ATTRIBUTE-DRIVEN FRAGILITY CURVES THROUGH CLASS DISAGGREGATION

A.K. Kazantzi[(1)], D. Vamvatsikos[(2)]

[(1)] *Research Engineer, National Technical University of Athens, kazantzi@mail.ntua.gr*
[(2)] *Assistant Professor, National Technical University of Athens, divamva@mail.ntua.gr*

Abstract

Fragility curves are an important ingredient in the seismic loss assessment process. For a regional scale loss estimation, to reduce to reasonable levels the computational burden associated with determining the seismic demands for individual buildings, analytical seismic fragilities are instead evaluated on a broad building class basis. The latter process essentially involves representing a population of buildings having similar characteristics with a set of characteristic “index” buildings to avoid analyzing every single building within this population. For the definition and modeling of index buildings, two main options are currently available, these being (a) defining a limited number of index buildings to represent the class and modeling them with relatively complex, yet more accurate, MDOF systems, and (b) defining numerous index buildings to represent the class and modeling them with simplified approximate SDOF systems. Apparently, the dilemma of defining the optimal way to sample the index buildings comes down to the use of few MDOFs or many SDOFs.

Despite the fact that the use of many SDOFs is a rather attractive option, given that they are an easy and computationally inexpensive choice in terms of both modeling and analysis, they are often a bad approximation of the actual problem. This is the case, for example, of tall or irregular buildings, where non-negligible higher modes render the SDOF approximation ineffective. Then, the more expensive and accurate MDOF option has to be employed. However, using a limited number of MDOFs to represent the class of interest inherently offers very little flexibility towards capturing individual buildings that might belong to that class yet their salient features do not necessarily match those of the “average” index building. Aggregating the results of all index buildings into a single class fragility means that one cannot provide a more accurate answer than the mean class fragility plus some dispersion, even if the building in question actually closely matches one of the underlying index ones. This may not matter for estimating long-term average losses over a region, but it becomes increasingly important as the size of the portfolio is reduced and individual structures stand out.

To resolve the aforementioned issue, we propose here a method for adding substance back to the class fragility and consequently obtaining fine-grained attribute-driven fragility estimates. The term attribute-driven is key in our approach, since it implies that the process explicitly accounts for the specific characteristics of the building in question. It is essentially a meso-scale approach that stands between the building-specific FEMA P-58 style approach (micro-scale) and the building-class approach (macro-scale). Our testbed is a population of modern high-rise reinforced concrete buildings, represented by seven index buildings, for which we have evaluated fragility functions. With this information at hand, our proposed approach employs statistical methodologies for effectively disaggregating the index building fragility functions, to provide attribute-aware response and collapse fragility spot estimates for individual sample buildings, other than the index ones, that belong to the same class.

Keywords: fragility; building class; loss assessment

1. Introduction

At present, vulnerability assessment methodologies exist both for individual buildings and building classes. These two approaches may be interchangeably adopted by the interested beneficiaries (engineers, insurance companies, emergency management agencies, building owners etc.) based on their need to perform a vulnerability, and consequently loss, assessment, either on a single building or an entire building portfolio. Building-specific analytical vulnerability methodologies are deemed to be well defined, nowadays with the FEMA P-58 [1] approach being currently considered as the state-of-art. By contrast, for the class vulnerability assessment, there are still several issues than need to be resolved, some of which were only relatively recently addressed by a grand research project undertaken by the Global Earthquake Model (GEM) for proffering a comprehensive open-source tool for large scale loss assessment studies.

The idea of adopting the FEMA P-58 [1] methodology for undertaking a regional loss assessment is hampered by the excessive required time and computational demands. For the aforementioned reason, alternative analytical methodologies have been proposed in the past few years for dealing with the problem of assessing building vulnerability on a large scale. These methodologies, no matter whether are empirical, analytical or hybrid may be characterized as being attribute-blind, in the sense that they do not explicitly account for the individual characteristics of every single building allocated in a particular class. By contrast, the entire building stock, is first categorized into a number of broad building classes, with each one of them representing a different group within the building population, having similar key characteristics, in the sense of these affecting substantially the seismic performance (e.g. material type, height range, use, construction era etc.). More importantly, for reasons related to economizing in terms of time and computational resources, building classes are often defined in a rather generic manner to allow modeling the representative structures with a limited number of relatively complex MDOF index buildings. The latter results to large discrepancies in the vulnerability estimates, even between the index buildings selected to represent a particular class. For instance, Kazantzi *et al*. [2] have shown for a building class of modern high-rise reinforced concrete buildings, that the vulnerability variability could be rather high even between the selected index buildings, with the most vulnerable index building being three to four times more vulnerable compared to the least. On the other hand, if numerous simpler SDOF models are utilized to represent the class, then the in-class variability could be more effectively depicted, but such simplified models often represent a bad approximation of the actual problem, in particular in higher-mode dominated structures (e.g. high-rise or irregular buildings).

