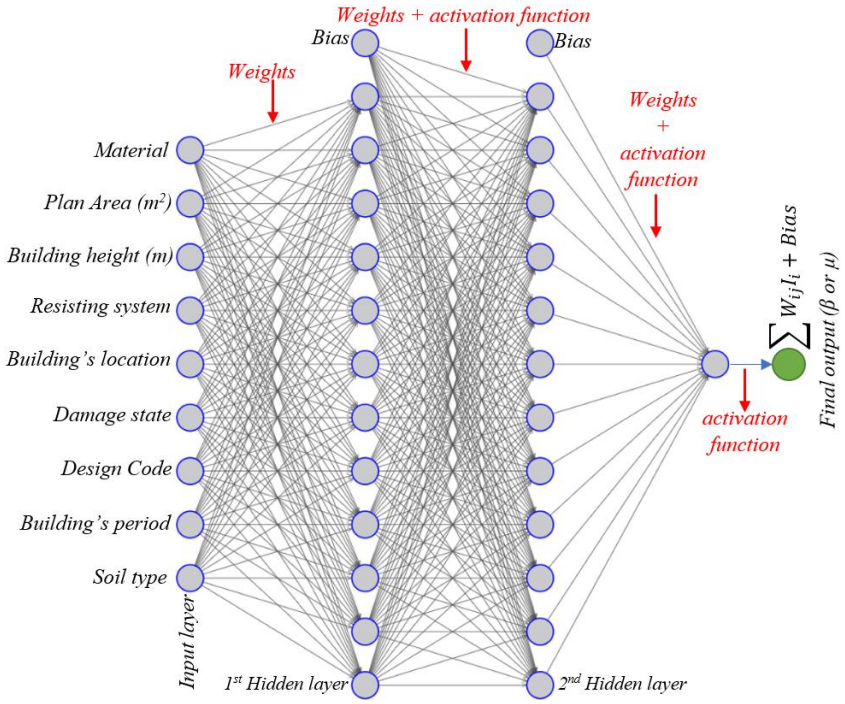



***Figure 5-7****. The architecture of the ANN models developed in this study for predicting* $\beta$ *and* $\mu$*.*

## 5.5 Results

The prediction models were developed by adjusting their characteristics using training sub-databases. Then, the target outputs of both training and testing datasets were predicted by the models. The correlation between the predicted and actual values is shown in Fig. 5.8 and Fig. 5.9 respectively for $\beta$ and $\mu$. The green solid line shows the ideal line (predicted = actual) while the read dotted lines illustrate 80% upper and lower values of the ideal line.

In order to figure out the accuracy of the models more precisely, each predicted value of $\beta$ and $\mu$ is compared with its corresponding actual value as displayed in Fig. 5.10 and Fig. 5.11, respectively.

