ORIGINAL ARTICLE

lajss.org© Latin American Journal of Solids and Structures

www.lajss.org

Epistemic Uncertainty Treatment Using Group Method of Data Handling Algorithm in Seismic Collapse Fragility

Fooad Karimi Ghaleh jough[a]*, Meisam Veghar[b], S.Bahram Beheshti-Aval[c]

[a] Final International University, Kyrenia, North Cyprus, via Mersin 10 Turkey. E-mail: Fooad.karimi@final.edu.tr
[b] Faculty of Civil Eng, Semnan University, Semnan, Iran. E-mail: Meisam.veghar@semnan.ac.ir
[c] K.N. Toosi University of Technology, Faculty of Civil Eng, Tehran, Iran. E-mail: beheshti@kntu.ac.ir

* Corresponding author

https://doi.org/10.1590/1679-78256254

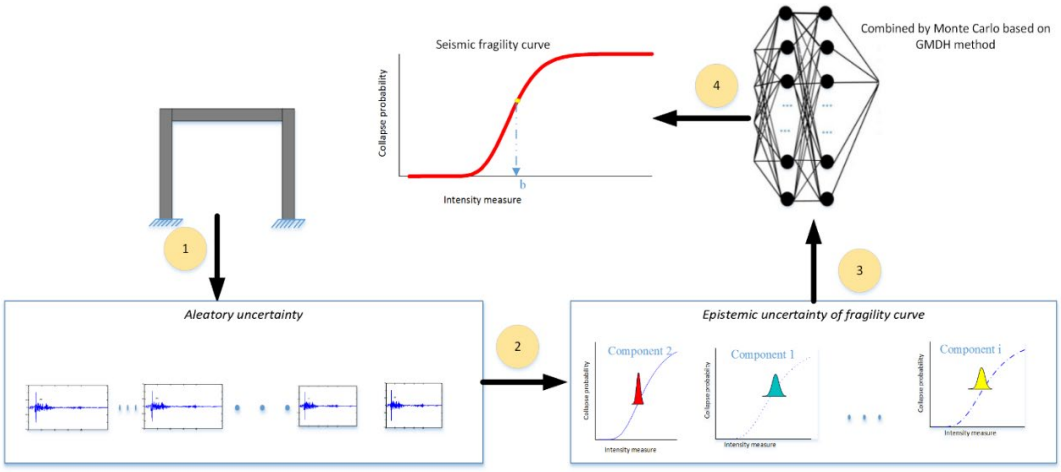
Abstract

Developing fragility functions is the essential step in incorporating important uncertainties in next-generation performance-based earthquake engineering (PBEE) methodology. The present paper is aimed to involve record-to-record variability as well as modelling uncertainty sources in developing the fragility curves at the collapse limit state. In this article, in order to reduce the dispersion of uncertainties, Group Method of Data Handling (GMDH) in combination with Monte Carlo Simulation (MCS) is used to develop structural collapse fragility curve, considering effects of epistemic and aleatory uncertainties. A Steel Moment Resisting Frame (SMRF) is chosen as the tested structure. The fragility curves obtained by the proposed method which belongs to GMDH approaches are compared with those resulted from simple and well-known available methods such as First-Order Second-Moment (FOSM), Approximate Second-Order Second-Moment (ASOSM) and Monte Carlo (MC)/Response Surface Method (RSM), somehow, as an accurate method. The results of the application of the proposed approach indicate increasing accuracy and precision of the outputs as well as power with the same computational time compared to aforementioned methods. The GMDH method introduced here can be applied to the other performance levels.

Keywords

Group Method of Data Handling (GMDH) algorithm; Monte Carlo method; Aleatory uncertainty; Epistemic uncertainty.

Graphical Abstract



Received September 09, 2020. In revised form December 12, 2020. Accepted February 01, 2021. Available online February 08, 2021.
https://doi.org/10.1590/1679-78256254

Latin American Journal of Solids and Structures. ISSN 1679-7825. Copyright © 2021. This is an Open Access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

1 INTRODUCTION

Safety analysis of structures in seismic conditions is one of the significant phases of structural design especially in urban areas with high hazard of earthquake. Over the years with collaboration among earthquake engineering research centers, the Performance-Based Earthquake Engineering (PBEE) methodology has received more achievements to employ more reliable approaches in determining the seismic performance of structures. Considering the limits of the first generation of performance-based guidelines and the necessity of explicit inclusion of existing uncertainties, the Pacific Earthquake Engineering Research (PEER) Center has provided a new framework for PBEE (Ghobarah, 2001). The proposed procedure includes the inclusion of data and models related to seismic hazard, seismic response of structures and earthquake damage and losses through triple integration of PEER framework. By combining these data, the ultimate result reveals the probable consequences of different earthquake scenarios, e.g., mean annual frequency of exceeding direct and indirect damage and fatality after an earthquake. The outcome of this integration as reliable criteria forms a common language among different interested groups with different attitudes. The first part of this triple integral is devoted to assess the expected mean annual exceedance frequency of Engineering Demand Parameter (EDP) illustrated in the following equation:

$$\boldsymbol{\lambda_{LS}} = \int \boldsymbol{G(LS|IM = im)}.|\boldsymbol{d\lambda(im)}| \tag{1}$$

In which, im is intensity measure, λ(im) is mean annual rate of exceedance for a given intensity (seismic hazard curve), G(.) is conditional cumulative probability distribution function for specified Limit State (LS) or in fact is constituting fragility function. As can be seen, the seismic fragility curves at different limit states are recognized as the crucial parts of the PBEE framework. The basic components of fragility curves are more tangible to the damage level at various performance level. Among various limit states, collapse safety is more important pertaining to increasing rapidly casualties and fatality after strong ground motions. This metric is going to be crucial safety assessment criteria for new and excising constructions. Recently, ASCE 7 (2016) would mandate taking the collapse probability of all new buildings at the intensity level of a maximum considered earthquake (MCE) less than 10%.

Collapse of buildings may be triggered through the lack of load carrying capacity of lateral and gravity load resisting systems. In this case L_{LS}= $L_{COLLAPSE}$ is mean annual frequency of collapse, which may be used as metric for safety assessment of structures against various probable intensity levels in site specific seismic hazard. Lesson learnt from past earthquake disaster indicated that due to the side-sway collapse, the majority of collapse was extended to vertical collapse accelerated by P-Delta effects for taller buildings. Side-sway collapse is defined as the total instability of building structural system due to lateral displacements of building under intensive ground motions and second-degree impacts of P-Δ effect of gravity loads, to the extent that emergent mechanisms cause the whole structural instability and collapse (Möller and Beer, 2013). Employing sequential dynamic analyses using a suit of proper seismic records so-called Incremental Dynamic Analysis (IDA) may be used for developing fragility curves based on two approaches, IM-based and EDP-based perspectives. Although in both cases, the collapse fragility curve is obtained by comparing seismic demand and the capacity of structure, difference is attributed to considering the variable in integration "Eq. (1)" where in the first one is IM and in the later is EDP. The collapse limit state was defined when the applied demand to structure exceeds its collapse capacity. In this study for simplicity and direct usage of last multiple analysis outcomes, IM-based approach is used based on direct usage of IM parameter ($S_{acollapse}$) in determining the collapse fragility curve [Ibarra and Krawinkler, 2005]. The collapse capacity in each IDA curve is the intensity measure of scaled record applied to structure before dynamic instability occurred in acceptable tolerances. The collapse limit state is assessed here through detecting sidesway collapse of the structure (Haselton, 2006). The collapse fragility curve which is collapse probability conditioned on the intensity measure is determined using the following IM-based method (Karimi Ghaleh Jough and Şensoy, 2016):

$$\boldsymbol{G(EDP|IM = im)} = \boldsymbol{p(EDP > edp|IM = im)} = \boldsymbol{P(im > IM_{collapse})} \tag{2}$$

The important aspects of the seismic performance assessment framework in next generation of PEER approach is the quantification and propagation of the various uncertainties throughout the seismic evaluation of structures. The fragility functions account for major uncertainty in PBEE methodology. In reliability assessment framework, the uncertainty sources are categorized into aleatory and epistemic uncertainties. Aleatory uncertainties are referred to ones that according to our current knowledge and data cannot be accurately predicted. Given the limited information pertaining to the mechanism of earthquake occurrence, nonce, the exact simulations of the characteristics of future ground motions are impossible. The aleatory uncertainty laid in predicting ground motions is categorized in two sources,

one from frequencies, amplitudes, phases, and durations is so-called 'record-to-record' variability and the others associated with variability in ground motion intensity were addressed by site-specific hazard curve. The first term is incorporated by sequential dynamic analyses such as IDA, and the second term is treated by computing mean annual occurring frequency through integration of the corresponding fragility curve over the mean hazard curve.