At the moment, when considering building-specific vulnerability assessment, one could either adopt the time-consuming FEMA P-58 [1] micro-scale approach, which requires not only a great deal of time but also a high-level of engineering expertise, or the building class macro-scale approach (with few, yet more realistic, index buildings modelled as MDOFs or several, yet more approximate, index buildings modelled as SDOFs to represent the class). In the case of few MDOFs we essentially assume that the mean class fragility satisfactorily depicts the fragility of every building belonging to the considered class whereas in the case of many SDOFs we assume that the one nearest to the considered building SDOF represents a realistic approximation of the structural system in question, no matter its structural complexity. The proposed attribute-driven approach stands between these two approaches, aiming to disaggregate effectively the macro-scale class vulnerability functions to end up with approximate, yet almost instantly obtained and satisfactorily accurate vulnerability estimates. In particular, the proposed approach utilizes pre-determined class vulnerability functions and with the aid of statistical methodologies it adds substance back to the class fragility to obtain fine-grained fragility estimates at low computational cost for individual sample buildings, other than the index ones.

2. High-rise building class

The present study considers a class of high-rise Reinforced-Concrete Moment-Resisting Frames (RCMRFs) that is described in detail by Kazantzi & Vamvatsikos [3]. The class refers to modern structures built to post-1980 seismic design provisions for high-seismicity regions in the United States. The main features differentiating the buildings within the class are (a) the building height, defined as the number of stories, (b) the design base shear, as this was determined by the code-based value of spectral acceleration at 1s, termed

SD1 in the US codes and (c) the vertical irregularity ratio, defined as the ratio of the first story height to the (constant) height of the upper stories. The high-rise index set consists of seven RCMRFs with heights ranging from 7 to 20 stories. The seven index buildings were obtained on the basis of three perimeter frames that were selected and accordingly modified by a database of thirty (30) archetypes. This database was developed by Haselton *et al*. [4]. The probability distributions of the building features were based on a set of 263 RCMRFs in California [5]. The main features of the analyzed index frames are summarized in Table 1.

The weights that represent the contribution of each index building to the total sample along with the discrete joint probability mass functions of the high-rise class key features (defined by means of a common central point and two "sigma set" points per attribute to establish higher moments) were obtained using the moment-matching sampling technique [6].

Table 1 – Main features of the index buildings and their moment-matching weights

Index	Feature $X1$ (No of stories)	Feature $X2$ (code design level)	Feature $X3$ (vertical irregularity)	T_1 (sec)	T_2 (sec)	moment-matching weights
No0	12	0.60g	1.744	2.14	0.73	0.1475
No1	7	0.60g	1.744	1.61	0.52	0.0962
No2	20	0.60g	1.744	2.85	0.92	0.0787
No3	12	0.60g	1.150	2.02	0.69	0.3152
No4	12	0.60g	2.745	2.42	0.81	0.0425
No5	12	0.26g	1.744	2.14	0.73	0.1700
No6	12	0.97g	1.744	2.14	0.73	0.1500

3. Modeling of the class index buildings

With respect to the structural modeling it can be said that any detailing level that is able to capture the dominant failure modes of the index buildings is suitable for application within a vulnerability assessment study. The aforementioned statement however, disregards potential time and computational resources limitations which are present in studies dealing with a class of buildings rather than a single one. In the methodology proposed by GEM, three levels of complexity are proposed, offering three distinct choices of structural detailing, these being [7]:

(a) Level A (MDOF): A detailed 2D or 3D multi-degree-of-freedom (MDOF) model of a structure, including elements for each identified lateral-load resisting component in the building, e.g. columns, beams, walls etc.