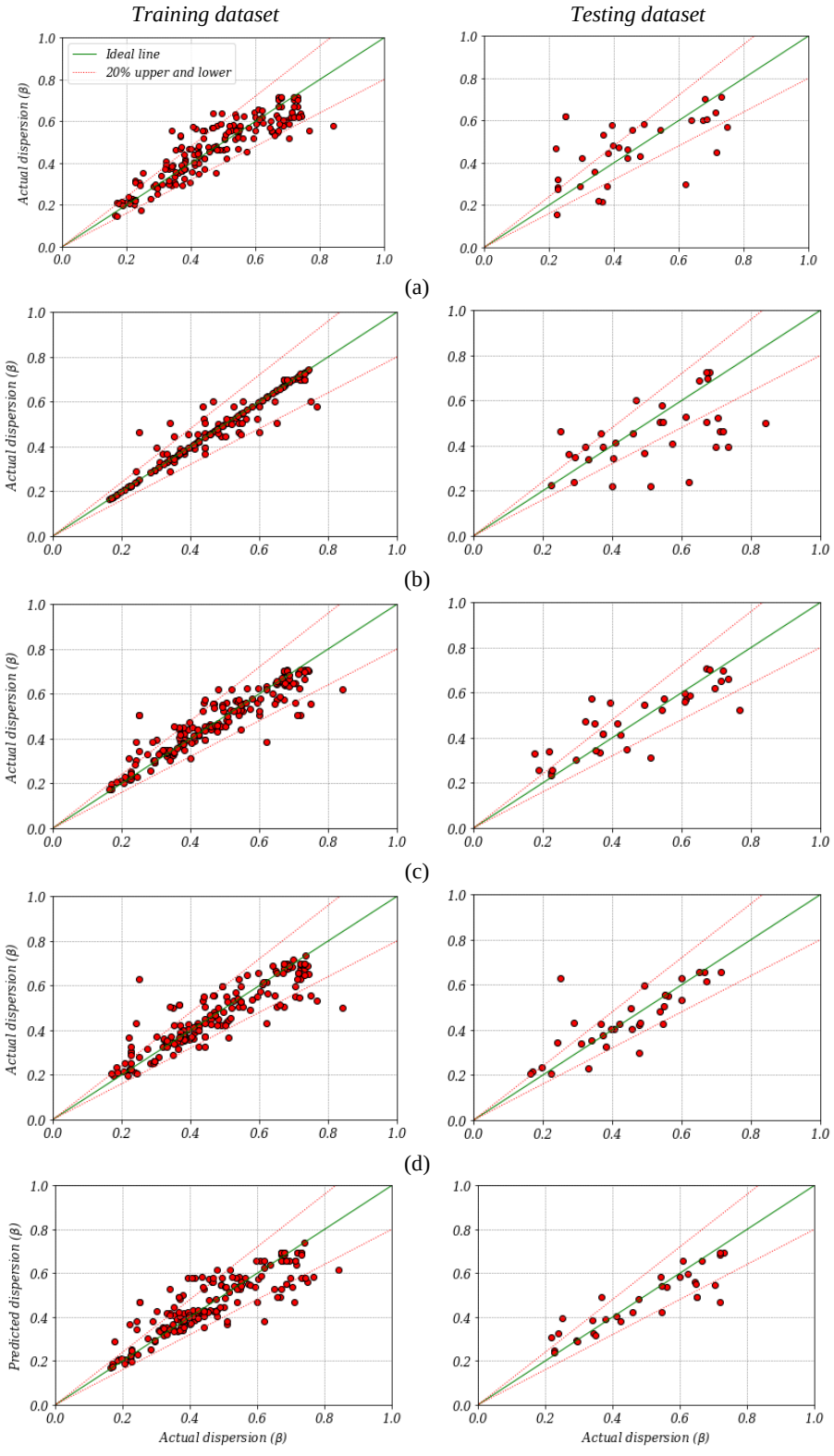



***Figure 5-8**. Correlation between actual and predicted values of β: (a) NLR, (b) DT, (c) RF, (d) KNN, (e) ANN.*

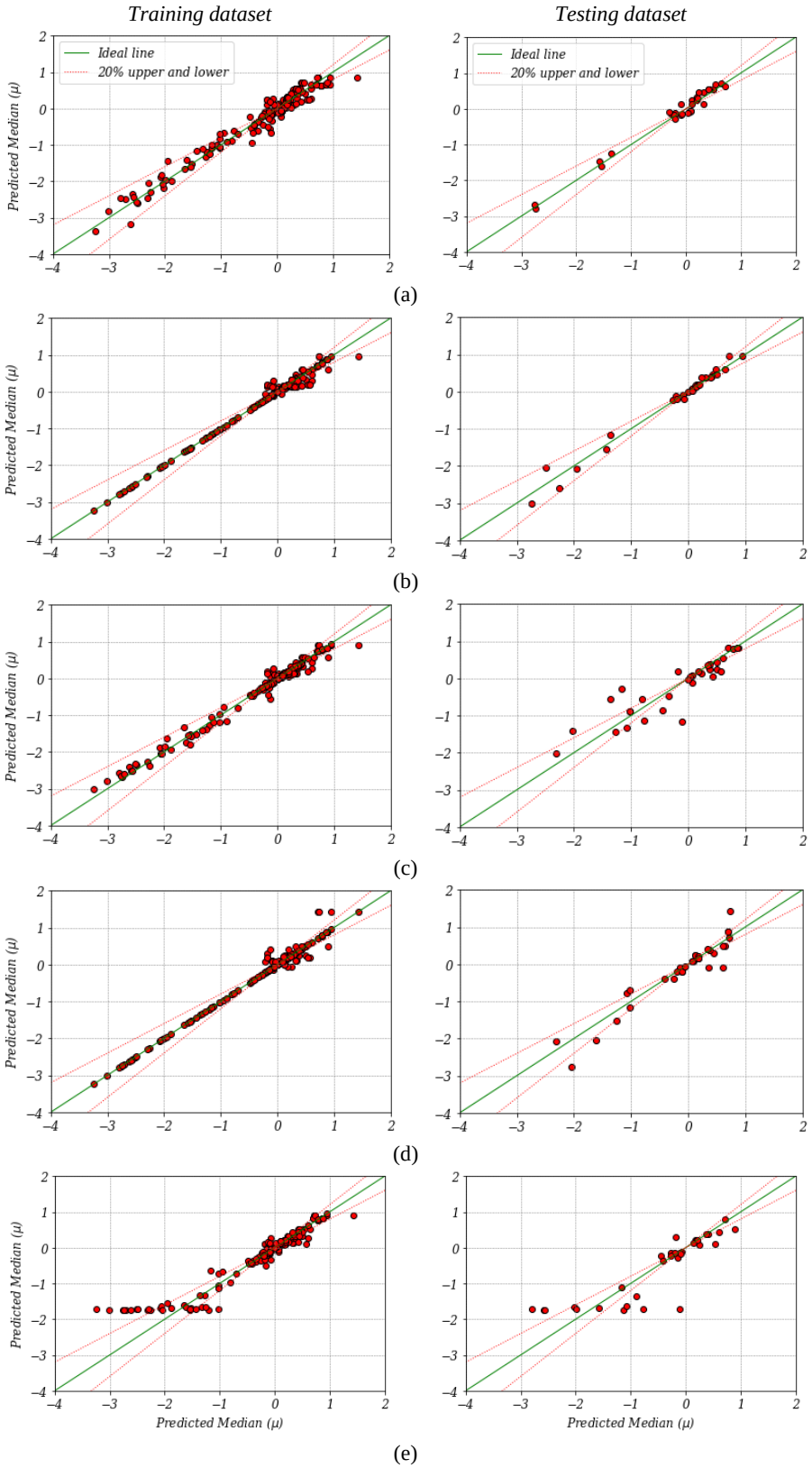



***Figure 5-9****. Correlation between actual and predicted values of μ (log (PGA)): (a) NLR, (b) DT, (c) RF, (d) KNN, (e) ANN.*