In contrast to aleatory ones, the epistemic uncertainties can be reduced by developing knowledge boarders. The effect of these uncertainty factors can be reduced by collecting more data or using more appropriate analytical model. The parameters of modeling assumptions (analytical model) are mainly sources of epistemic uncertainties, which are propagated into the structure responses through numerical analysis. The parameters of structural modeling, building construction quality and computational methods in predicting more accurate structural behavior of buildings can be incorporated by epistemic uncertainties (Ibarra et al., 2005). In addition to inherent uncertainties in structural idealization and analysis strategies, modeling uncertainties involved in uncertain variables may be divided into physical and non-physical ones. Ones which are raised from direct measure of variable involved in analysis such as loading (dead loads and live loads), mass, material properties, and they are considered as physical variables. Other modeling uncertainties maybe indirectly measured and estimated using predictive function of physical variables such as damping and the parameters for definition of generalized force-deformation relationship in analytical models. These are so-called non-physical parameters (Bradley, 2013).

Selecting which types of modeling uncertainties to be considered in analysis depends on the type of analytical model. If the material stress-strain relationship is used as primary structural behavior, the mechanical properties of members' materials are the source of uncertainties (such as elastic modulus, yield stress or strain, ultimate strength and strain and so on). Two following finite element models belong to this category:

(i) In microscopic finite-element models where members are modeled by assembly of numerous two or three-dimensional finite elements

(ii) In discrete finite-element modeling of skeletal structures and using fiber sections at integration points along frame members for evaluating moment-curvature relationship. This model is categorized into the distributed plasticity model where inelasticity may propagate along element at integration point locations.

In the nonlinear range, the structural behavior may be controlled with formation of plastic hinges at members' ends. This model is categorized as concentrated plasticity model where inelasticity is lumped at members' end-points or finite end-length. Using nonlinear spring at ends and linear behaviour within its length is a usual assumption for this model. Epistemic uncertainty in this analytical model is applied at the generalized force-displacement relationship for ends' springs as plastic hinges. Taking benefit of considering inherent uncertainties of analytical assumptions from experimental evidence for evaluating generalized force-deformation relationship may be considered as the major advantage of this method in respect to distributed plasticity models.

Where the most research works are conducted on considering the aleatory uncertainty in seismic performance assessment, the least works are devoted to incorporating modelling uncertainty pertaining to high computational cost. Although in moderate to low-seismic regions using deterministic modeling parameters (mean values) is adequate in estimation of seismic response demands, in high-seismic regions, incorporating these uncertainties is an important issue in this respect (Celik and Ellingwood, 2010). Not only for ductile buildings designed based on current building codes, but also for existing deficient buildings designed based on old codes, the record-to-record variability as well as modeling uncertainty is of the source of variability. The modeling uncertainty is more crucial, especially in highly nonlinear range of structure deformation (collapse capacity), because more uncertain parameters define the post-strength of material (Challa and Hall, 1994 and Vamvatsikos and Fragiadakis, 2010). For example, where the logarithmic standard deviation attributed to record-to-record uncertainties ranges between 0.35 and 0.45 depending on the structure of interest, the logarithmic standard deviation associated with modeling uncertainties may be as much as 0.45 (Haselton, 2006). Combining these uncertainties can make a significant impact on collapse probabilities, revealing the importance of incorporating modeling uncertainties in the seismic risk assessment. The modeling uncertainties increase the dispersion and also usually decrease the prediction of the median collapse capacity (Liel et al., 2009).

In this research, GMDH methodology is proposed to consider modelling uncertainties and to derive collapse fragility curves, incorporating epistemic and aleatory uncertainties, and it is finally compared with the other methods.

2 Statistical-based method

The full Monte Carlo Simulation (MCS) is the common method in estimation of failure probability in the presence of various uncertain variables. Employing MCS demands more computational cost and it is known as a time-consuming

method to cover all probabilistic distributions allied with each source of uncertainty. Although this approach is not feasible for large and huge structure with numerous random variables, it is presently only one to achieve precise and more accurate results.

To reduce computational effort, some reliability-based approaches for incorporating epistemic uncertainties, including First-Order Second-Moment (FOSM) (Ibarra and Krawinkler, 2005), (Celarec and Dolšek, 2013), First-Order Reliability Method (FORM), Second Order Reliability Method (SORM), Mean Estimate Method, Confidence Interval Method (Zareian and Krawinkler, 2007), ASOSM (Liel et al., 2009) and Response Surface Method (Karimi Ghaleh Jough and Beheshti Aval, 2018) were utilized.

The FOSM and FORM methods are more effective when the large number of model random variables wished to incorporate in uncertainty treatment. Both methods use the same assumption of linear approximation of failure surface. The main difference is that in FORM, the nonlinear failure surface is replaced at the design point (the nearest distance to origin in random variable space associated with highest failure probability) by linear approximation through Taylor expansion in spite of FOSM where failure probability computed based on two moments of probability distributions, i.e., mean and standard deviation of failure surface. For high curvature failure surfaces, using this linear approximation leads to imprecise failure probability. To overcome this issue quadratic approximation of SORM was introduced. Due to incomplete fitting of quadratic form to the actual failure surface and also more computational cumbersome inherent in this method, the linear approximations approach is more attractive than SORM for incorporating modeling uncertainties into fragility functions (Liel et al., 2009). Moreover, the application of these approximation methods needs explicit form of failure function to find the gradient close to design point. However, for complex structure, using IDA in conjunction with accurate nonlinear finite element model provides the implicit failure surface as a bunch of points (Hurtado, 2004).

To overcome aforementioned restrictions, Response Surface Method (RSM) has been suggested (Bucher and Bourgund, 1990). The response surface surrogates the time consuming nonlinear dynamic analyses and hence decreases computational cost of MCS.

Due to requiring reselection of the response function order (usually restricted to quadratic function), achievement of the right function is not guaranteed. Indeed, the RSM still requires significant computational cost and hence is considered as a time-consuming method for complex structural models (Liel et al., 2009).

For simplicity and reducing more computational and analysis time, the method like FOSM, so-called approximated second order second moment (ASOSM) was recently developed to treat the effects of modeling uncertainties (Liel et al., 2009). In spite of FOSM, while by incorporating modeling uncertainties, median of fragility function was maintained, in this approach, the median was shifted.

More recently the application of Heuristic based techniques in seismic risk assessment of structures is received more attention. Reducing computational effort across limiting results error and generality and efficiency were reported by developers. Met-heuristic algorithm doesn't require certain mathematical and problem solving, regardless of internal performance of the problem, optimization is done. This algorithm can solve every goal with every restriction (for example, linear or nonlinear), which is the continuous search space, or mixed batch is defined.