(b) Level B (Stick): A simplified 2D lumped representation of the building where each one of the N floors is represented by one node having 3 to 6 degrees of freedom.

(c) Level C (SDOF): A simple single-degree-of-freedom (SDOF) representation of the building by a 1D nonlinear spring.

In this study, the simplest, yet efficient, representation of the index buildings was considered to be a 2D model idealization of the MDOF structure (i.e. Level A). The 2D modeling is justified by the insignificant plan asymmetry encountered by the index buildings. A Level B model is not recommended for high-rise buildings (e.g. buildings with their height more than three times their width), since irregular shortening or lengthening of the columns may well violate the fundamental planar floor assumption and induce significant secondary moments [7]. Level C is not suitable for high-rise buildings for the apparent reason that their seismic performance in not anchored to a single mode of vibration.

Regarding the structural members, their behavior was depicted by lumped plasticity elements having their properties evaluated from empirical equations proposed by Panagiotakos & Fardis [8]. Adopting lumped plasticity elements, as opposed to the more sophisticated but less well documented fiber section models (especially when close to collapse), was considered to be a reasonable compromise for a vulnerability assessment. To properly account for the destabilizing P-Δ effects, the perimeter frames were assumed to carry only a fraction of the tributary gravity loads that were considered for evaluating the seismic floor masses whereas the remaining portion was applied on a leaning column. This is anticipated to be a reasonable assumption that also accounts for the contribution of the internal gravity columns on the lateral load resisting mechanism of the building. Fig. 1 depicts the model idealization of the No1 7-story perimeter frame along with its corresponding P-Δ leaning column.

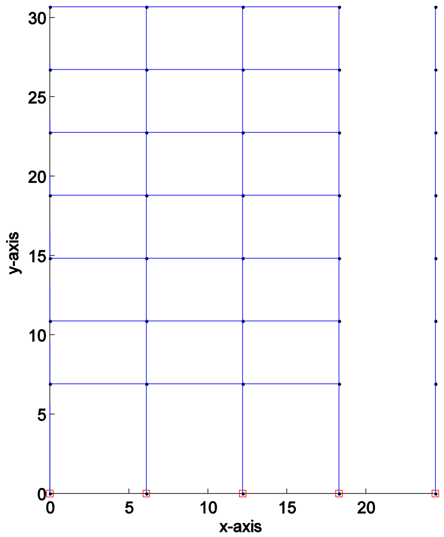

Fig. 1 – Model idealization of the No1 7-story perimeter RC index frame

4. Evaluation of structural response

For evaluating the seismic performance of the high-rise index buildings, the Incremental Dynamic Analysis (IDA) [9] is proposed as the benchmark analysis methodology. The ground motion records that are needed for the IDAs come from the far-field record set of FEMA P695 [10] which contains 22 ground motion records with two components each (i.e. 44 individual components).

The Intensity Measure (IM) selection is a rather important task towards the development of analytical seismic vulnerability functions, irrespectively of whether we consider a single building or a class of buildings. The key difference between the aforementioned cases, is that the consideration of a building class, imposes to the problem an additional requirement for adopting a common, yet practical, IM that maintains efficiency and sufficiency for all index structures within the class. One such common IM, was shown by Kazantzi & Vamvatsikos [3] to be a geometric mean scalar IM which uses five periods, $S_{agm}(T_i)$, ranging from the second-mode (T_2) to twice the first-mode period ($2\cdot T_1$) that may be computed as follows,

$$S_{agm}(T_i)=\sqrt[5]{S_a(T_2)\cdot S_a(\min[(T_2+T_1)/2, 1.5\cdot T_2])\cdot S_a(T_1)\cdot S_a(1.5\cdot T_1)\cdot S_a(2\cdot T_1)} \quad (1)$$