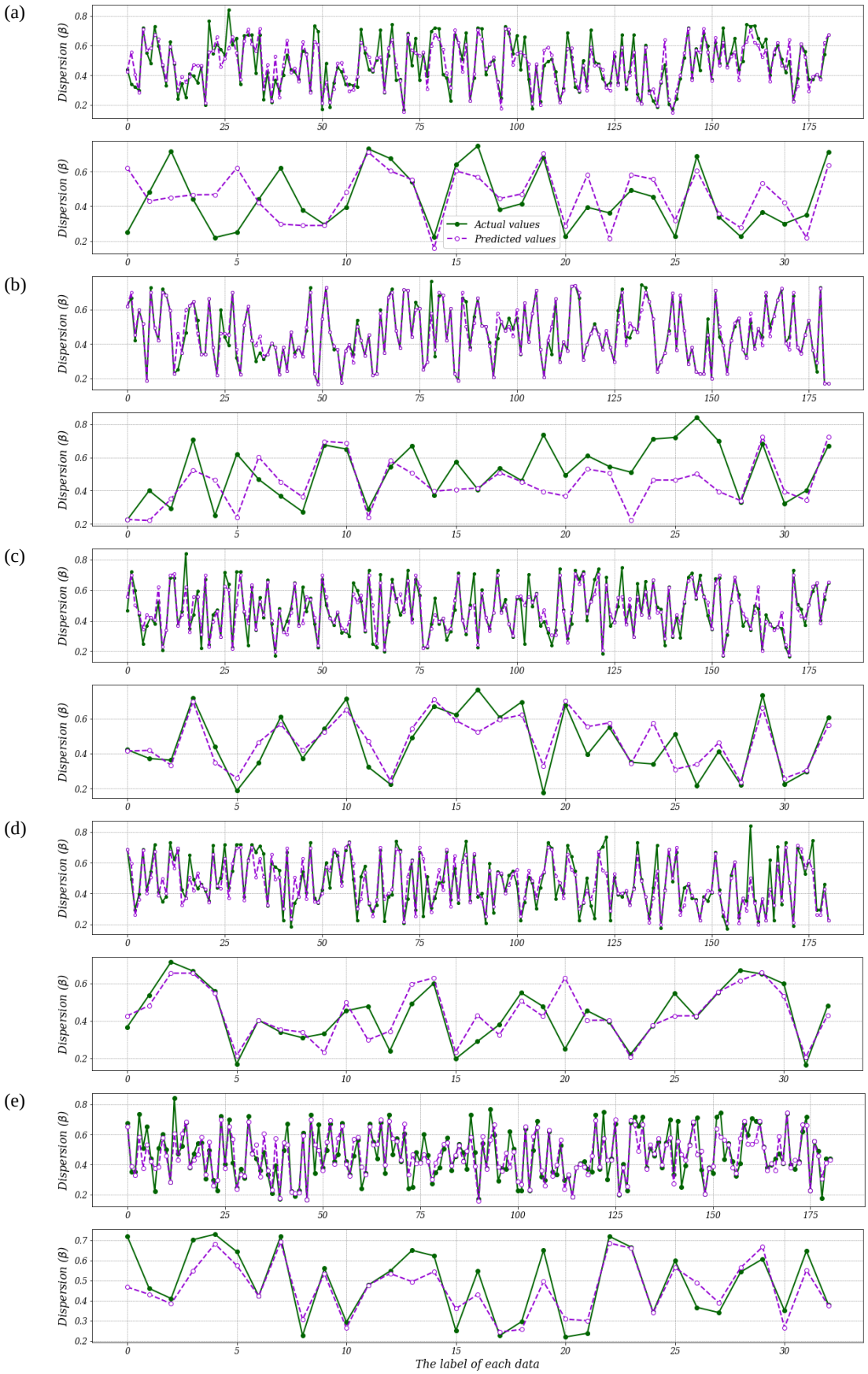



***Figure 5-10****. Comparing the predicted and actual values of β (a) NLR, (b) DT, (c) RF, (d) KNN, (e) ANN.*

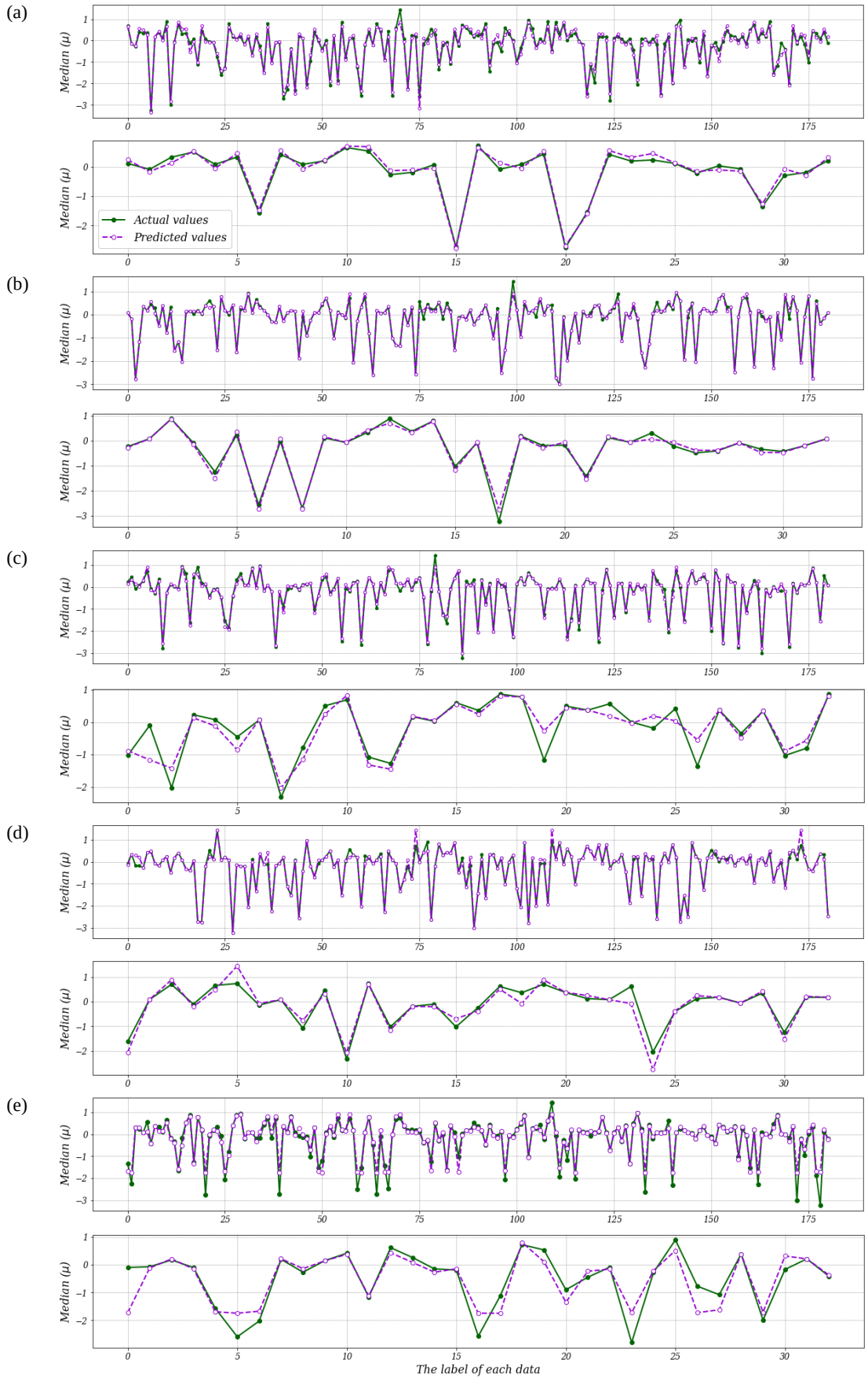



***Figure 5-11****. Comparing the predicted and actual values of μ (log (PGA)) (a) NLR, (b) DT, (c) RF, (d) KNN, (e) ANN.*

At the first glance on figures 5.7-5.10, it could be claimed that almost all the methods exhibited the ability to learn the relationship between input variables and two target outputs. More specifically, DT showed good correlation between actual and predicted values of both $\beta$ and $\mu$ in Figures 5. 10(b) and 5.11(a). However, the models need to be assessed more precisely as presented in the next section.

## 5.6 Accuracy assessment

The accuracy of the prediction models proposed in this study is assessed through common performance metrics and Taylor diagram.

### *5.6.1 Performance metrics*

The performance metrics which are commonly used for assessing accuracy of a model are Root of Mean Square Error (RMSE), Mean Absolut Error (MAE), Mean Absolute Percentage Error (MAPE) and R2-score. These parameters are calculated using Eq. (5.6-5.9).

$$\mathrm{RMSE} = \left(\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2\right)^{0,5} \tag{5.6}$$

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|\hat{y}_i - y_i| \tag{5.7}$$

$$\mathrm{MAPE} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{\hat{y}_i - y_i}{y_i}\right| \tag{5.8}$$

$$R^2 = 1 - \frac{\sum_i(\hat{y}_i - y_i)^2}{\sum_i(y_i - \bar{y}_i)^2} \tag{5.9}$$

where $y$ is the actual output, $\hat{y}$ is the predicted output, $n$ is the number of data records and $\bar{y}$ is the mean of the dataset. The higher $R^2$-score and the lower RMSE, MAE and MAPE are, the more accurate the model will be. Table 5.6 compares the performance metrics of all the models for estimating $\beta$ and $\mu$ quantitatively.