2.1 Heuristic-based methods

As mentioned, the epistemic uncertainties are a part of factors that cause of change in structure capacity and for which, a predictive model can be designed based on the existing data. Recently the soft-computing approaches are received much more attentions in pattern recognition science like the Fisher Discrimination, Bayesian Classification, Classification Trees, Wavelet Analysis (WA), Artificial Neural Network (ANN) and Probabilistic Support Vector Machines (PSUV) (Hurtado, 2004). The last three i.e. WA, ANN, and PSUV were recently utilized in seismic reliability assessment of structures (Noh et al., 2012) and (Mahmoudi and Chouinard, 2016). Despite aforementioned traditional uncertainty treatment methods explained in previous subsection, these heuristic fitting algorithms can simulate high nonlinear functions.

The WA was successfully proposed as a framework for a probabilistic damage classification model of steel moment-resisting frame under seismic loading. In this circumstance, the fragility functions were derived from the WA (Noh et al., 2012). In other effort PSVM was employed to develop the system fragility analysis of an existing bridge structure using results of IDA (Mahmoudi and Chouinard, 2016). The efficiency of application of the method in comparison with using MCS to develop fragility functions was reported. The epistemic uncertainty was not incorporated in both efforts.

Among aforementioned Artificial Intelligence approaches, ANN was employed to incorporate a wide range of uncertainty in seismic performance of structures, (Lagaros and Fragiadakis, 2007). In this respect, the response surface method was substituted by a trained Artificial Neural Network (ANN) method. The application of ANN method not only reduces the computational effort but also causes the reduction of estimation error compared with response surface-

based method. The key success factor of this heuristic method is a substitution of more time consuming performing nonlinear finite element analysis with trained associative memory of ANN. After training, ANN can be able to generate output responses from multiple realizations of random variables generated by MCS with less computational effort. The ANN-based MCS was also used for the sensitivity analysis of non-structure building such as large concrete dams (Mitropoulou and Papadrakakis, 2011). In another effort various uncertainties from lexical to informal and stochastic types were involved to develop comprehensive collapse fragility function for steel moment resisting frame (Karimi Ghaleh Jough and Şensoy, 2020).

So far, various ANN architectures have been proposed including the multilayer feed-forward, the radial basis function, Bayesian regularized networks with self-organizing maps, and recurrent neural networks. Among them, recently the multilayer feed-forward (Giovanis et al., 2015) and the radial basis function in Pang et al. (2014) have been used to develop fragility curves.

In this study, a new class of polynomial neural networks combining Group Method of Data Handling (GMDH) with ANNs are used to predict mean and standard deviation of fragility curve. This method is explored for modelling highly nonlinear and function approximation (Wang and Hu, 1999), (He and Wang, 2012). The GMDH network, that is an adaptive machine learning, is used for software reliability to predict variables during analysis. This algorithm has been created on the principle of heuristic self-organization. Their versatilities and comparison between these kinds of artificial intelligence across regular reliability based method against proposed method is the merit of this research.

2.1.1 Monte Carlo method based on GMDH approach

GMDH aims to simulate the elementary operation of biological self-organizing approaches of artificial intelligence and presents a computational function to predict approximate results based on input data [Farlow, 1984]. The standard description of the recognition problem is to achieve a function $\bar{g}$, which could be approximately applied instead of real one g, for predicting outcome $\overline{\boldsymbol{O}}$ for the input data vector $A = \left(a_1, a_2, a_3, ..., a_n\right)$ close to its real outcome *O*. So, the following N perceptions of multilayered structure is:

$$\boldsymbol{O_i = g(a_{i1}, a_{i2}, a_{i3} \dots, a_{in})\ i = 1, 2, 3, \dots, N} \tag{3}$$

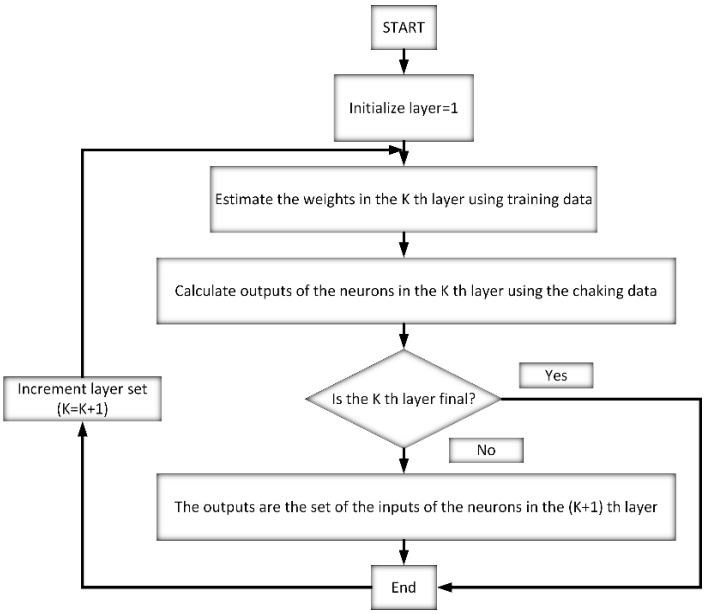


Figure 1: A flowchart of GMDH-type approach.

To simulate the outcome values $\overline{o_l}$ for each input variable$A = (a_{i1}, a_{i2}, a_{i3}, \dots, a_{in})$, it is available to train GMDH-type neural network. So that:

$$\boldsymbol{\overline{O_i} = \bar{g}(a_{i1}, a_{i2}, a_{i3}, \dots, a_{in})\ i = 1, 2, \dots, N} \tag{4}$$

The problem is now to decide the type of GMDH. Therefore, the square of variation in the real outcome and the simulated one is minimized, as:

$$\boldsymbol{min} = \sum_{i=1}^{N}[\bar{g}(a_{i1}, a_{i2}, a_{i3}, \dots, a_{in}) - O_i]^2 \quad (5)$$

Complex polynomial function can explain the relation between *A* and *O* in the following form:

$$\bar{O} = G(a_i, a_j) = b_0 + b_1 a_i + b_2 a_j + b_3 a_i a_j + b_4 {a_i}^2 + b_5 {a_j}^2 \quad (6)$$

A common flowchart GMDH-type is represented in Figure 1.

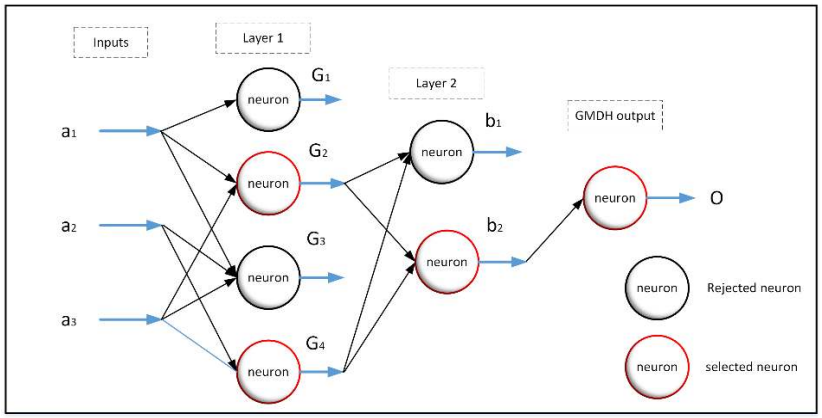


Figure 2: Final structure of GMDH network.