5. Index building collapse fragility curves

The fragility is defined here as the probability function of the limit-state capacity C being exceeded by the demand D for a given intensity level (i.e. IM-value), s. If demand and capacity are expressed in terms of intensity levels, then we get the simplest representation of fragility:

$$P_{LS}(s)=P(C<D\,|\,s)=P(s_c<s\,|\,s)=F(s_c\,|\,s) \quad (2)$$

where, s_c is the (random) IM-value of capacity that when exceeded signals violation of the limit-state and F[·] is the cumulative distribution function (CDF) of its arguments. Essentially, the fragility curve then becomes the CDF of s_c evaluated at the intensity level s. By assuming that the probability distribution of s_c is lognormal, a reasonable assumption according to existing literature, we can also get the following expression:

$$P_{LS}(s) = \Phi\left(\frac{\ln s - \ln \hat{s}_c}{\beta_{s_c}}\right) \tag{3}$$

where, $\hat{s}_c$ is the median IM-value of capacity, β_{s_c} the corresponding dispersion (i.e. standard deviation of the log data) and Φ represents the CDF of the standard normal distribution, readily available, e.g. in Excel.

Fig. 2 illustrates the collapse fragility curves for all index buildings and for the selected scalar geometric mean IM. For deriving the fragility curves by means of Eq. (3) the collapse capacities were assumed to be lognormally distributed. Table 2 summarizes the statistics (median and dispersion) of the collapse capacity evaluated for each one of the index buildings as well as the total expectation and total variance for the class fragility evaluated by combining the sample means and dispersions via the laws of total expectation [11] and total variance [12], respectively.

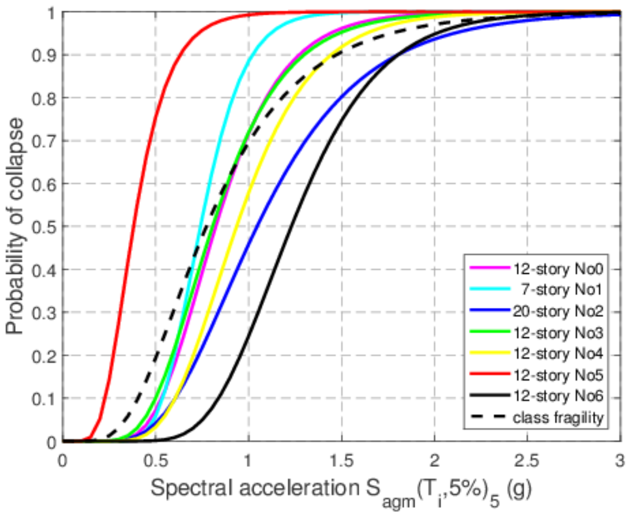

Fig. 2 – Collapse fragility curves for the seven high-rise RC index buildings along with the class fragility curve

Table 2 – Median and dispersion as well as the class statistics of the collapse capacity for $S_{agm}(T_i)_5$

ID	$\hat{s}_c$ [g]	β_{s_c}
No0	0.82	0.34
No1	0.74	0.25
No2	1.05	0.43
No3	0.81	0.37
No4	0.94	0.34
No5	0.38	0.39
No6	1.23	0.30
class total	0.77	0.50

6. k-Nearest Neighbor (k-NN) classification

For determining the collapse fragility of a random sample building that belongs to the considered high-rise class, we shall interpolate the fragilities estimated from the reference index buildings based on the salient characteristics X1 – X3 that distinguish them. Non-parameteric interpolation/regression approaches are the preferred approach, and there are several to choose from: Neural Networks, Decision Trees, and all pertinent machine-learning approaches. Perhaps the simplest effective method is the k-Nearest Neighbor (k-NN) regression. It is based on estimating values (here the collapse fragility median and dispersion) for a test sample (i.e., any given highrise building) by weighted averaging of the k closest members of the pre-computed training samples (i.e., the index buildings). k-NN is considered to be an instance-based or lazy learning algorithm [13], in the sense that the training data points are not being used to fit any generalized regression expression (e.g., as in standard linear regression), but are instead they are employed directly to compute the regression/interpolation result. Ideally k should be chosen so as to represent a small fraction of the training samples, since in principle we want the k nearest samples to be as close as possible to the sample case. Apparently, as the number of training samples increases the performance of the k-nearest rule increases [14].