***Table 5-6****. Performance metrics of the developed models.*

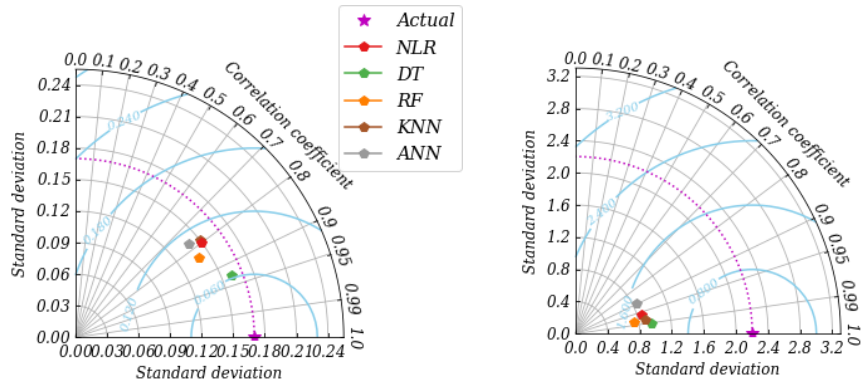
| | *Dispersion ($\beta$) prediction models* | | | | *Median ($\mu$, log (PGA)) prediction models* | | | |
|---|---|---|---|---|---|---|---|---|
| | *$R^2$* | *RMSE* | *MAE* | *MAPE* | *$R^2$* | *RMSE* | *MAE* | *MAPE* |
| *Nonlinear* | *0.80* | *0.07* | *0.05* | *12.08* | *0.96* | *0.19* | *0.15* | *104.69* |
| ***DT*** | ***0.93*** | ***0.04*** | ***0.02*** | ***3.91*** | ***0.99*** | ***0.10*** | ***0.04*** | ***32.71*** |
| *RF* | *0.84* | *0.06* | *0.04* | *9.57* | *0.98* | *0.12* | *0.08* | *41.99* |
| *KNN* | *0.80* | *0.07* | *0.05* | *44.81* | *0.98* | *0.13* | *0.05* | *669.15* |
| *ANN* | *0.75* | *0.08* | *0.06* | *13.29* | *0.91* | *0.28* | *0.15* | *51.71* |

Regarding the values reported for the proposed models, it could be observed that the DT model processes the highest $R^2$-score (93% and 99% for estimating $\beta$ and $\mu$, respectively) and hence is introduced as the most accurate model.

It is worth mentioning that almost all the proposed models have an acceptable level of ability to learn the relationship between the inputs and outputs which reflect the high reliability of the ML-based methods for predicting fragility curves of buildings.

### *5.6.2 Taylor diagram*

In order to compare the accuracy of the models more easily, they are compared in Taylor diagram shown in Fig. 5.12. It should be explained that Taylor diagram consists of three main parts: (i) horizontal and vertical axis which reflects standard deviation, (ii) circular curves centred at actual values which reflect RSME, and (iii) radial line which reflects $R^2$-score. In other words, each model is plotted in a Taylor diagram by its standard deviation, RMSE and $R^2$-score. The closest model to the actual value is known as the most accurate model (Shariati, Mafipour et al. 2020, Dabiri, Rahimzadeh et al. 2022, Wadoux, Walvoort et al. 2022).




(a) (b)

***Figure 5-12****. Using Taylor diagram for comparing the models developed for predicting (a) β and (b) μ (log(PGA)).*

Considering Diagrams illustrated in Fig. 5.12, and the above-mentioned explanations, the DT models are the closest models to the star point (actual values) and thus are introduced in as the most accurate models for predicting both $\beta$ and $\mu$, as concluded in the previous section.

## 5.7 Sensitivity analysis

In this section, the influence of buildings properties namely construction material (i.e., RC, steel and masonry), plan area, building height and period, and soil type on the fragility parameters is assessed. A two-story (6.20m height) masonry building located on soil type C (according to (Eurocode 8 2004) soil classification) in L'Aquila, Italy is selected from the literature (Suzuki and Iervolino 2021) and considered as the reference dataset. Then, the above-

mentioned characteristics are varied to generate a series of 26 datasets, as reported in Table 5.10. The first row of Table 5.10 represents the reference building which is used for assessing the influence of parameters' variation on both $\beta$ and $\mu$. The material of the second and third building is changed to reinforced concrete and steel, respectively. Then, the plan area of the reference building is varied in the range of 100-1000 $m^2$ by the step of 100. The building height is also increased from 3m to 30 m by the step of 3. It should be noted that building period is influenced by its height and therefore the variation of either $\beta$ or $\mu$ by changing building's period is almost the same as that of changing building's height. However, to figure out the effect of building period on the fragility parameters, it is varied according to height variation. Eventually, three soil classifications including A (rock), B (stiff) and C (soft) are taken into account. Fragility parameters, $\beta$ and $\mu$, of the generated database are estimated by the prediction model developed using the Decision Tree model. It is worth explaining that the DT model is used for prediction because it showed the highest accuracy for predicting both $\beta$ and $\mu$. The predicted values and their differences compared to the reference case is provided in Table 5.7.

***Table 5-7****. Variation of β and μ by changing building's properties and soil type.*

| | Material | Area (m2) | Height (m) | Period | Soil type | β | $(\beta_i-\beta_0)/\beta_0$ | μ | $(\mu_i-\mu_0)/\mu_0$ |
|---|---|---|---|---|---|---|---|---|---|
| 1 | **Masonry** | 220.00 | 6.20 | 0.12 | C (soft) | 0.38 | 0.00 | -0.08 | 0.00 |
| 2 | **RC** | 220.00 | 6.20 | 0.12 | C (soft) | 0.74 | 0.95 | -0.07 | -13.11 |
| 3 | **Steel** | 220.00 | 6.20 | 0.12 | C (soft) | 0.72 | 0.88 | 0.47 | -652.91 |
| 4 | Masonry | **100.00** | 6.20 | 0.12 | C (soft) | 0.62 | 0.63 | -0.07 | -13.20 |
| 5 | Masonry | **200.00** | 6.20 | 0.12 | C (soft) | 0.36 | -0.07 | -0.07 | -13.20 |
| 6 | Masonry | **300.00** | 6.20 | 0.12 | C (soft) | 0.37 | -0.03 | -0.07 | -13.20 |
| 7 | Masonry | **400.00** | 6.20 | 0.12 | C (soft) | 0.37 | -0.03 | -0.09 | 4.58 |
| 8 | Masonry | **500.00** | 6.20 | 0.12 | C (soft) | 0.37 | -0.03 | -0.09 | 4.58 |
| 9 | Masonry | **600.00** | 6.20 | 0.12 | C (soft) | 0.37 | -0.03 | -0.09 | 4.58 |
| 10 | Masonry | **700.00** | 6.20 | 0.12 | C (soft) | 0.37 | -0.03 | 0.96 | -1238.88 |
| 11 | Masonry | **800.00** | 6.20 | 0.12 | C (soft) | 0.37 | -0.03 | 0.96 | -1238.88 |
| 12 | Masonry | **900.00** | 6.20 | 0.12 | C (soft) | 0.37 | -0.03 | 0.96 | -1238.88 |
| 13 | Masonry | **1000.00** | 6.20 | 0.12 | C (soft) | 0.37 | -0.03 | 0.96 | -1238.88 |
| 14 | Masonry | 220.00 | **3.00** | 0.13 | C (soft) | 0.36 | -0.07 | -0.07 | -13.20 |
| 15 | Masonry | 220.00 | **6.00** | 0.21 | C (soft) | 0.42 | 0.11 | -0.07 | -13.20 |
| 16 | Masonry | 220.00 | **9.00** | 0.29 | C (soft) | 0.24 | -0.38 | -0.09 | 6.18 |
| 17 | Masonry | 220.00 | **12.00** | 0.35 | C (soft) | 0.18 | -0.54 | -0.09 | 6.18 |
| 18 | Masonry | 220.00 | **15.00** | 0.42 | C (soft) | 0.18 | -0.54 | -0.09 | 6.18 |
| 19 | Masonry | 220.00 | **18.00** | 0.48 | C (soft) | 0.18 | -0.54 | -0.09 | 6.18 |
| 20 | Masonry | 220.00 | **21.00** | 0.54 | C (soft) | 0.18 | -0.54 | -0.09 | 6.18 |
| 21 | Masonry | 220.00 | **24.00** | 0.60 | C (soft) | 0.18 | -0.54 | -0.09 | 6.18 |
| 22 | Masonry | 220.00 | **27.00** | 0.65 | C (soft) | 0.18 | -0.54 | -0.09 | 6.18 |
| 23 | Masonry | 220.00 | **30.00** | 0.71 | C (soft) | 0.18 | -0.54 | -0.09 | 6.18 |
| 24 | Masonry | 220.00 | 6.20 | 0.12 | **A (rock)** | 0.45 | 0.18 | 0.02 | -124.61 |
| 25 | Masonry | 220.00 | 6.20 | 0.12 | **B (stiff)** | 0.45 | 0.18 | -0.07 | -13.20 |
| 26 | Masonry | 220.00 | 6.20 | 0.12 | **C (soft)** | 0.36 | -0.07 | -0.07 | -13.20 |