It can be usually observed that a tree chart is made by applying the quadratic function in equation 6 which its coefficients are calculated in a LSE (Last Square Error).

$$R^2 = \frac{\sum_{i=1}^{N}(o_i - \bar{o})^2}{\sum_{i=1}^{N} {o_i}^2} \quad (7)$$

The LS method from nonlinear regression results to solve the simple equations in the following form:

$$a = (A^T A)^{-1} A^T Y \quad (8)$$

Where,

$$A = \begin{bmatrix} 1 & a_{1p} & a_{1q} & a_{1p}a_{1q} & {a_{1p}}^2 & {a_{1q}}^2 \\ 1 & a_{2p} & a_{2q} & a_{2p}a_{2q} & {a_{2p}}^2 & {a_{2q}}^2 \\ 1 & a_{Mp} & a_{Mq} & a_{Mp}a_{Mq} & {a_{Mp}}^2 & {a_{Mq}}^2 \end{bmatrix} \quad (9)$$

That is the best vector of the quadratic function of Equation 6 for the entire set of M triple variables. The final structure of GMDH network is shown in Figure 2. After training the GMDH, Monte Carlo approach is used based on GMDH function. At first, a large number of scenario variables are simulated. This scenario is used to simulate the modeling variable and characteristic value of records. Second, mean and standard deviation of Sacollapse are obtained by trained GMDH. Figure 3 demonstrates the procedure considered in this study.

3 Application of the proposed method to the case study structure

In the proposed method, parameters with aleatory uncertainty are considered as random variables with specified probability distributions while parameters with epistemic uncertainty are taken as 125 scenarios of modelling parameters.

Table 1 Design sections for considered structure

story	C1	C2	B1	B2
1	BOX 180x180x1.6	BOX 200x200x1.6	IPE 300	IPE 330
2	BOX 180x180x1.6	BOX 200x200x1.6	IPE 300	IPE 330
3	BOX 180x180x1.6	BOX 200x200x1.6	IPE 300	IPE 330

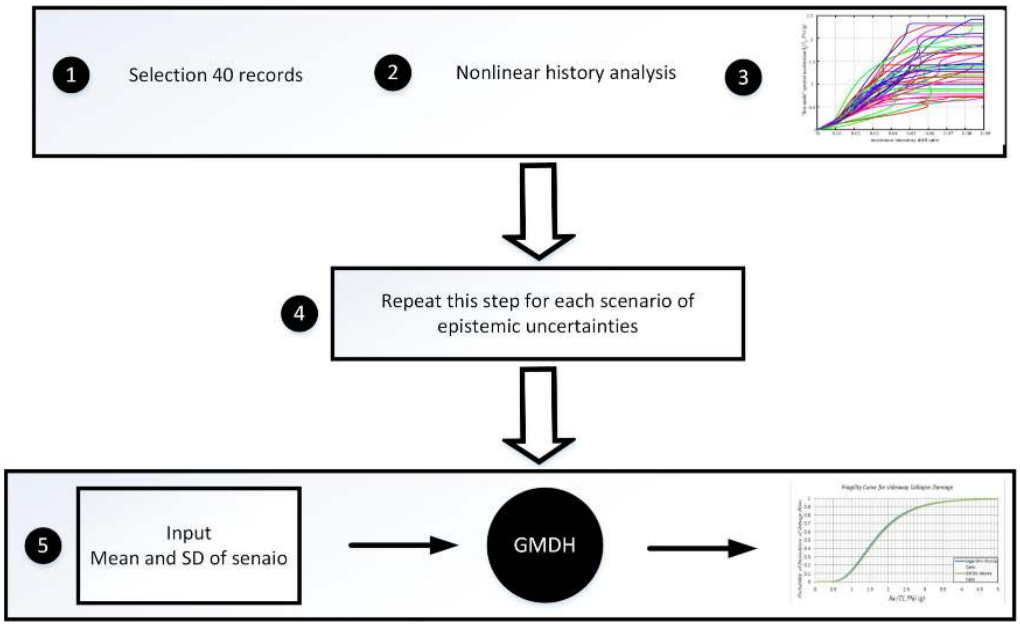


Figure 3: Fragility analysis with considering epistemic and aletory uncertainties by GMDH network

In the first stage, the mean and standard deviation of the fragility curve are determined, using Incremental Dynamic Analysis (IDA), as functions of parameters, which contains epistemic uncertainty; then parameters of the model are determined. The mean and standard deviation values of the fragility curve are predicted by Group Method of Data Handling (GMDH) approach for certain combinations of model parameter values. The Monte Carlo based on GMDH method instead of Monte Carlo response surface is used to predict these values. Monte Carlo based on the GMDH is considered both epistemic and aleatory uncertainties in deriving the fragility curves in the same way. This capability has not yet been employed to incorporate uncertainties in performance-based seismic risk analysis.

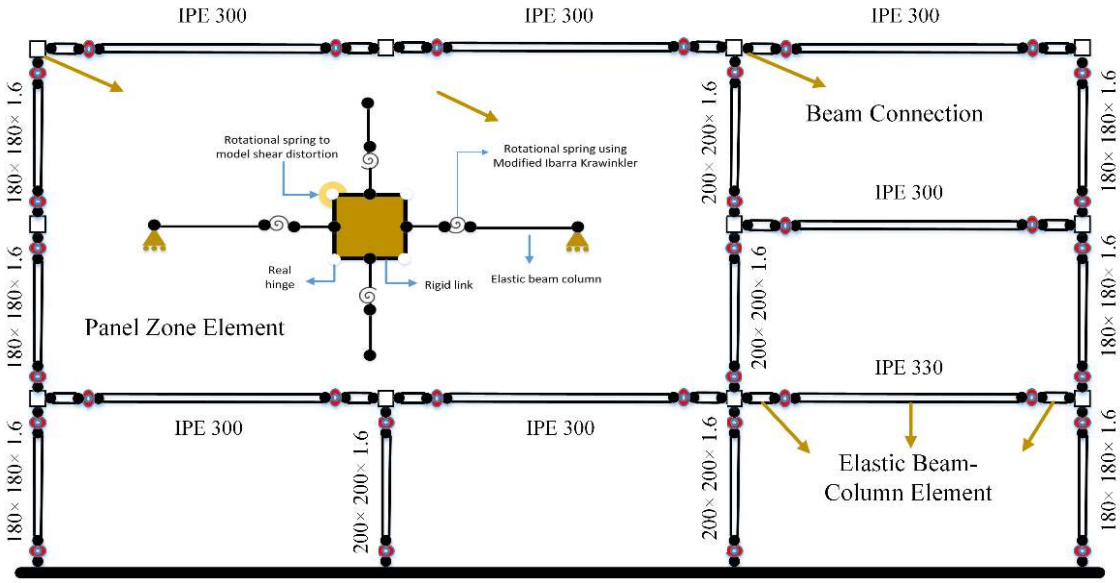


Figure 4: The frame model of the example structure and M2-WO panel zone

Three-story moment-resisting steel frame is considered as the case study to introduce the proposed method. It was designed according to UBC design provisions (UBC, 1997). The design acceleration response spectra of Iranian seismic code No-2800 (Code No. 2800, 2007) for seismic hazard zonation of Tehran, the capital city of Iran was used. The structural frame model is shown in Figure 4.

Example steel building is symmetric in plane and height; therefore 2D model is used in this study. Table 1 shows the cross sections of beams and columns in the example structure. The first mode period of the structure is 0.98 sec. The OpenSees program (OpenSEES, 2006) is used for modeling the case study structure. To represent the epistemic uncertainties in the proposed approach, Ibarra–Medina–Krawinkler (IMK) model has been applied based on rotational springs in the connections of the structure [Foutch and Yun, 2002]. Aleatory uncertainties are considered by applying a set of records in IDA analysis. IM-Based approach is used to compute the fragility curves based on the IDA approach [Zareian et al., 2010]. The strength and deformation limitation is determined by using the hysteresis backbone curve of IMK model that is shown in figure 5. The main parameters of backbone curve in IMK model include the initial stiffness (K_e), yield strength (My), stiffness of hardening branch ($K_s = \alpha_s Ke$), maximum strength (Mc) and its corresponding displacement (θc), post-capping stiffness point ($K_c = \alpha_c Ke$), and residual strength (M_r) and its corresponding displacement (θr). The backbone curve and hysteresis model of IMK of above parameters have been illustrated in Figure 5 [Ibarra and Krawinkler, 2005]. The cyclic deterioration ratio obeys the rule provided by [Rahnama and Krawinkler, 1993] which is formulated based on dissipated energy in each cycle of reciprocal loading.