To determine closeness, one may select from several distance metrics that are available in the literature, e.g. Euclidean, Standardized Euclidean, Minkowski etc. For example, the Euclidean distance between two *N*-dimensional vectors *x*, *y* is evaluated as:

$$d_{x,y} = \sqrt{\sum_{j=1}^{N} (x_j - y_j)^2} \tag{4}$$

For the case at hand, the Standardized Euclidean distance is an ideal choice. In fact, the standardized Euclidean distance is evaluated in an identical to the Euclidean distance manner, but prior to estimation each variable x is normalized to have a mean of zero and a variance of one, simply by subtracting the mean and dividing by the standard deviation. This transformation results in the output being unaffected by the different scaling between the class feature attributes (i.e. height, spectral acceleration and vertical irregularity ratio in our case).

7. Case-study

The testbed for demonstrating the methodology, will be a post-1980 designed RC building with 8 stories, having a vertical irregularity ratio of 1.150 and designed for an $S_a(T_1)$ of 0.4g. The fundamental period of the sample building was found to be 1.70sec while the second mode period was estimated to be 0.55sec. The considered building, apparently, could be considered to belong to the reference building class, but according to its salient characteristics it is difficult to guess which of the seven class index buildings is the closest to it in terms of vulnerability. However, if we have to pick three such candidate index building, we can say that its vulnerability will more likely lie somewhere between index building No1 as it has 7 stories, index building No3 that has the same vertical irregularity ratio, and index building No5 that was designed for an S_a of 0.26g (see Table 1). Further to the above, its fragility curve should lie somewhere between the red and the green fragility curves with reference to Fig. 2.

The k-NN method provides not only a classification of the sample building vis-a-vis its nearest neighbor index building, but also a ranking based on how close that sample is to its k-nearest neighbors. The ranking assumes a weight that is inversely proportional to the evaluated standardized Euclidean distance. Hence, the larger the distance of the sample building from each index building, the smaller the contribution of the vulnerability of this index building to the vulnerability of the sample. Finally, all weights are normalized so as their sum adds up to 100%.

Table 3 provides the weight percentages associated with the case study sample building, using a 3-NN classification, i.e., inversely proportional to the distance of the sample building from the *k*=3 closest index buildings. As can be inferred from the tabulated data the 3-NN classification selects as closest indices the No1, No3 and No5 buildings, with No3 considered to be most influential. Considering the weights and the median collapse capacities of the three index buildings, the 3-NN weighted average estimate of the median collapse capacity for the test 8-story building is equal to

exp(ln(0.74g)*32.90%+ln(0.81g)*37.16%+ln(0.38g)*29.94%) = 0.63g which is lower than the total expectation of the class collapse capacity by a factor of ~1.22 (0.77g/0.63g). With reference to the weighted dispersion of the test sample building, this is equal to

$$\sqrt{(32.90\%*0.25^2+37.16\%*0.37^2+29.94\%*0.39^2)} = 0.34$$

which is substantially lower from the class total variance that was found to be 0.50 (see Table 2).

To assess the accuracy of the presented methodology the fragility of the 8-storey sample building was evaluated explicitly via IDA, in the same manner done for the index buildings. Fig. 3 presents the fragility of the sample building along with the fragility curves of the 7 index buildings that were deemed to represent the class of interest. The sample building was found to have an actual median collapse capacity of 0.46g and a dispersion of 0.27.

Table 3 – Ranking percentage (weights) of the sample 8-story building using a 3-NN regression along with the median collapse capacities of the index buildings in S_{agm} terms

Index building	**S_{agm} [g]**	**Ranking percentage – 3NN [%]**
No0	0.82	—
No1	0.74	32.90
No2	1.05	—
No3	0.81	37.16
No4	0.94	—
No5	0.38	29.94
No6	1.23	—