### *5.7.1 Influence of soil type and building properties on β*

Figure 5.13 displays the variation of β when soil type or building parameters change. In terms of construction material (Fig. 5.13a), it could be claimed that RC and steel structures with similar response, exhibited considerably higher dispersion than that of masonry buildings. As far as building geometry is concerned, it could be realized that by increasing both plan area and height of a building, dispersion alters notably. Fig. 5.13b, reveals that dispersion drops when building area is increased and then, it remains almost constant for larger areas (in this case study 200 $m^2$ with the height to area ratio of 3.1%). Roughly the same variation could be figured out that for building's height variation according to Fig. 5.13c. More clearly, when the height increases from 3 (low-rise) to 6 (mid-rise), dispersion increases by 27.61%. Then, a sharp decrease is observed for higher buildings with 6-12 m (high-rise) height. Dispersion of tall

buildings (more than 12 m height in this study), however, remains approximately unchanged. Fig. 5.13d shows that the relationship between building's period and dispersion is the same as that of building's height, since period is directly affected by the building's height as explained previously. Using the dispersion values predicted by the developed DT model and the quantitative building properties namely plan area (A, $m^2$), height (H, m) and period (T, s), fitting curves are drawn as depicted by black dotted lines in Fig. 5.13 b-d. Accordingly, the following equations are suggested for obtaining dispersion:

$$\beta = 0{,}05083 \times 10^{-7}A^2 - 0{,}0007A + 0{,}5845 \quad (5.10)$$

$$\beta = 6 \times 10^{-4}H^2 - 0{,}0283H + 0{,}4708 \quad (5.11)$$

$$\beta = 1{,}268T^2 - 1{,}425T + 0{,}5615 \quad (5.12)$$

Fig. 5.13e illustrates how soil type could affect dispersion. Based on this figure, it could be claimed that dispersion of the buildings located on soft soil (type C) possess lower dispersion compared to those located on either rock (type A) or stiff soil (type B). The same results have been reported in similar studies evaluating the effect of soil type on fragility parameters of buildings (Pejovic and Jankovic 2016, Suzuki and Iervolino 2021).

### *5.7.2 In fluence of soil type and building properties on μ*

The effect of building properties and soil type on median is demonstrated in Fig. 5.14. It is noteworthy that $\mu$ is the logarithmic median of ln (PGA). Just like $\beta$, median values of RC and steel structures are higher than that of the masonry building (Fig. 5.14a). Furthermore, by increasing the building's plan area, median value increases notably first. Then, it remains almost the same for larger areas (600 $m^2$ in this study), as displayed in Fig. 5.14b. The buildings on soil types A have higher $\mu$ in comparison to those on soil type B or C as depicted in Fig. 5.14e. By increasing the building's height or period, like dispersion variation, median values reduce first (building height=9 m and height/area=4.09% in this study). Then after, it remains roughly constant for higher buildings representing mid- and high-rise buildings. The prediction equations suggested based on the estimated values and corresponding fitting curves are given in Eq. (5.13-5.15):