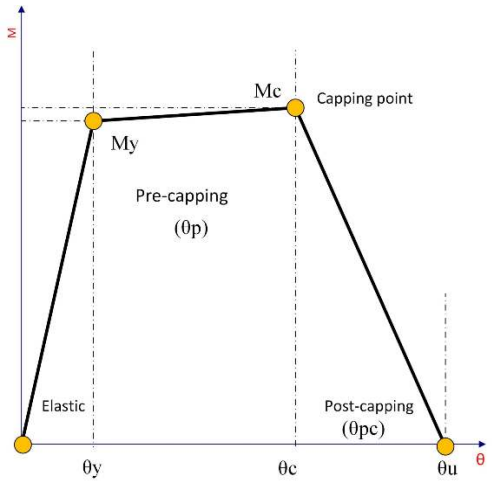


Figure 5: The backbone curve of IMK in Zolfaghari (2014)

Table 2 Mean and standard deviation and correlation calibration of modelling parameters.

Median θp	$\sigma_{\theta p}$ (rad)	Median θ_{pc}	$\sigma_{\theta pc}$	Median Λ	σ_{Λ}	$\rho_{\theta p,\theta pc}$	$\rho_{\theta p,\Lambda}$	$\rho_{\theta pc,\Lambda}$
0.025	0.43	0.16	0.41	1.00	0.43	0.69	0.44	0.67

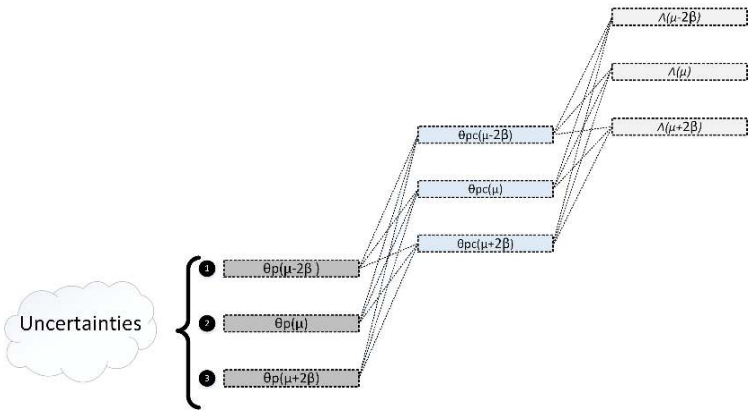


Figure 6: 125 scenarios for modelling parameters

The capacity of base energy is determined by (Equetion 10):

$$\boldsymbol{E_t = \gamma M_Y \theta_Y = \Lambda \theta_Y} \quad (10)$$

According to Equation 10, the capacity of hysteresis energy is defined as multiplying a coefficient to yield rotation. The capacity coefficient of dissipated hysteresis energy is determined by laboratory data and considered as an uncertainty modeling parameter. Considering the epistemic uncertain parameters as the moment-rotation curves of target springs in Ibarra-Krawinkler model, (θ_p), (θ_{pc}) and (λ) have been found to be the most important parameters related to ductility capacity and stiffness and strength deterioration in former studies (Ibarra et al., 2005). Probability distribution of Log-normal function is used to apply uncertainties due to (θ_p), (θ_{pc}) and (λ). Means, standard deviations, and correlation coefficients of these probability distributions, based on laboratory tests, are presented in Table 2(Lignos, 2008).

In this paper, a set of 40 records used by Medina have been applied for IDA analysis to simulate GMDH approach. To achieve input data for evaluated Monte Carlo based GMDH method, five scenario of each modeling variable are shown which are based on $[\mu - 2\beta,\ \mu - \beta, \mu, \mu + \beta, \mu + 2\beta]$(125 scenarios shown in Figure 6). For each scenario of input variable, the IDA is performed for each record and $Sa_{collapse}$ for each scenario is achieved and finally mean and standard deviation of collapse fragility curve is derived, considering spectral acceleration for first mode as IM and maximum drift as EDP. The sample of IDA and related collapse fragility curves are represented in figure 7.

Response Surface Method (RSM) used for simulating the mean and standard deviation of fragility curves at sidesway collapse limit state defined in equations is based on following logarithmic (equations 11 and 12).

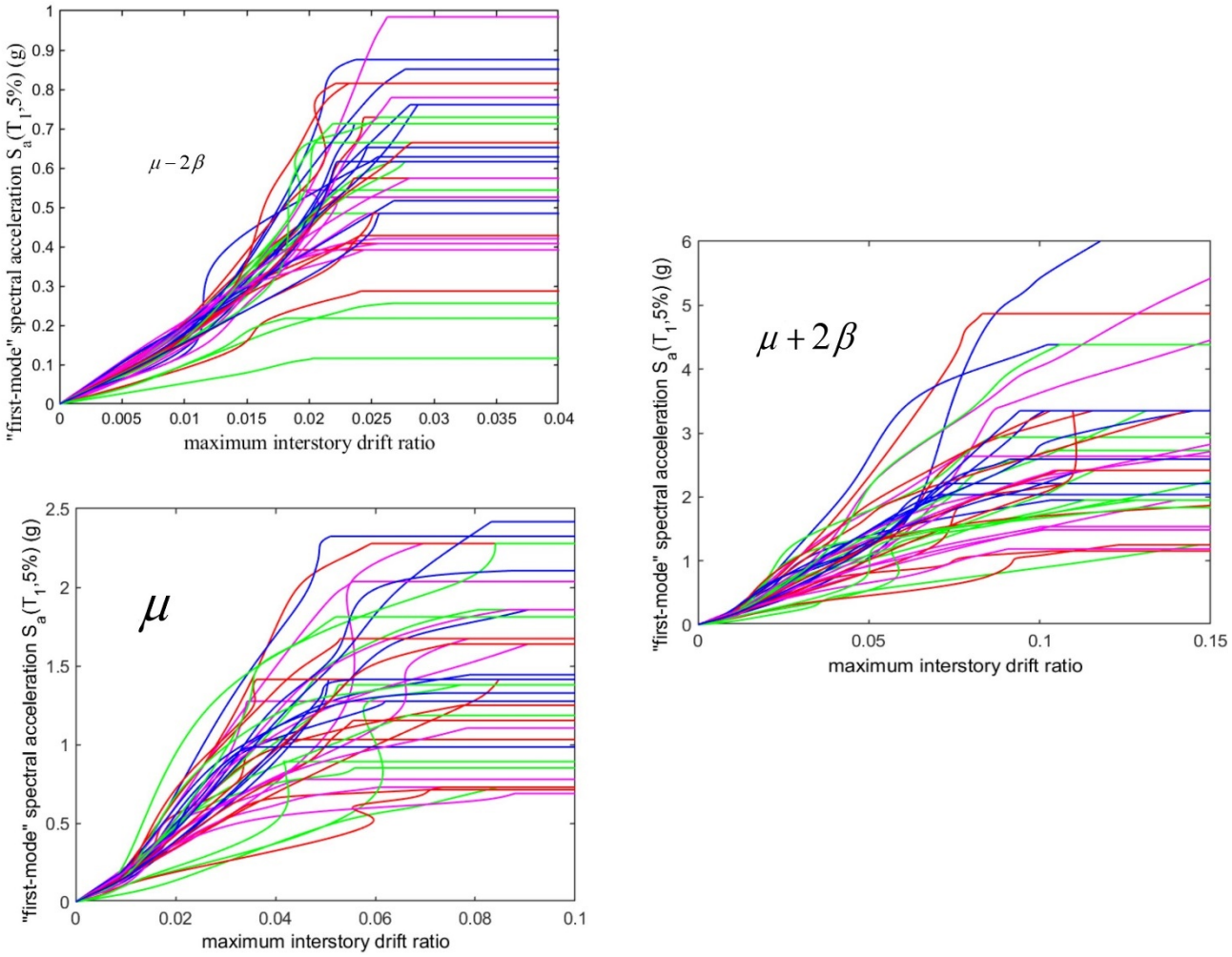


Figure 7: The sample IDA for mean value of variables [μ-2β, μ, μ+2β]

