Efficient seismic fragility analysis of structures based on probabilistic machine learning method

Jia-Yi Ding

Ph.D. Student, School of Civil Engineering, Southeast University, Nanjing 211189, China

Qin-Cheng Hu

Ph.D. Student, School of Civil Engineering, Southeast University, Nanjing 211189, China

De-Cheng Feng

Professor, School of Civil Engineering, Southeast University, Nanjing 211189, China

ABSTRACT: Machine learning (ML) technology have been widely adopted in engineering practices recently by establishing the relationship between seismic intensities and the corresponding structural demands based on very limited experimental or numerical simulation databases. Nevertheless, its predictions are essentially deterministic, which largely have ignored the uncertainty inherent within the structural system. In this study, a new ML algorithm, that is natural gradient boosting (NGBoost), is applied to directly evaluate the conditional probability distribution $P(y|x)$ instead of producing a point estimate $E[y|x]$ for each value of x. This type of probability prediction can directly obtain the failure probability of the structure under an intensity measure (IM), while avoiding the extra input of uncertainties from structure properties in traditional ML methods. Therefore, the evaluation of structural seismic fragility is especially simple and efficient. Finally, seismic fragility analysis of a typical 3-span-6-storey reinforced concrete frame (RCF) is carried out to illustrate and demonstrate the NGBoost model. The results indicate that the average accuracy of the NGBoost agrees with conventional ML algorithms reasonably well in seismic fragility analysis, while without re-input for each additional set of structural material properties and geometric parameters.

Reliable and rapid estimate of the fragility of structures against seismic hazards have become one of the most essential tasks in civil engineering. The fragility curve is defined as the conditional failure probability that the damage measure (DM) of a structure reaches or surpasses the limit states under a selected ground motion intensity measure (IM). The conventional approaches to obtain DMs in seismic fragility analysis include very limited experimental data and expensive high-fidelity numerical simulation (Guan et al. 2021; Lyu et al. 2020). Experimental models often bear inevitable simplifications or idealizations, and offer high accuracy only when the prediction effort targets structures or components that closely resemble the validation experiments (boundary conditions, scale, configuration, geometry, and so forth). Predicting complex behavior using physics and first-principle-based approaches, on the other hand, requires highly complex (e.g., three-dimensional finite-element) simulation models, which currently have limited use in routine civil engineering design. With the aid of numerical analysis software with enormous computing power, accurate DMs can be extracted by the nonlinear time history analysis (NLTHA) (Mangalathu and Jeon 2019). Moreover, in the fragility analysis based on numerical simulation, the uncertainty of ground motion and structural parameters can be easily considered. However, the fragility calculation using accurate and conceptually straightforward Monte Carlo simulation (MCS) is time-consuming, especially for complex structures, which currently have limited use in routine civil engineering design.

The existing methods for generating fragility curves of structural are cloud (Cornell et al. 2002), stripe (Mangalathu and Jeon 2019) and incremental dynamic analysis (IDA) (Vamvatsikos and Cornell 2002) approaches. In the cloud method a linear regression between seismic response and IMs in logarithmic space is established with unscaled ground motions for NLTHA. Although this method requires relatively small amount of NLTHA and is easy to implement, its prediction accuracy may be poor due to without a wide range of ground motion IM. For the stripe method, the ground motion is expanded to the same IM level, and the probability distribution of DMs under this intensity is calculated. The IDA method continuously expands a ground motion to obtain the IDA curve reflecting the structural response under different IM. Then the IDA curves corresponding to all ground motion records are obtained repeatedly. Obviously, the perform of a large number of NLTHAs in the stripe and IDA approaches requires considerable computing time.