$$\mu = 2 \times 10^{-6}A^2 - 0{,}0003A - 0{,}1404 \quad (5.13)$$

$$\mu = 5 \times 10^{-5}H^2 - 0{,}0023H - 0{,}0666 \quad (5.14)$$

$$\mu = 0{,}1054T^2 - 0{,}1129T - 0{,}0605 \quad (5.15)$$

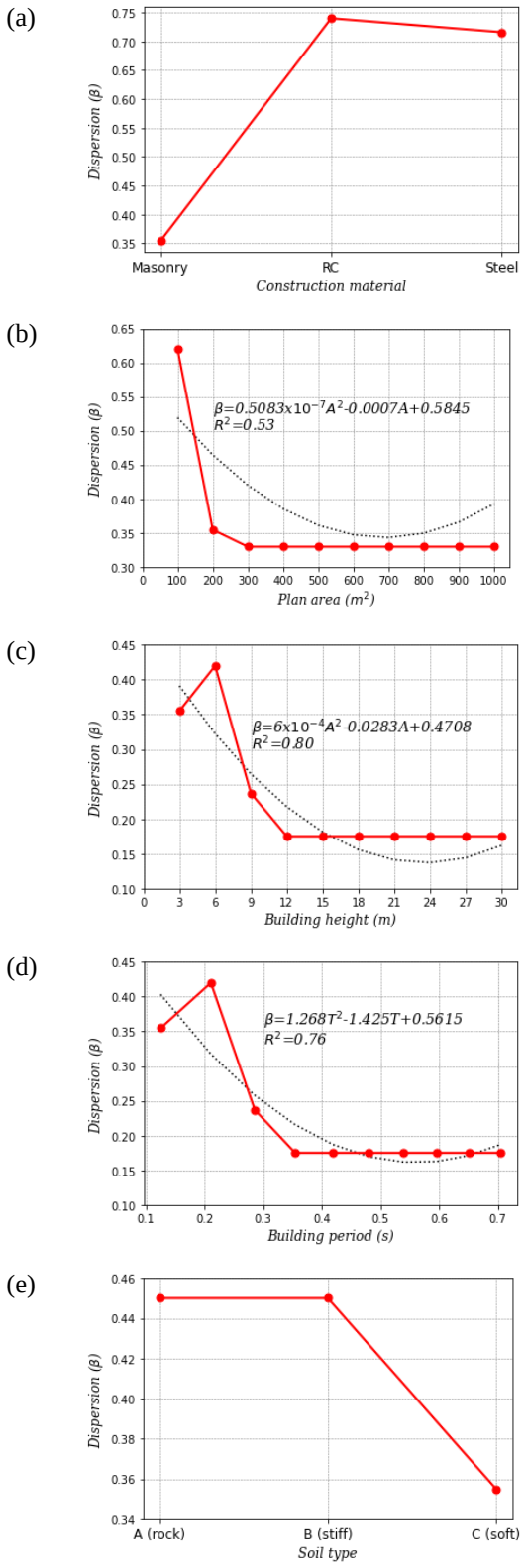



***Figure 5-13**. Influence of building properties and soil type on dispersion (β).*

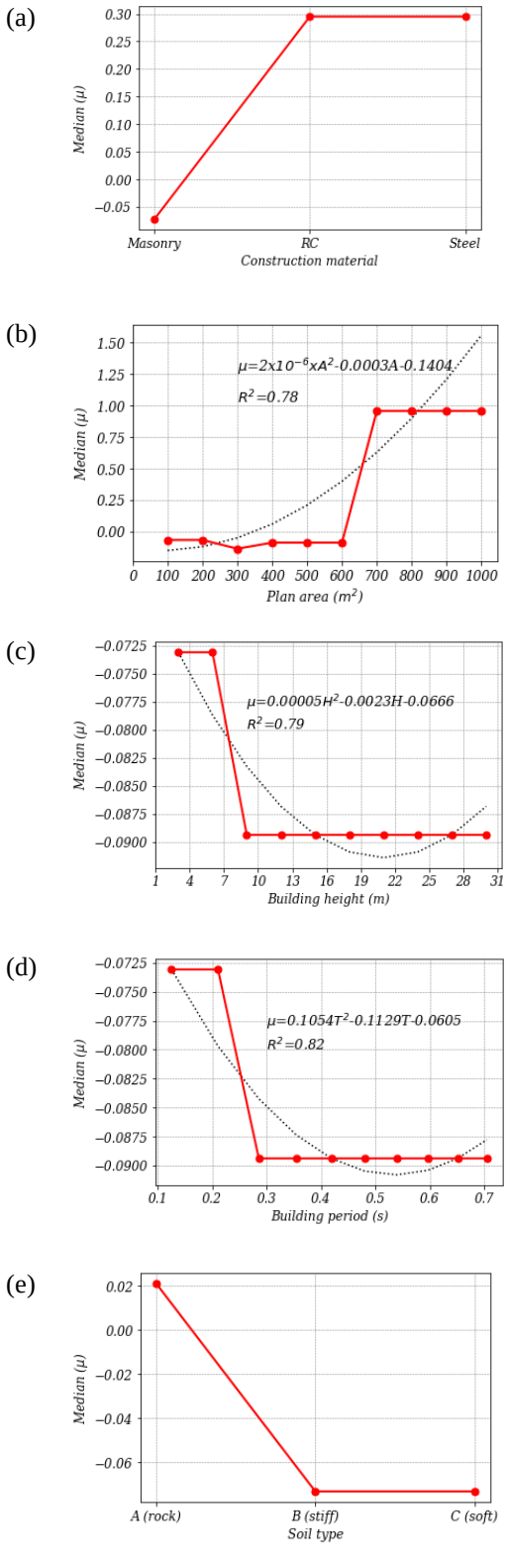



***Figure 5-14**. Influence of building properties and soil type on median (μ).*

## 5.8 Summary and conclusion

Fragility curves are one of the crucial means which should be obtained for risk assessment of buildings in the PEER framework. The process of deriving fragility curves, however, is time-consuming and complicated. These issues might increase inaccuracy of the fragility curves. Therefore, proposing a quick and error-free alternative approach for generating fragility curves has become one of the main researchers' concerns in the field of structural and earthquake engineering. In this study, an attempt has been made to develop ML-based models for predicting fragility parameters of buildings and hence deriving fragility curves. To this end, a dataset including 214 fragility curve parameters are collected from peer-reviewed publications. They were then divided into training (85%) and testing (15%) sub-databases. Parameters which are proven to have the highest effect on fragility parameters are considered as the inputs for predicting dispersion ($\beta$) and median ($\mu$) of fragility curves. Various ML-based prediction models namely NLR, DT, RF, KNN and ANN were developed. The predicted values were compared with the actual values and the accuracy of the models were assessed through common performance metrics and Taylor diagram. Eventually, a parametric study conducted and equations were presented for calculating fragility parameters. The main conclusions are:

- Almost all the ML-based techniques showed a high ability to learn the relationship between input and output parameters. They could be therefore considered as a quick and accurate model for estimating fragility curves instead of time-consuming and inaccurate analytical analysis.
- All the developed models for predicting $\beta$ exhibited high reliability with $R^2$-score≥0.75. The DT model with $R^2$-score=0.91, however, was the most accurate model.
- Among the models proposed for estimating $\mu$, DT possessed the highest $R^2$-score compared to other models and hence is introduced as the most accurate model.
- Based on the study results, DT is recommended for deriving fragility curves of buildings.
- It is also worth mentioning that, unlike other studies carried out for developing models for obtaining fragility curves, the models of the present study have the following benefits: (i) generating fragility curves directly by estimating $\beta$ and $\mu$, (ii) considering significant parameters which influence fragility curves notably, and (iii) being generalized which means that they could be applied for RC, steel and masonry buildings with different heights and plan areas located on different soil types.

- The parametric study on a case study illustrated that, (i) $\beta$ and $\mu$ of RC or steel buildings were higher than that of the masonry building, (ii) $\beta$ of the buildings located on either rock or stiff soil were higher than those of the buildings located on soft soil, while $\mu$ of the buildings located on rock was higher than those located on either stiff and soft soil, (iii) by increasing building plan area, $\beta$ reduced first (up to height/area ratio=3.1) and then remained almost constant; moreover, $\mu$ of the building with area equal or higher than 600-700 $m^2$ are higher than other buildings considered in the sensitivity analysis and (iv) increasing building height led to initial reduction and increase respectively in $\beta$ and $\mu$, while fragility curve parameters of tall buildings remained unchanged.

# References

Abo-El-Ezz, A., et al. (2013). "Seismic fragility assessment of low-rise stone masonry buildings." Earthquake Engineering and Engineering Vibration **12**(1): 87-97.

Alani, A. M. and A. Faramarzi (2014). "An evolutionary approach to modelling concrete degradation due to sulphuric acid attack." Applied Soft Computing **24**: 985-993.

Aljawhari, K., et al. (2021). "Effects of ground-motion sequences on fragility and vulnerability of case-study reinforced concrete frames." Bulletin of Earthquake Engineering **19**(15): 6329-6359.

Alwaeli, W., et al. (2020). "Rigorous versus less-demanding fragility relations for RC high-rise buildings." Bulletin of Earthquake Engineering **18**(13): 5885-5918.

Beilic, D., et al. (2017). "Seismic fragility curves of single storey RC precast structures by comparing different Italian codes." Earthq. Struct **12**(3): 359-374.

Bellini, T. (2019). IFRS 9 and CECL Credit Risk Modelling and Validation: A Practical Guide with Examples Worked in R and SAS, Academic Press.

Berman, J. J. (2016). Data simplification: taming information with open source tools, Morgan Kaufmann.

Berman, J. J. (2018). Principles and practice of big data: preparing, sharing, and analyzing complex information, Academic Press.

Cardone, D., et al. (2019). "Developing collapse fragility curves for base-isolated buildings." Earthquake Engineering & Structural Dynamics **48**(1): 78-102.

Cardone, D., et al. (2018). "Estimating fragility curves of pre-70 RC frame buildings considering different performance limit states." Soil Dynamics and Earthquake Engineering **115**: 868-881.