$$\boldsymbol{ln(\mu) = a_0 + a_1\, ln(x_1) + a_2\, ln\left(\frac{x_2}{x_1}\right) + a_3 ln\,(x_3)} \tag{11}$$

$$\boldsymbol{ln(\beta) = b_0 + b_1\, ln(x_1) + b_2\, ln\left(\frac{x_2}{x_1}\right) + b_3 ln\,(x_3)} \tag{12}$$

In RSM, 125 inputs are considered, corresponding to 125 scenarios for epistemic parameters. Computed coefficients in logarithmic function are represented in Table 3. In these inputs, 65% are chosen to simulate related functions, and applied for training the GMDH, while the GMDH model simulates mean and standard deviation of fragility curves. The outcome of this method has been compared against Monte Carlo-based logarithm function and FOSM and ASOSM approaches based on capability of fragility curve. Figure 8 shows the statistical results for predicting mean of fragility curve. In these figures, horizontal axis shows values that are evaluated by direct IDA, and vertical axis estimates regressed analytical functions. The solid green and red line shows the positions where approximate values are equal to IDA-based values.

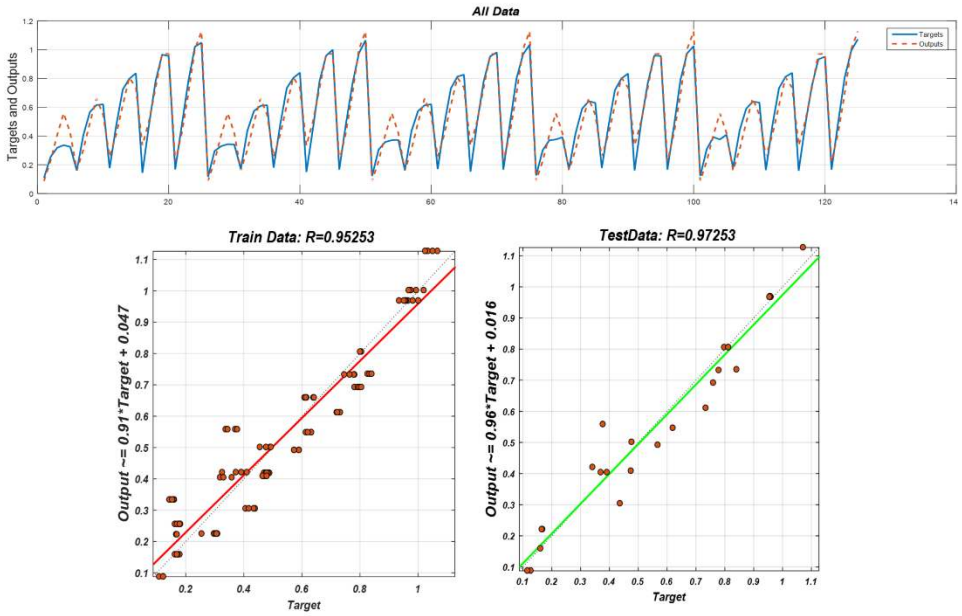


Figure 8: The statistical values obtained for mean of fragility curve by GMDH network-type method

Table 3 Constant coefficients of Monte Carlo based logarithm

	a_0	a_1	a_2	a_3	MSE
Mean	0.395	0.456	0.334	0.719	0.0104
	b_0	b_1	b_2	b_3	MSE
SD	-1.612	-0.063	0.143	0.083	0.0035

In order to evaluate the accuracy of mean and standard deviation predicted by GMDH, the Mean Square Error (MSE), the Root Mean Square Error (RMSE) and R-square (R^2) statistic measurements are determined as the accuracy criterion against the test data. The MSE, RMSE and R^2 between the exact and predicted responses are as (equations 13 and 14):

$$\boldsymbol{RMSE = \frac{n_i \sum_{i=1}^{n_i}(y_i - \bar{y}_i)^2}{(n_i - 1)\sum_{i=1}^{n_i}(y_i)^2}} \tag{13}$$

$$\boldsymbol{R^2 = 1 - \left[\frac{\sum_{i=1}^{n_i}(y_i - \bar{y}_i)^2}{\sum_{i=1}^{n_i}(\bar{y}_i)^2}\right]} \tag{14}$$

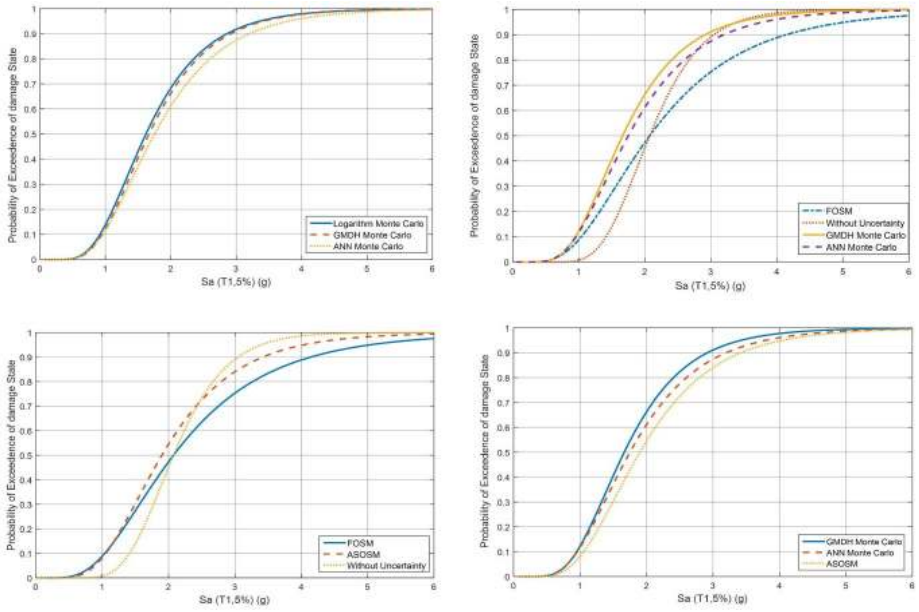


Figure 9: Compression of various methods in developing fragility curve with neglecting epistemic uncertainty

Table 4 Results of various methods on collapse fragility curves

Method	Logarithm	GMDH	ANN	FOSM	ASOSM	Neglecting
Mean value	0.4821	0.5097	0.56	0.7297	0.6417	0.7297
Standard deviation value	0.4436	0.4374	0.47	0.5398	0.4588	0.3
Error based GMDH	-5.41%	-	9.86%	43.16%	25.89%	43.16%

Table 5 Probability of collapse and mean annual frequency with considering various method

Method	Logarithm	GMDH	ANN	FOSM	ASOSM
P(Collapse|MCE)	0.1335	0.1372	0.1391	0.148	0.142
MAF ×(10-5)	5.59	6.24	6.29	6.33	6.39
Error based GMDH	-2.69%	-	1.38%	7.8%	3.4%