Various machine learning technologies, support vector machines (SVM) (Cherkassky 1997), random forest (RF) (Breiman 2001), extreme gradient boosting (XGBoost) (Chen 2016), and artificial neural networks (ANN) (de Lautour and Omenzetter 2009), have been used to solve various problems in the field of civil engineering, such as the classification of failure mode and the prediction of associated shear strength for reinforced concrete beam-column joints (Mangalathu and Jeon 2018), performance classifications and predictions for reinforced masonry shear walls (Siam et al. 2019), and the prediction of shear strength for reinforced concrete deep beams (Feng et al. 2021). Similarly, these ML methods can also be applied to the seismic fragility analysis of structures, because their model establishment relies on data-driven without considering the physical mechanism. ML methods can establish the relationship between seismic intensities and the corresponding structural demands based on experimental or numerical simulation databases. ML methods has the advantages of low calculation cost and high estimation accuracy in predicting DM and evaluating fragility of structure. Nevertheless, its predictions are deterministic, largely have ignored the uncertainty inherent within training and testing data.

Instead of producing a point estimate like $E[y|x]$, the conditional probability distribution $P(y|x)$ for each value of x should be evaluated (i.e., probabilistic estimation includes the Bayesian approaches (Kim et al. 2020) and NGBoost (Chen et al. 2022; Duan 2019)).

In this paper, an efficient probabilistic ML method is proposed for seismic fragility analysis of structures. The present paper is organized as follows: In Section 1, natural gradient boosting (NGBoost) is detailly introduced. The framework for seismic fragility analysis using the NGBoost method is provided in Section 2. A reinforced concrete frame (RCF) is analyzed and the performance of the proposed approach is discussed in Section 3 and Section 4, respectively. Conclusions are provided in Section 6.

1. NATURAL GRADIENT BOOSTING FOR PROBABILISTIC PREDICTION

The NGBoost can produce a prediction of probability distribution $P_\theta(y|\boldsymbol{x})$, where parameters $\boldsymbol{\theta}$ (e.g., mean μ and standard deviation σ) is the functions of $\boldsymbol{x}$. The loss function and the gradient in above gradient boosting will not be feasible to gradually obtaining the optimal probability prediction (Chen et al. 2022; Duan 2019). There are a new scoring rule S and Natural Gradient be selected in the NGBoost.

Scoring rule $S(\boldsymbol{\theta}, y)$ is used to evaluate the difference between a predicted distribution and an actual observation distribution, which is similar to the loss function $L(F(\boldsymbol{x}), y)$ of point prediction in the deterministic ML model. The maximum-likelihood estimation (MLE) (1993 Eliason) score $\mathcal{L}$ is commonly used scoring rule, which is given by

$$\mathcal{L}\left[\boldsymbol{\theta}, y\right] = -\log P\left(\boldsymbol{\theta} \mid \boldsymbol{x}, y\right) \qquad (1)$$

where $P(\boldsymbol{\theta}|\boldsymbol{x}, y)$ is the likelihood of parameters $\boldsymbol{\theta}$. Continuous ranked probability score (CRPS) is another scoring rule(Gebetsberger et al. 2018).

The negative gradient is the steepest descent direction would be to minimize the loss function, which relies on Euclidean distance between two parameter vectors and cannot represent the true "distance" between two distributions. In distribution space, the distance between two probability distributions determined by the parameter vector is expressed by KL divergence (Duan 2019) as follows

$$D_{KL}(Q \parallel P) = \underset{Q(x|\theta)}{\mathrm{E}}\left[\log P(\boldsymbol{x} \mid \boldsymbol{\theta})\right] - \underset{Q(x|\theta)}{\mathrm{E}}\left[\log Q(\boldsymbol{x} \mid \boldsymbol{\theta})\right] \tag{2}$$