Chanal, D., et al. (2021). "Online Diagnosis of PEM Fuel Cell by Fuzzy C-Means Clustering."

Dabiri, H., et al. (2022). "Compressive strength of concrete with recycled aggregate; a machine learning-based evaluation." Cleaner Materials **3**: 100044.

Dabiri, H., et al. (2022). A comparison of machine learning-and regression-based models for predicting ductility ratio of RC beam-column joints. Structures, Elsevier.

Dall'Asta, A., et al. (2021). "Influence of time-dependent seismic hazard on structural design." Bulletin of Earthquake Engineering **19**(6): 2505-2529.

Del Gaudio, C., et al. (2019). "Empirical fragility curves for masonry buildings after the 2009 L'Aquila, Italy, earthquake." Bulletin of Earthquake Engineering **17**(11): 6301-6330.

Donà, M., et al. (2021). "Mechanics-based fragility curves for Italian residential URM buildings." Bulletin of Earthquake Engineering **19**(8): 3099-3127.

Eurocode 8 (2004). Design of structures for earthquake resistance - Part 1: General rules, seismic actions and rules for buildings, European Standard EN 1998-1:2004.

Faramarzi, A., et al. (2014). "An EPR-based self-learning approach to material modelling." Computers & Structures **137**: 63-71.

Faramarzi, A., et al. (2013). "Numerical implementation of EPR-based material models in finite element analysis." Computers & Structures **118**: 100-108.

Frankie, T. M., et al. (2013). "Simulation-based fragility relationships for unreinforced masonry buildings." Journal of Structural Engineering **139**(3): 400-410.

Hasanuzzaman, M. and N. Abd Rahim (2019). Energy for sustainable development: demand, supply, conversion and management, Academic Press.

Huang, H.-H., et al. (2010). "Nonlinear regression analysis." International encyclopedia of education: 339-346.

Hwang, S.-H., et al. (2021). "Machine learning-based approaches for seismic demand and collapse of ductile reinforced concrete building frames." Journal of Building Engineering **34**: 101905.

Jia, D.-W. and Z.-Y. Wu (2021). "Seismic fragility analysis of RC frame-shear wall structure under multidimensional performance limit state based on ensemble neural network." Engineering Structures **246**: 112975.

Karafagka, S., et al. (2021). "Fragility curves of non-ductile RC frame buildings on saturated soils including liquefaction effects and soil–structure interaction." Bulletin of Earthquake Engineering **19**(15): 6443-6468.

Kiani, J., et al. (2019). "On the application of machine learning techniques to derive seismic fragility curves." Computers & Structures **218**: 108-122.

Kirçil, M. S. and Z. Polat (2006). "Fragility analysis of mid-rise R/C frame buildings." Engineering Structures **28**(9): 1335-1345.

Kotu, V. and B. Deshpande (2018). Data science: concepts and practice, Morgan Kaufmann.

Kumar, P. and A. Samanta (2020). Seismic fragility assessment of existing reinforced concrete buildings in Patna, India. Structures, Elsevier.

Lange, D., et al. (2014). "An application of the PEER performance based earthquake engineering framework to structures in fire." Engineering Structures **66**: 100-115.

Liu, S., et al. (2015). Computational and statistical methods for analysing big data with applications, Academic Press.

Malekian, A. and N. Chitsaz (2021). Concepts, procedures, and applications of artificial neural network models in streamflow forecasting. Advances in Streamflow Forecasting, Elsevier**:** 115-147.

Mao, W. and F. Wang (2012). New advances in intelligence and security informatics, Academic Press.

Mitropoulou, C. C. and M. Papadrakakis (2011). "Developing fragility curves based on neural network IDA predictions." Engineering Structures **33**(12): 3409-3421.

Nettleton, D. (2014). Commercial data mining: processing, analysis and modeling for predictive analytics projects, Elsevier.

Özel, A. E. and E. M. Güneyisi (2011). "Effects of eccentric steel bracing systems on seismic fragility curves of mid-rise R/C buildings: A case study." Structural Safety **33**(1): 82-95.

Park, J., et al. (2009). "Seismic fragility analysis of low-rise unreinforced masonry structures." Engineering Structures **31**(1): 125-137.

Paul, S. and D. Bhatia (2020). Smart Healthcare for Disease Diagnosis and Prevention, Academic Press.

Pejovic, J. and S. Jankovic (2016). "Seismic fragility assessment for reinforced concrete high-rise buildings in Southern Euro-Mediterranean zone." Bulletin of Earthquake Engineering **14**(1): 185-212.

Prieto, J. A., et al. (2018). "Development of structural debris flow fragility curves (debris flow buildings resistance) using momentum flux rate as a hazard parameter." Engineering Geology **239**: 144-157.

Profillidis, V. and G. Botzoris (2019). "Chapter 5—statistical methods for transport demand modeling." Modeling of Transport Demand; Profillidis, VA, Botzoris, GN, Eds.

Richman, J. S. (2011). Multivariate neighborhood sample entropy: a method for data reduction and prediction of complex data. Methods in enzymology, Elsevier. **487:** 397-408.

Rota, M., et al. (2010). "A methodology for deriving analytical fragility curves for masonry buildings based on stochastic nonlinear analyses." Engineering Structures **32**(5): 1312-1323.

Sadiq, R., et al. (2019). "Empirical models to predict disinfection by-products (DBPs) in drinking water: an updated review."

Salehi, H. and R. Burgueño (2018). "Emerging artificial intelligence methods in structural engineering." Engineering Structures **171**: 170-189.

Sandoli, A., et al. (2021). "Fragility curves for Italian URM buildings based on a hybrid method." Bulletin of Earthquake Engineering **19**(12): 4979-5013.

Sanft, R. and A. Walter (2020). Exploring mathematical modeling in biology through case studies and experimental activities, Academic Press.

Saouma, V. E. and M. A. Hariri-Ardebili (2021). Performance Based Earthquake Engineering. Aging, Shaking, and Cracking of Infrastructures, Springer**:** 517-528.

Saruddin, S. N. A. and F. M. Nazri (2015). "Fragility curves for low-and mid-rise buildings in Malaysia." Procedia Engineering **125**: 873-878.

Shariati, M., et al. (2020). "A novel hybrid extreme learning machine–grey wolf optimizer (ELM-GWO) model to predict compressive strength of concrete with partial replacements for cement." Engineering with Computers: 1-23.

Shinozuka, M., et al. (2000). "Statistical analysis of fragility curves." Journal of engineering mechanics **126**(12): 1224-1231.

Shobha, G. and S. Rangaswamy (2018). Chapter 8-Machine Learning Handbook of Statistics, Elsevier.

Subasi, A., et al. (2020). Human activity recognition using machine learning methods in a smart healthcare environment. Innovation in health informatics, Elsevier**:** 123-144.

Sun, H., et al. (2021). "Machine learning applications for building structural design and performance assessment: state-of-the-art review." Journal of Building Engineering **33**: 101816.

Suzuki, A. and I. Iervolino (2021). "Seismic fragility of code-conforming Italian buildings based on SDoF approximation." Journal of Earthquake Engineering **25**(14): 2873-2907.