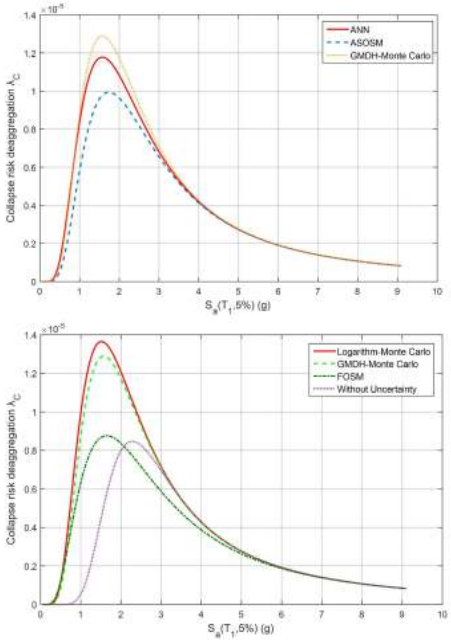


Figure 10: The result of λ_C deaggregation with considering the proposed method

In order to gain the best results, various GMDH network-type with the different number of neurons are used. Each of the GMDH network-type was trained and the accuracy of mean and SD was assessed by Equations 13 and 14. The best GMDH model corresponding to the LSE is recognized. The parameters of the GMDH model used to obtain these results are: Maximum Number of Neurons in a Layer= 5, Maximum Number of Layers= 7, Selection Pressure= 0.9 and Train Ratio= 0.9.In order to gain the best results, various GMDH network-type with different number of neurons are used. Each of the GMDH network-type was trained, and the accuracy of mean and SD was assessed by Equations 16 and 17. The best GMDH model corresponding to the LSE is recognized. The parameters of the GMDH model used to obtain these results are: Maximum Number of Neurons in a Layer= 5, Maximum Number of Layers= 7, Selection Pressure= 0.9 and Train Ratio= 0.9. The mean square error, applying trained GMDH network-type in prediction of mean and standard deviation values, are 0.124% and 0.0137%, respectively, which are compared with the other methods. Comparison of proposed fragility curves of structure, applying FOSM and ASOSM approaches is shown in Figure 9.

Using FOSM and ASOSM outcomes resulted in change of standard deviation of collapse fragility curves are compared without epistemic uncertainties. The mean of fragility curve value doesn't change using FOSM and ASOSM results. As it is shown in Figure 9, mean and standard deviation of example building are 0.5097 and 0.4374 in the GMDH network-type Monte Carlo, respectively. Also in this figure, it is observed that FOSM and ASOSM methods cannot predict more accurate results than GMDH network-type Monte Carlo and ANN methods. In order to compare these methods, different ANN (Multi-Layer perceptron) configuration with variable number of neurons in each layer is used. Each of the ANN configurations was trained, and the performance error was evaluated. Using such a trial-and-error approach, the best ANN model corresponding to the least error measure is determined. As a result, a 5-35-35-1 neural network containing two hidden layers with 35 neurons for each is finally selected. Besides, more care was taken to assess the effect of different training algorithms (such as Levenberg–Marquardt, Broyden– Fletcher–Goldfarb– Shanno (BFGS), Quasi-Newton and Scaled Conjugate Gradient). Again, by a trial-and error approach, Levenberg–Marquardt training algorithm is chosen.

The collapse fragility curve and mean annual frequencies of collapse (λc) are determined here to show the effect of the epistemic and aleatory uncertainties with the proposed method in comparison with other available methods. Computing (λ_C) comprises integrating the collapse fragility curve of the case study structure over the hazard curve at the site applying the following integration:

$$\boldsymbol{\lambda_c = \int_0^{\infty} P(C|im).|\frac{d\lambda_{IM}(im)}{d(im)}|d(im)} \tag{15}$$

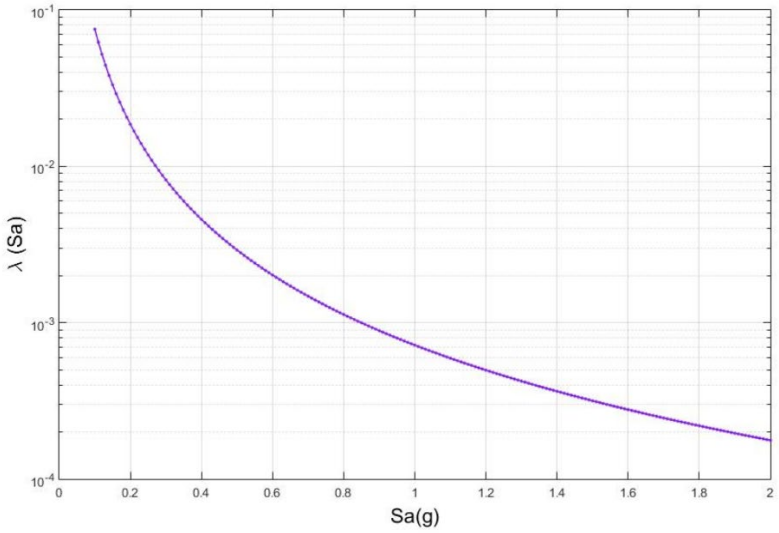


Figure 11: Hazard curve for sample structure in Tehran

Where the first and the second part are the collapse fragility function and the slop of the hazard curve, respectively. If the product of two mentioned functions is drawn versus intensity measure (here acceleration response spectra), the collapse risk deaggregation curve is achieved. This curve illustrates the contribution of various ground motion intensities in the collapse risk at the specified site. As shown in the deaggregation curve, spectral acceleration between 0.5 to 2 g shows the frame collapse risk. Figure 10 represents the deaggregation curves when applying the the slope of Tehran

seismic hazard curves with the collapse fragility curves achieved by the proposed method against other uncertainty propagation methods for the case study structure. As represented in this figure, the reasonable match occurs among GMDH Monte Carlo method, Logarithm Monte Carlo approaches and others. The best match in estimation of values at the mode of the deaggregation curve can be observed between GMDH Monte Carlo method and accurate Logarithm Monte Carlo method.

While this difference reaches approximately 15% in the first mode spectral acceleration, when epistemic uncertainties are not considered, it results in conservative response that reaches approximately 60% in the first mode spectral acceleration.

The mean and standard deviation of the collapse fragility curve are depicted in Table 4. The results of proposed approach represent good correlation with respect to those of Monte Carlo based logarithm simulation approach as well as in comparison with other methods. Table 5 illustrates probability of collapse in discreet hazard levels MCE (Maximum Considered Earthquake acceleration) Figure 11 and Mean Annual Frequency of exceedance (MAF) in a sample structure. The Probabilistic hazard analysis have been done for Tehran region (Zolfaghari, 2014). It can be observed that including modelling uncertainties increases about 8% at MCE hazard level.

4 CONCLUSION

In this study, Monte Carlo based GMDH is used to derive the sidesway collapse fragility curve considering modeling and aleatory uncertainties effects. The accuracy of the proposed approach was compared with other existing methods in respect of probabilistic outcomes. The effects of aleatory and modelling uncertainties are compared by aforementioned approaches on deriving the fragility curves. This paper represents the efficiency of the GMDH Monte Carlo method against the probabilistic approach through applying methods to a case study Steel Moment Resisting Frame (SMRF) structure.

It can be seen that while the results perfectly conform to those obtained using the Monte Carlo based logarithm method (which is used as the basis for comparison and validation of the proposed method), consideredepistemic uncertainties made changes in mean and standard deviation value (especially decrease of mean value of the fragility curve).