The natural gradient does is the direction of steepest descent in distribution space that minimizes the scoring rule $\mathcal{L}(\boldsymbol{\theta}, y)$ is expressed as

$$\begin{aligned}\tilde{\nabla}_{\theta}\mathcal{L}\left[\boldsymbol{\theta}, y\right] &= \lim_{\varepsilon\to 0} \underset{d:\, D_{KL}\left[P_{\theta}\|P_{\theta+d}\right]}{\arg\max} \mathcal{L}\left[\boldsymbol{\theta}+d, y\right] \\ &= I_{\mathcal{L}}(\boldsymbol{\theta})^{-1}\nabla_{\theta}\mathcal{L}\left[\boldsymbol{\theta}, y\right]\end{aligned} \tag{3}$$

where $P_{\theta}=P_{\theta}(y|\boldsymbol{x})$; $\nabla_{\theta}\mathcal{L}[\boldsymbol{\theta}, y]$ is conventional gradient with respect to $\boldsymbol{\theta}$; $I_{\mathcal{L}}(\boldsymbol{\theta})$ is Fisher Information that the target variable y carries about the probability distribution, which can be written as

$$I_{\mathcal{L}}(\boldsymbol{\theta}) = \mathrm{E}_{y\sim P_{\theta}}\left[\nabla_{\theta}\mathcal{L}\left[\boldsymbol{\theta}, y\right]\nabla_{\theta}\mathcal{L}\left[\boldsymbol{\theta}, y\right]^{\mathrm{T}}\right] \tag{4}$$

To specify the algorithm, a set of training samples $D=(\boldsymbol{x}_i, y_i)$, $i=1, 2, ..., N$, are given, boosting iterations M, learning rate η, type of the probability distribution P_{θ}, scoring rule S, and weak prediction model f are determined. The type of the probability distribution for regression problem (studied in this paper) in the NGBoost include normal, lognormal, and exponential distributions.

At the beginning of the algorithm, an initial parameter $\boldsymbol{\theta}^0$ based on the training set is obtained:

$$\boldsymbol{\theta}^0 = \arg\max_{\theta} \sum_{i=1}^{N} \mathcal{L}\left[\boldsymbol{\theta}, y_i\right] \tag{5}$$

Then, for each decision tree the individual natural gradients $G_i^{(j)}$ with respect to the estimated parameters up to that stage as $G_i^{(j)} = \tilde{\nabla}_{\theta}\mathcal{L}\left[\boldsymbol{\theta}_i^{j-1}, y_i\right]$.

The natural gradients $G_i^{(n)}$ together with the input vectors $\boldsymbol{x}_i$ to train the set of base learners of that stage $f^{(j)}(\boldsymbol{x})$. Next, the distribution parameters of this stage are updated as follows：

$$\theta_i^{(j)} = \theta_i^{(j-1)} - \eta\left(\rho^{(j)} f^{(j)}\left(x_i\right)\right) \tag{6}$$

$$\rho^{(j)} = \arg\min_{\rho} \sum_{i=1}^{M} \mathcal{L}\left(\boldsymbol{\theta}_i^{(j-1)} - \rho f^{(j)}(x_i), y_i\right) \tag{7}$$

where, $\rho^{(j)}$ is the scaling factor, which scales equally on all distribution parameters. When the boosting iterations M is reached, the training is over. Finally, the mean and the variance prediction models $F^{(M)}= (F_{\mu}^{(M)}, F_{\sigma}^{(M)})$ are obtained.

2. FRAMEWORK FOR SEISMIC FRAGILITY ANALYSIS

In the seismic fragility analysis, the conditional probability that the structural responses exceed the threshold of a certain damage level under each ground motion IM needs to be calculated. The responses of buildings and bridges typically obtained by employing NLTHA, where the input variables consist of structural parameters x_S and ground motion parameters x_{GM}. The outputs are the structural response y_R that represented by different damage measures of the structure under any earthquake, such as maximum inter-story drift ratio (MIDR) and curvature ductility (CD). The damages of a structure are generally divided into four states: slight, moderate, extensive, and complete damages (HAZUS 2022). As a surrogate model to approximate the expensive NLRHAs, the ML model is trained based on sufficient and perfect samples set to efficiently evaluate the output responses. The implementation procedure of NGBoost-based seismic fragility analysis is shown in Fig. 1.