Wadoux, A. M.-C., et al. (2022). "An integrated approach for the evaluation of quantitative soil maps through Taylor and solar diagrams." Geoderma **405**: 115332.

Walczak, S. and N. Cerpa (1999). "Heuristic principles for the design of artificial neural networks." Information and software technology **41**(2): 107-117.

Williams, B., et al. (2020). Data-Driven Model Development for Cardiomyocyte Production Experimental Failure Prediction. Computer Aided Chemical Engineering, Elsevier. **48:** 1639-1644.

Wood, D. A. and J. Cai (2021). Sustainable Geoscience for Natural Gas SubSurface Systems, Elsevier.

Zucconi, M., et al. (2018). "Simplified survey form of unreinforced masonry buildings calibrated on data from the 2009 L'Aquila earthquake." Bulletin of Earthquake Engineering **16**(7): 2877-2911.

# 6 Summary and conclusions

This thesis was organized in four main sections. Summary and conclusions drawn in each section are presented briefly in this section.

- **Chapter 1**: *Research background, motivation, objectives and organization*
  In the first chapter, research background is presented briefly by providing definitions and applications of the study subject matter. The needs for conducting this study, is stated and finally the main aims of the study are listed and discussed.
- **Chapter 2**: *a state-of-the-art review on time-dependent seismic hazard analysis*
  The main aim of reviewing previous studies was to evaluate the recent researches conducted on the subject matter and use their remarkable conclusions. More specifically, the time-dependent hazard models developed to date were gathered and the most accurate one (BPT) was selected for our study. Moreover, the adjustable parameters (i.e., site-to-source distance, structural response dispersion and aperiodicity parameter in BPT) were assessed carefully in relevant studies.
  As the main conclusion, the research gap in the study area was recognized and the thesis was organized so that it could cover the existing gap and needs. From the literature review, it was concluded that time-dependent seismic hazard has not been implemented in structural design so far. As a results, this gap formed the main objective of this research and the results could be considered for further revisions of design codes which currently use time-independent seismic hazard in their provisions (e.g., Eurocode 8).
- **Chapter 3**: *Influence of time-dependent seismic hazard on structural design*
  The second section aimed at evaluating the effect of time-dependent seismic hazard on structural capacity. Otherwise noted, the application of time-dependent seismic hazard in structural capacity, which is defined as the capacity required to provide a fixed

reliability level measured by the failure rate, was investigated for the first time. To this end, the analysis conducted on two main seismic sources: (i) a point source where the results of time-dependent seismic hazard was compared with the results of time-dependent model, and (ii) a more realistic source combined of a time-dependent area source (generating low- to medium-magnitude earthquakes) and aa time-dependent line source (generating high-magnitude earthquakes). The results of the combined source were compared with the results of the area source which was considered as the reference source. It is worth explaining that in the second source, only the line source (causing high-magnitude earthquakes) was considered time-dependent because it has been proven that sources with the capability of producing great magnitude follow time-dependent hazard models, while the sources with low- to medium- magnitude motions could be efficiently simulated by time-independent models (e.g., POI). Bothe the PSHA and fragility curves were implemented in our methodology to assess the influence of parameters namely site-to-source distance (source characteristics), dispersion of the structure response (structures' property) and aperiodicity factor in BPT model (time-dependent model) on seismic hazard and structural capacity.

The results highlighted the significance of time-dependent seismic hazard on structural capacity. The main conclusions are: (i) In the both point and combined source, high variations in hazard rates do not translate into similar variations in the capacity (in the combined source variations are smaller), (ii) The time-dependent behavior can be neglected in the design when the time-dependent source is sufficiently far (30 km in the case study), (iii) The dispersion of the structural capacity plays an important role in all the considered cases, (iv) variations of α plays an opposite role on short-term and long term required capacity: reducing (increasing) α will lead to reduction (increment) of required capacity for small elapsed times, associated to a hazard rate lower than the constant Poisson value, (v) Variations due to α changes in the parametric analysis were generally limited (lower than 10%).

- **Chapter 4**: *Influence of GMPE selection, soil type and period on structural capacity in the case of time-dependent seismic hazard*

  The effect of time-dependent hazard was deeper evaluated in the third section by taking soil classification, period and GMPE into consideration. To achieve the objective, the GMPE proposed by Ambraseys et al. was used in PSHA. The main asset of this GMPE is that it considers soil conditions and period. Accordingly, demand and structural

capacity for different soil classifications (soft or very soft, stiff soil and rock) and different periods (in the range of 0.1-2 s) were obtained and compared in different times elapsed since the last event (earthquake) using the methodology introduced in the second section. More importantly, the effect of the GMPE used in the analysis was assessed by comparing the obtained results of Ambraseys et al. with those obtained by implementing Sabetta and Puliese GMPE.

The notable conclusions according to the results are: (i) The hazard due to combined source was affected mostly by area and line source in short and long periods, respectively for all soil types, (ii) The difference between demand and capacity for a site located on soil is larger than a site located on rock, (iii) Regardless of the soil type and GMPE, by elapsing time, both of demand and capacity increase. However, the capacity changes is slightly larger than demand changes, (iv) For very short periods (less than 0.2s) or very long periods, elapsing time does not affect neither capacity nor demand notably, (v) Elapsing time affect both of demand and capacity corresponding to moderate periods (0.4s in this study) significantly, and (vi) The GMPE used for the analysis affects the results remarkably, therefore the most appropriate GMPE for the region of the study should be used

- **Chapter 5**: *Fragility curves of buildings, a critical review and ML-based study*

  In this section, an attempt was made to develop ML-based models for generating fragility curves of buildings. To this end a comprehensive database including 236 analytical results published in peer-reviewed journals were collected. The database was then divided into two sub-databases: training and testing which were respectively used for training the relationship between inputs and outputs, and assessing the accuracy of the models. Nine input variables including plan area, height, period, construction material, lateral resisting system, soil classification of the building location, design code and damage state were defined to predict dispersion and median values which are known as fragility parameters. The prediction models were developed using techniques namely nonlinear regression, Decision Tree, Random Forest, K-nearest Neighbours and Artificial Neural Networks. The predicted values are compared with actual values and the accuracy of the models was assessed through performance metrics and Taylor diagram. The most accurate model was introduced and used for conducting a sensitivity analysis on the quantitative inputs. The results illustrated the high ability of ML-based methods for learning the relationship between the inputs and outputs properly.

Moreover, the models could be introduced as an acceptable quick alternative method for obtaining fragility curves instead of time-consuming and complicated conventional methods. The parametric study clarified the effect of each quantitative input variable on fragility of buildings. The proposed models could be used in further studies for predicting structures' capacity during different times elapsed since the last earthquake. The main conclusions are: (i) All the proposed models could learn the relationship between the input variables and the output properly, (ii) The high accuracy of the predicted values proved the reliability of the ML-based prediction models, (iii) For predicting Beta: the Decision Tree model had the highest accuracy with $R^2$=0.91, (iv) For predicting Median, the Decision Tree model exhibited the highest accuracy with $R^2$=0.99, and (v) Therefore, Decision Tree method could be applied for estimating fragility curves of steel, RC and masonry buildings.