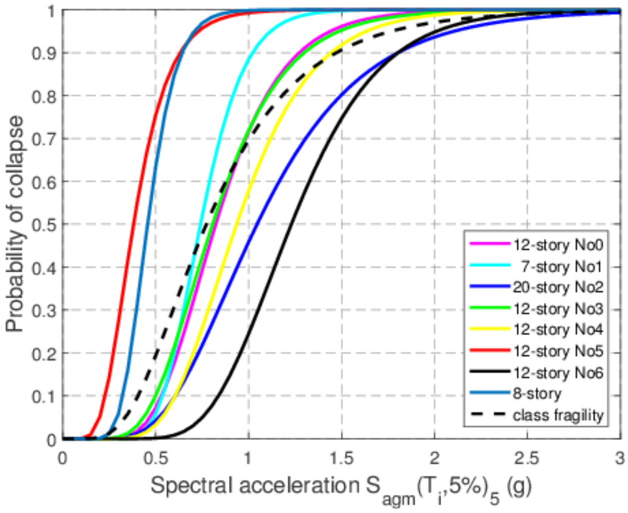

Fig. 3 – Collapse fragility curves for the seven high-rise RC index buildings along with the class fragility curve as well as the fragility curve that was explicitly estimated for an 8-story RC building that belongs to the class, yet its characteristic features are not identical to those of the index buildings

The aforementioned findings imply that, if the class fragility was used for assessing the median collapse risk of the sample building, this would have resulted into a collapse capacity overestimation by 67% (0.77g as opposed to 0.46g which is the actual median collapse capacity), consequently leading to lower loss estimates as well as lower insurance premiums not reflecting realistically the actual loss distribution. By contrast, if the k-NN method is used, we would have predicted a better median collapse capacity for the building in question, that would have been about 37% higher than the actual one (0.63g as opposed to 0.46g), still leading to unconservative, yet highly improved estimate compared to those obtained using the median collapse capacity of the class. The benefits are even more substantial for the dispersion estimate, where the 3-NN approach is clearly superior to using the class fragility, by providing a much more reasonable dispersion estimate of 0.34 vis-à-vis the much higher class-level value of 0.50.

To further improve the accuracy of the proposed methodology, one could consider assigning a weight to the characteristic features of the index buildings. Identifying the most influential features could be a straightforward task but defining their weights to be considered in the k-NN process is likely a more challenging one. For the case at hand, the least influential feature seems to be the Feature X3 that refers to the vertical irregularity ratio. This is apparent even by simply inspecting the collapse fragilities of the index buildings No0, No3 and No4. Those index buildings differ only in Feature 3 and apparently their collapse fragilities are very close, with those of index buildings No0 and No3 (see Figure 3, pink and green curves) being almost identical. That implies that for the considered population of buildings, which is designed to modern seismic design provisions, typical vertical irregularity ratios have little effect on the collapse capacity.

As an illustrative example, we have reapplied the k-NN methodology disregarding Feature X3. In addition, to avoid biasing the estimate by the lack of a sufficient number of index buildings, we sought just the two nearest samples in a $k = 2$, 2-NN regression scheme. The methodology resulted in assigning the sample building closest to the index buildings No1 and No5 with ranking percentages (weights) of 55.30% and 44.7%, respectively. This yields a weighted median collapse capacity equal to 0.55g which is only about 20% higher than the actual one (i.e. 0.46g) and a dispersion equal to 0.32, that is only slightly higher than the actual one (i.e. 0.27). Figure 4 illustrates the actual collapse fragility of the 8-story building along with the class fragility and the fragilities obtained following a 3-NN and a 2-NN approach. The differences between the collapse fragilities (actual and disaggregated) of the sample building with the class fragility are apparent. For instance, by simply inspecting Figure 4, for a S_{agm} equal to 0.8g the actual collapse probability of the 8-storey building is approximately 98%. For the same intensity level, the class fragility yields a substantially lower collapse probability of 53%, whereas the 3-NN and the 2-NN approaches yield collapse probabilities of 76% and 88%, respectively. The improvement in accuracy becomes more apparent if we consider that it comes at the cost of only a few additions, multiplications and divisions.

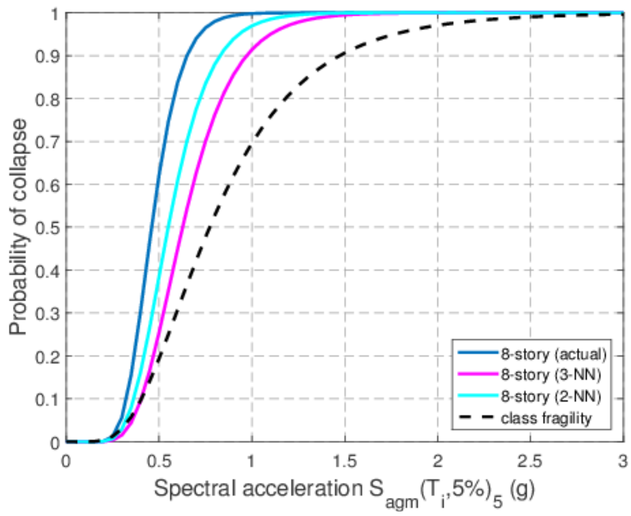

Fig. 4 – Comparison of the collapse fragilities estimated for the 8-story RC building via the k-NN approach and the class fragility curve.