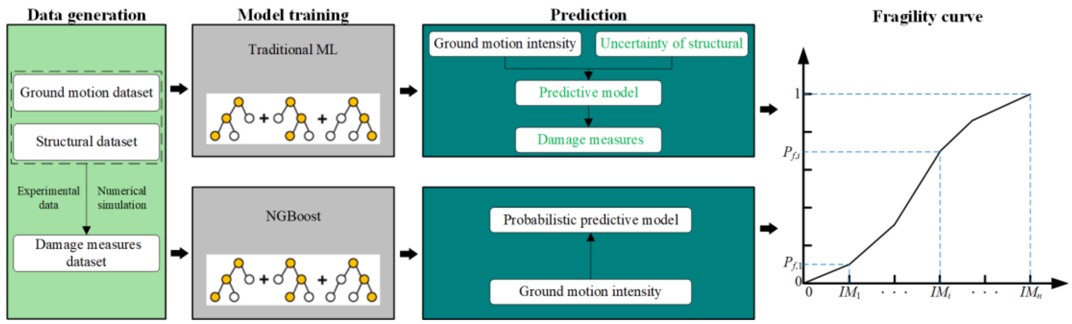

Fig. 1. Implementation procedure of NGBoost-based seismic fragility analysis.

2.1 Data generation

The uncertainty in geometry and material parameters of structure needs to be reasonably characterized. In this study, Latin hypercube sampling is used to conduct efficient sampling according to the preset distribution of input variables. Then Open System for Earthquake Engineering Simulation Platform (OpenSees) (McKenna 2006) is employed to construct the NLTHA models. Finally, a large number of numerical models based on variable input samples (including structural parameters x_S and ground motions parameters x_{GM}) are conducted to obtain datasets (x_S, x_{GM}, y_R) where y_R is output responses.

2.2 Model training

A mathematical model to predict the probability distribution $P_\theta(y_R|x_S, x_{GM})$ for each the structural and ground motion parameter vectors (x_S, x_{GM}) can be constructed based on the NGBoost. In order to perform NGBoost training and testing, the datasets are split into training and testing sets, e.g., 75% and 25% of the total data, respectively.
Before starting training, like the traditional gradient boosting algorithm, five key hyperparameters of the NGBoost needs to be initialized, including the frame-level hyperparameters (i.e., the number of weak learners M, and learning rate η) and model-level hyperparameters (i.e., maximum depth D, minimum sample numbers for splitting N_S, and minimum sample numbers for a leaf N_L). Firstly, η, D, N_S, and N_L are set as the initial value and the training will be stopped when M is reached. It is also necessary to predefine the distribution type of response variables in NGBoost, such as normal, lognormal, and exponential distributions. Then, the model training is performed based on the training data set, where the hyperparameters are iteratively optimized through the grid-searching method (Chen et al. 2022). In grid-searching, the 10-fold cross validation (CV) score is used to evaluate the performance of the model, which usually avoids the unexpected deviation caused by random sampling.

Finally, the accuracy of the model is investigated further through the testing set using various metrics such as root mean squared error (*RMSE*) and the coefficient of determination (R^2). If *RMSE* is near zero and R^2 is near 1.0, this means good accuracy of the prediction potential.

2.3 Fragility curves

For NGBoost-based seismic fragility assessment after the probability prediction model is established, the basic steps are:

a) Determine a group of ground motion intensities.

b) Conduct probability prediction model with IMs to obtain the conditional distribution of corresponding output responses (i.e., mean μ_y and standard deviation σ_y), then directly calculate the exceedance probability of structure damage under given value of IM, as follows

$$P\left(IMs = x_{GM}\right) = 1 - \Phi\left(\frac{y_{LS} - \mu}{\sigma}\right) \tag{8}$$

c) Repeat **Step** b) for each value of IM to obtain the fragility curves.

Obviously, the acquisition process of fragility curves based on the NGBoost is very concise, that avoid the generation of random samples for structural parameters and the time-consuming Monte Carlo method in calculating the exceedance probability.