Finally, sidesway collapse fragility curves of the case study SMRF were computed through ASOSM, FOSM, ANN and Monte Carlo based-logarithm function and GMDH. Modelling parameter has been considered as (θ_p), (θ_{pc}) and (λ) which was abbreviated as IMK model. As shown in Figure 10, neglecting epistemic uncertainty causes underestimation of collapse fragility curves. Since the prediction error by GMDH approach is less than Monte Carlo based-logarithm function, it can be resulted that prediction of mean and SD of fragility curve by Monte Carlo method based on GMDH approach would obtain more accurate results. Achieved outcomes represent that GMDH -based approach shows less reduction of mean values compared with logarithmic response surface ASOSM and FOSM methods, while more increments of standard deviation are obtained.

In comparison the proposed method with counterpart method, i.e. ANN, although dispersions in mean and standard deviation are large, finding the proper layer in ANN to predict accurate values is more cumbersome. Advantage of this method is to overcome the problem of choosing proper network size in applying multi-layer perception ANN. Such system of network's topology prepares the best networks in hidden layers and their connectivity so that there is a perfect match between this network and logarithm function, which is represented in Figure 10. It should be noted that GMDH approach is a universal simulator that is able to accurately predict the target output and is not limited to this example.

Author Contributions: Conceptualization, FK Jough, SB Beheshti; Methodology, FK Jough, SB Beheshti; Software, FK Jough, M Vegar; Investigation, FK Jough, M Vegar; Writing - original draft,, FK Jough, SB Beheshti; Writing - review & editing, FK Jough, M Vegar, SB Beheshti; Funding acquisition, FK Jough, M Vegar; Resources,FK Jough; Supervision, FK Jough, SB Beheshti.

Editor: Pablo Andrés Muñoz Rojas.

References

Bradley, Brendon A. (2013). A critical examination of seismic response uncertainty analysis in earthquake engineering. Earthquake Engineering & Structural Dynamics, Vol. 42, Issue 11: 1717-1729.

Bucher C G., Bourgund U. (1990). A fast and efficient response surface approach for structural reliability problems. Structural Safety, Vol. 7: 57-66.

Challa V R Murty., Hall John F. (1994) Earthquake collapse analysis of steel frames. Earthquake Engineering & Structural Dynamics, Vol. 23, Issue 11: 1199-1218.

Celarec D., Dolšek M. (2013). The impact of modelling uncertainties on the seismic performance assessment of reinforced concrete frame buildings. Eng Struct, Vol. 52: 340–354.

Celik OC., Ellingwood BR. (2010). Seismic fragilities for non-ductile reinforced concrete frames: role of aleatoric and epistemic uncertainties. J Struct Saf Vol 32. Issue 1: 1–12.

Code No. 2800. (2007). Iranian Code of Practice for Seismic Resistant Design of Buildings. vol. 3rd Edition, ed. Iran: Building and Housing Research Center.

Farlow S J. (1984). Self-organizing methods in modeling: GMDH type algorithms, 54, CrC Press.

Foutch D A., Yun Y S. (2002). Modeling of steel moment frames for seismic loads. Journal of Constructional Steel Research, Vol. 58: 529-564.

Ghobarah. A. (2001). Performance-based design in earthquake engineering: state of development, Engineering structures, Vol. 23: 878-884.

Giovanis D G., Fragiadakis M., Papadopoulos V. (2015). Epistemic uncertainty assessment using Incremental Dynamic Analysis and Neural Networks, Bull Earthquake Eng.

Haselton C B. (2006) Assessing seismic collapse safety of modern reinforced concrete frame buildings. Ph.D. thesis, (Stanford University).

He J N., Wang Z. (2012). Analysis on System Reliability of Steel Framework Structure and Optimal Design. in Applied Mechanics and Materials: 902-906.

Hurtado J E. (2004). An examination of methods for approximating implicit limit state functions from the viewpoint of statistical learning theory. Structural Safety, Vol. 26: 271–293

Ibarra L F., Medina R A., Krawinkler H. (2005). Hysteretic models that incorporate strength and stiffness deterioration. Earthquake engineering and structural dynamics, Vol. 34: 1489-1512.

Ibarra L F., Krawinkler H. (2005). Global collapse of frame structures under seismic excitations: Pacific Earthquake Engineering Research Center.

Jough, F. K. G., & Şensoy, S. (2016). Prediction of seismic collapse risk of steel moment frame mid-rise structures by meta-heuristic algorithms. Earthquake Engineering and Engineering Vibration, 15(4), 743-757.

Karimi Ghaleh Jough, F., & Şensoy, S. (2020). Steel Moment-Resisting Frame Reliability via the Interval Analysis by FCM-PSO Approach considering Various Uncertainties. Journal of Earthquake Engineering, 24(1), 109-128.

Karimi Ghaleh Jough, F., Beheshti Aval, S. (2018). Uncertainty analysis through development of seismic fragility curve for an SMRF structure using an adaptive neuro-fuzzy inference system based on fuzzy C-means algorithm. *Scientia Iranica*, 25(6), 2938-2953.

Lagaros N D., Fragiadakis M. (2007). Fragility assessment of steel frames using neural networks, Earthquake Spectra, Vol. 23: 735-752.

Liel A B., Haselton C B., Deierlein G G., Baker J W. (2009). Incorporating modeling uncertainties in the assessment of seismic collapse risk of buildings. Structural Safety, Vol. 31: 197-211.

Lignos D. (2008). Sidesway collapse of deteriorating structural systems under seismic excitations (ProQuest).

Mahmoudi S N., Chouinard L. (2016). Seismic fragility assessment of highway bridges using support vector machines, Bull Earthquake Eng, Vol. 14: 1571–1587.

Mitropoulou CC., Papadrakakis M. (2011). Developing fragility curves based on neural network IDA predictions, Engineering Structures, Vol. 33: 3409-3421.

Möller B., Beer M. (2013). Fuzzy randomness: uncertainty in civil engineering and computational mechanics: Springer Science & Business Media.

Noh H Y., Lignos D.G., Krishnan Nair K., Kiremidjian A S. (2012). Development of fragility functions as a damage classification/prediction method for steel moment resisting frames using a wavelet-based damage sensitive feature. Earthquake Engng Struct. Dyn, Vol. 41: 681–696.

OpenSEES. (2006). Open System for Earthquake Engineering Simulation, Pacific Earthquake Engineering Research Centre, University of California, Berkeley.

Pang Y., Dang X., Yuan W. (2014). An Artificial Neural Network Based Method for Seismic Fragility Analysis of Highway Bridges, Advances in Structural Engineering, Vol. 17 No.

Rahnama M., Krawinkler H. (1993). Effect of soft soils and hysteresis models on seismic design spectra. John A. Blume Earthquake Engineering Center Report.

UBC. (1997). Uniform building code, in Int. Conf. Building Officials.

Vamvatsikos D., Fragiadakis M. (2010). Incremental dynamic analysis for estimating seismic performance sensitivity and uncertainty. Earthq Eng Struct Dyn Vol. 39, Issue 2: 141–163.

Wang X Y., Hu WY. (1999). The self-organizing polynomial network algorithm based on the self-organizing theory. System Engineering --- Theory & Practice, Vol. 19, Issue 4: 51-56

Zareian F., Krawinkler H. (2007). Assessment of probability of collapse and design for collapse safety," Earthquake \ Engineering & Structural Dynamics, Vol. 36: 1901-1914.

Zareian F., Krawinkler H., Ibarra L., Lignos D. (2010). Basic concepts and performance measures in prediction of collapse of buildings under earthquake ground motions, The Structural Design of Tall and Special Buildings, Vol. 19, Issue 1-2: 167-181.

Zolfaghari M R. (2014). Development of a synthetically generated earthquake catalogue towards assessment of probabilistic seismic hazard for Tehran. Natural Hazards, Vol. 76, Issue 1: 497-514.