8. Conclusions

The study proposes a methodology for evaluating attribute-driven fragility curves for individual buildings using a simple, yet efficient, k-NN based class-disaggregation technique. It was showcased that the proposed methodology yields results of superior accuracy compared to the use of the class fragility (macro-scale) and of acceptable accuracy as well as in a miniscule fraction of the time required to explicitly evaluate the fragility by means of the building-specific FEMA P-58 style approach (micro-scale). The methodology can be applied in cases where the individual constituent index buildings used to form class fragilities are available to obtain almost instantaneous fragility estimates to be used in loss assessment, insurance premium estimates or even property evaluations.

9. Acknowledgements

Financial support has been provided by the Executive Agency for Small and Medium-sized Enterprises (EASME) under the powers delegated by the European Commission through the Horizon 2020 program "HYPERION- Development of a decision support system for improved resilience & sustainable reconstruction of historic areas to cope with climate change & extreme events based on novel sensors and modelling tools", Grant Agreement number 821054.

10. References

[1] FEMA (2012): Seismic performance assessment of buildings, FEMA P-58-1, Applied Technology Council for the Federal Emergency Management Agency, Washington, DC.

[2] Kazantzi A, Vamvatsikos D, Porter K, Cho I (2014): Analytical vulnerability assessment of modern highrise RC moment-resisting frame buildings in the Western USA for the Global Earthquake Model. *2nd European Conference on Earthquake Engineering*, Istanbul, Turkey.

[3] Kazantzi A, Vamvatsikos D (2015): Intensity measure selection for vulnerability studies of building classes. *Earthquake Engineering & Structural Dynamics*, **44** (15), 2677-2694.

[4] Haselton CB, Liel AB, Deierlein GG, Dean BS, Chou JH (2011): Seismic collapse safety of reinforced concrete buildings. I: Assessment of ductile moment frames. J*ournal of Structural Engineering*, **137** (4), 481–491.

[5] Porter K, Cho I (2013): Characterizing a building class via key features and index buildings for class-level vulnerability functions. *Proceedings of the 11th International Conference on Structural Safety & Reliability (ICOSSAR)*, Columbia University, New York, NY.

[6] Ching JY, Porter KA, Beck JL (2009): Propagating uncertainties for loss estimation in performance-based earthquake engineering using moment matching. *Structure and Infrastructure Engineering*, **5** (3), 245–262.

[7] D'Ayala D, Meslem A, Vamvatsikos D, Porter K, Rossetto T, Crowley H, Silva V (2013): Guidelines for analytical vulnerability assessment. Global Vulnerability Consortium, www.nexus.globalquakemodel.org/gem-vulnerability/posts/.

[8] Panagiotakos TB, Fardis MN. (2001): Deformations of reinforced concrete members at yielding and ultimate. *ACI Structural Journal*, **98** (2), 135–148.

[9] Vamvatsikos D, Cornell CA (2002): Incremental dynamic analysis. *Earthquake Engineering & Structural Dynamics*, **31** (3), 491-514.

[10] FEMA (2009): Quantification of building seismic performance factors, FEMA P695. Applied Technology Council for the Federal Emergency Management Agency, Washington, DC.

[11] Wikipedia (2019): Law of total expectation. https://en.wikipedia.org/wiki/Law_of_total_expectation. Last modified October 14, 2019.

[12] Wikipedia (2020): Law of total variance. https://en.wikipedia.org/wiki/Law_of_total_variance. Last modified January 8, 2020.

[13] Wikipedia (2020): k-nearest neighbors algorithm. https://en.wikipedia.org/wiki/K-nearest_neighbors_algorithm. Last modified January 30, 2020.

[14] Duda RO, Hart PE, Stork DG (2000): Pattern classification. John Wiley & Sons.