3. DATABASE PREPARATION FOR NGBOOST PREDICT

3.1 Ground motion records

A suite of 100 ground motion records were obtained from the PEER database to evaluate seismic structural fragility. However, the most of them are between 0.3g and 0.6g. Under such earthquake intensity, it is difficult to produce enough structural damage at different levels, especially complete collapse. Therefore, the randomly selected 15 ground motion records are successively scaled according to *Sa* from 0.05g to 1.6g by 0.05g per step to obtain 32×15=480 ground motions records. The unselected 100−15=85 ground motions are randomly scaled 2 times to obtain 2×85=170 ground motions. The corresponding 650 set of IMs including PGA, PGV, SAI, EPA, and Sa_{T1} are calculated according to Ref. (Wang et al. 2017) as ground motion inputs.

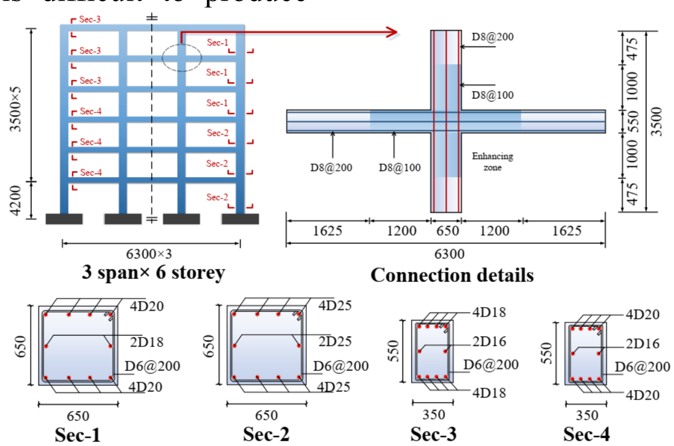

Fig. 3. Fiber-based model of 3-span-6-storey reinforced concrete frame.

Table 1. RCF attributes and its probability distribution

Features	Type	Mean	COV	Lower	Upper
Concrete bulk density	N	26.5 (kN/m^3)	0.0698	20.95	32.05
Core concrete compressive strength	L	33.6 (MPa)	0.21	17.63	61.32
Core concrete peak strain	L	0.0022	0.17	0.0013	0.0036
Core concrete ultimate strain	L	0.0113	0.52	0.0023	0.044
Cover concrete compressive strength	L	26.1 (MPa)	0.14	17.02	39.26
Cover concrete ultimate strain	L	0.004	0.2	0.0022	0.0071
Rebar diameter in columns	N	25 (mm)	0.04	22	28
Rebar diameter in beams	N	20 (mm)	0.04	17.6	22.4
Rebar diameter in beams	N	18 (mm)	0.04	15.84	20.16
Rebar yielding strength	L	378 (MPa)	0.074	302.0	470.5
Rebar elastic modulus	L	201 (kMPa)	0.033	181.96	221.79
Rebar hardening ratio	L	0.02	0.2	0.0108	0.0355
Damping ratio	N	0.05	0.1	0.035	0.065

Note: LN, lognormal; N, normal; U, uniform.

3.2 Building numerical model and associated uncertainties

A typical 3-span-6-storey reinforced concrete frame (RCF) shown in Fig. 3, is explored to demonstrate the proposed approach in this study.

The RCF can be efficiently and accurately simulated based on the fiber model using OpenSees. Compared with the macro model based on solid elements, the fiber-based model can reduce the time cost while maintaining the structural features, which is very advantageous for the seismic fragility analysis requiring a large number of model calculations. The process of fiber-based numerical model establishment on OpenSees software is detailed in Ref. (Cao 2023).

The prototype of the frame structure is a 3-span-6-storey high rest building. Due to the errors of design properties and construction conditions, the uncertainties of geometric and material properties should be considered in the numerical model of the frame structure. Based on the insights from previous research, 13 structure-related uncertainties are considered as random variables. Table 1 summarizes the considered random variables and their statistical properties. The LHS is used to generate random samples to transfer uncertainties of structural parameter to fiber-based numerical models. In accordance with the number of ground motions, a total 480+170=650 sets of random samples for these 13 structural variables are implemented.

Then, 650 fiber-based models were batch calculated based on OpenSees software to obtain the MIDR of the RCF as the output response. It should be pointed out that, the structural parameters are randomly matched with the ground motions for model analysis. Finally, the datasets were obtained with input variables, including 13 structural parameters x_S and 5 ground motion parameters x_{GM}, and output variables y_R.

3.3 Training and prediction

The whole datasets were divided into training sets with 480 selected ground motions and test sets with 170 unselected ground motions, respectively. At the beginning of training, hyper-parameters M, η, D, N_S, and N_L are recommended to be 300, 0.01, 5, 4 and 1. In order to obtain high performance probability prediction model, the grid-searching method based on the 10-fold CV is used optimize the hyperparameters. In addition, the common normal and lognormal distribution are compared for discussing the impact of the selection of response distribution type on the model.

The optimal hyperparameters and evaluation indicators corresponding to two different distributions are compared in Table 2. Normal distribution is better than lognormal distribution in R^2 index and *RMSE* of training set, but lognormal distribution is better in test set. It should be noted that the test set is used to investigate the prediction accuracy of the model, thus its corresponding indicators are expected to be better.

4. SEISMIC FRAGILITY ANALYSIS USING PROBABILISTIC NGBOOST MODEL

4.1 Performance of the prediction

To illustrate the applicability of the NGBoost method in seismic fragility analysis, common ML algorithms, such as support vector machines (SVM), random forest (RF), and extreme gradient boosting (XGBoost) are also used to predict the MIDR. Table 3 summarizes the R^2 and *RMSE* of

Table 3. Comparison of R^2 and RMSE for various ML methods (training set and test set) in this RCF

Error Measure	SVM	RF	XGBoot	NGBoost
R^2 (training set)	0.98	0.98	**0.99**	0.95
R^2 (test set)	0.70	0.86	**0.88**	0.86
RMSE (training set)	0.13	0.12	**0.09**	0.21
RMSE (test set)	0.43	0.28	**0.27**	0.30

MIDR. Table 3 summarizes the R^2 and *RMSE* of
above-mentioned four methods. The XGBoost-based prediction model shows a best performance for point estimation, while the NGBoost-based probability prediction model performs is very similar to the XGBoost, but obviously better than SVM. It should be further explained that in addition to providing relatively accurate point estimation, the NGBoost-based probability model can also give the variance of the point to reflect the inherent uncertainty of the response.

The established NGBoost model is applied to evaluate the seismic fragility of the RCF, and the specific steps are described in Section 4.3. In this study, the MIDR corresponding to the four earthquake-induced damage level are slight (less than 1/550), moderate (1/550-1/100), extensive (1/50-1/25) and complete damages (more than 1/25). The fragilities for the RCF at the above four damage states are evaluated by traditional, various ML and NGBoost methods in Fig. 4. The fragility curves are variable under the influence of different methods, but basically keep consistent prediction results, such as slight, moderate and extensive damages. In Fig. 4(d), the large difference of the curve is derived from the assumption that there is a linear relationship between the logarithm of IM and the logarithm of DM in the traditional method. Another reason may be that the performance of the surrogate model is poor due to the very few training samples corresponding to the damage state. In addition, the fragility curves the RCF at the slight, moderate and extensive damages obtained by using NGBoost and traditional methods are compared in Fig. 5. The curves are very close, which indicates that the NGBoost method has good applicability in structural seismic fragility analysis.

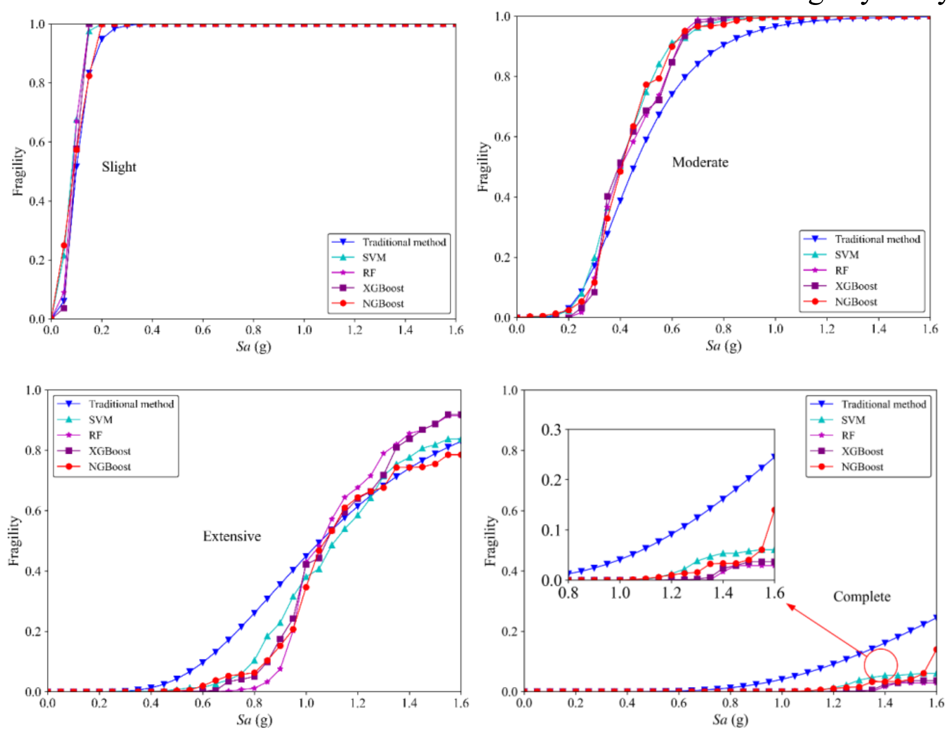

Fig. 4. Fragility curves using various methods for the RCF. (a) Slight damage; (b) Moderate damage; (c) Extensive damage, and (d) Complete damage.

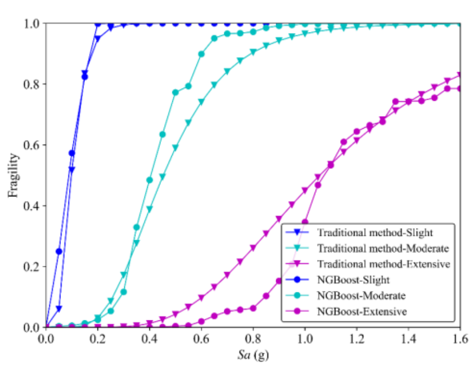

Fig. 5. Fragility curves using the traditional and NGBoost methods for the RCF.

5. CONCLUSIONS

In this study, a probabilistic machine learning method is proposed for seismic fragility analysis. The probabilistic ML method, i.e., NGBoost, can efficiently predict probability distribution, e.g., mean and standard deviation, of damage measures by establishing the mapping of ground motion and structural parameters to structural response. This feature is very convenient for the fragility estimation of structure, compared with traditional ML methods.

A typical structure is analyzed to illustrate and demonstrate the efficiency of the proposed approach. Compared with the traditional ML methods, such as SVM, RF, and XGBoost, the NGBoost method's point estimation accuracy is equivalent to the optimal NGBoost, but significantly better than the SVM. The exceedance probability of the structure under a given IM can be directly calculated by the obtained distribution parameters of the response. The fragility curves drawn are consistent with the traditional methods and the above-mentioned three ML methods. However, it should be pointed out that the fragility assessment process by the proposed method is more simplified.

This proposed framework can be further extended to develop a ground motion IMs selection algorithm for the purpose of assessing the structural system, because one can readily identify the ground motions that significantly affect the variability of the fragility. In addition, the advantages of the NGBoost are more significant in the seismic fragility assessment of regional buildings because it avoids repeated training models for different building.

6. REFERENCES

Breiman, L. (2001). "Random Forests." *Machine Learning*, 45(1), 5-32.

Chen (2016). "XGBoost- A Scalable Tree Boosting System."

Chen, S. Z., Feng, D. C., Wang, W. J., and Taciroglu, E. (2022). "Probabilistic machine-learning methods for performance prediction of structure and infrastructures through natural gradient boosting." *Journal of Structural Engineering*, 148(8), 04022096.

Cherkassky, V. (1997). "The nature of statistical learning theory." *IEEE transactions on neural networks*, 8(6), 1564-1564.

Cornell, C. A., Jalayer, F., Hamburger, R. O., and Foutch, D. A. (2002). "Probabilistic basis for 2000 SAC Federal Emergency Management Agency steel moment frame guidelines." *Journal of Structural Engineering*, 128(4), 526-533.

Duan (2019). "NGBoost- Natural Gradient Boosting for Probabilistic Prediction."

Feng, D. C., Wang, W. J., Mangalathu, S., Hu, G., and Wu, T. (2021). "Implementing ensemble learning methods to predict the shear strength of RC deep beams with/without web reinforcements." *Engineering Structures*, 235, 111979.

Gebetsberger, M., Messner, J. W., Mayr, G. J., and Zeileis, A. (2018). "Estimation methods for nonhomogeneous regression models: minimum continuous ranked probability score versus maximum likelihood." *Monthly Weather Review*, 146(12), 4323-4338.

Guan, X. Q., Burton, H., Shokrabadi, M., and Yi, Z. X. (2021). "Seismic drift demand estimation for steel moment frame buildings: from mechanics-based to data-driven models." *Journal of Structural Engineering*, 147(6), 04021058.

Huang, X., Ye, Y., Xiong, L., Lau, R. Y. K., Jiang, N., and Wang, S. (2016). "Time series k -means: A new K-means type smooth subspace clustering for time series data." *Information Sciences*, 367-368, 1-13.

Kim, T., Song, J., and Kwon, O.-S. (2020). "Probabilistic evaluation of seismic responses using deep learning method." *Structural Safety*, 84, 101913.

Lyu, Y. F., Li, G. Q., and Wang, Y. B. (2020). "Behavior-based resistance model for bearing-type connection in high-strength steels." *Journal of Structural Engineering*, 146(7), 04020109.

Mangalathu, S., and Jeon, J. S. (2018). "Classification of failure mode and prediction of shear strength for reinforced concrete beam-column joints using machine learning techniques." *Engineering Structures*, 160, 85-94.

Mangalathu, S., and Jeon, J. S. (2019). "Stripe-based fragility analysis of multispan concrete bridge classes using machine learning techniques." *Earthquake Engineering & Structural Dynamics*, 48(11), 1238-1255.

McKenna F., Fenves G. L., Scott M. H., et al., Open system for earthquake engineering simulation, University of California, Berkeley, CA.

Pacific Earthquake Engineering Research Center. PEER NGA-West2 Database. Berkeley: University of California; 2013.

Siam, A., Ezzeldin, M., and El-Dakhakhni, W. (2019). "Machine learning algorithms for structural performance classifications and predictions: Application to reinforced masonry shear walls." *Structures*, 22, 252-265.

Vamvatsikos, D., and Cornell, C. A. (2002). "Incremental dynamic analysis." *Earthquake Engineering & Structural Dynamics*, 31(3), 491-514.

Wang, X., Shafieezadeh, A., and Ye, A. (2017). "Optimal intensity measures for probabilistic seismic demand modeling of extended pile-shaft-supported bridges in liquefied and laterally spreading ground." *Bulletin of Earthquake Engineering*, 16(1), 229-257.

Zhang, R., Chen, Z., Chen, S., Zheng, J., Büyüköztürk, O., and Sun, H. (2019). "Deep long short-term memory networks for nonlinear structural seismic response prediction." *Computers & Structures*, 220, 55-68..