Structural Safety 115 (2025) 102596

Contents lists available at ScienceDirect

Structural Safety

journal homepage: www.elsevier.com/locate/strusafe

Reliability-based quantification of the benefits of machine learning predictive models in seismic structural design and performance assessment

Muneera A. Aladsani *, Henry V. Burton

Department of Civil and Environmental Engineering, University of California Los Angeles, CA 90095, USA

ARTICLE INFO

Keywords:
Reliability-based assessment
Machine learning
Predictive model accuracy
Epistemic uncertainty
Structural reliability

ABSTRACT

Machine Learning (ML) techniques have been used extensively in research within the field of structural engineering due to their high level of accuracy in predicting the behavior of different structural elements. In fact, the superior predictive performance relative to traditional statistical models is often suggested as the primary motivation for the adoption of ML models. However, the implications of such improvements in predictive accuracy in the design and performance of structural systems have not been studied. This paper presents a reliability-based investigation of the tangible benefits provided by ML models in terms of structural design and performance. To quantify these benefits, the increase in predictive accuracy is interpreted as a reduction in epistemic uncertainty. The specific focus is on a predictive model that estimates the drift capacity of reinforced concrete shear walls (RCSWs) with special boundary elements. The accuracy of an extreme gradient boosting (XGBoost) model relative to a basic linear regression equation is quantified in terms of reduced epistemic uncertainty. Then, using 36 RCSW archetype buildings, a Monte Carlo-based procedure is implemented to evaluate the implication of the improved predictive accuracy to seismic design and performance. The study provides insights into how much improvement in accuracy (i.e., ML relative to traditional model) is needed to have a tangible effect on the seismic design and performance.

1. Introduction

1.1. Machine learning in structural engineering

Research on the application of machine learning (ML) in structural engineering (SE) has received significant attention in recent years. The potential of ML to capture complex nonlinear relationships between the different variables contained in the data coupled with the increased accuracy (relative to traditional statistical models) is one of the most cited reasons for their widespread use. Several studies showed the superiority of ML models in accurately predicting the behavior of different types of structural systems and components. Backbone curve parameters [1], deformation capacity [2,3], bearing capacity [4], and failure mode [4,5] of reinforced concrete (RC) columns have all been predicted using ML models. In addition, ML models for RC beams have been developed to estimate shear strength [6,7] and long-term deflection [8]. ML has also been used to predict shear strength [9,10], drift capacity [11], and failure modes [12] of RC walls. The behavior of RC slabs [8,13,14] and beam-column joints [15,16,17] has also been investigated using ML algorithms. Based on this limited review, there were several key findings that can be generalized across the various studies. For example, the improvement in predictive performance by the ML model relative to more traditional methods (e.g., analytical equations) was almost always highlighted. Most studies conducted a feature importance analysis, which elucidates the factors that have the greatest effect on predictive performance. Most (if not all) of the studies limited their evaluation of the predictive models to the feature ranges contained in the training data. This approach brings into question the ability (or lack thereof) of the ML models to generalize (or extrapolate) beyond the observed data. Readers are directed to several review papers that have summarized the application of ML to structural engineering problems [18–21].

Despite their popularity, ML models are faced with the challenge of being difficult to explain i.e., they are generally viewed as “black-box” models. Explainable Artificial Intelligence (XAI) [22], a branch of study that concentrates research on ML interpretability and seeks to move toward a more transparent AI, is being developed in an effort to overcome this problem. Interpretability methods can be classified based on their purpose (e.g., explaining black-box models, creating white-box

* Corresponding author.
E-mail address: maladsani@ucla.edu (M.A. Aladsani).

https://doi.org/10.1016/j.strusafe.2025.102596
Received 15 March 2024; Received in revised form 16 March 2025; Accepted 17 March 2025
Available online 19 March 2025

models, enhancing fairness of a model, and testing the sensitivity of predictions). These methods can be further divided based on the type of ML algorithms that are considered. If the method is limited to a single model or a group of models, it is described as model-specific. On the other hand, if the method can be used with any ML model, it is deemed model-agnostic. The scale of interpretation is another essential factor that distinguishes different interpretability techniques. The technique is considered local if it only describes one particular instance, and it is considered a global method if it explains the entire model. Lastly, some XAI methods are more suitable to specific data types (e.g., tabular, images, and text) [23].

As previously indicated, the limitations of conventional data driven methods in terms of predictive accuracy are one of the main drivers of ML-SE applications. A review of ML-based regression models of different structural elements that were compared with other empirical models (e.g., linear regression, from a design code, other non-ML models reported in literature) was conducted to determine the improvements based on different evaluation metrics including root mean square error (*RMSE*), coefficient of determination (R^2), mean absolute error (*MAE*), mean absolute percentage error (*MAPE*), and the coefficient of variation (*COV*) of the ratio of predicted-to-experimental target variable. Table 1 provides a comprehensive summary of the percentage improvements in *RMSE*, R^2, *MAE*, *MAPE*, and *COV* achieved through the implementation of ML models [2–4,7,8,10,11,14,16,24–58].

An important observation from Table 1 is that the relative improvement in predictive performance varies significantly based on the considered metric and the type of component. For instance, when predicting wall shear strength, the average *COV* improvement is approximately 38 %, which is significantly lower compared to the other metrics (R^2, *RMSE*, *MAE*, and, *MAPE*). Whereas for column axial strength, the average *COV* improvement is greater than the other metrics. There is also significant variance in the improvement as measured by the different metrics for beam deflection and rotation capacity. A similar conclusion can be drawn when examining the improvements provided when predicting the flexural and torsional strength of beams. Based on these observations, it is important to recognize that while some of these improvements may appear significant, this may be partly due to the chosen predictive performance metric. The studies that form the basis of the results in Table 1 have exclusively focused on improving the predictive power of their proposed models without discussing the tangible benefits of these improvements. This is particularly important because, for most ML models, there exists a trade-off between predictive accuracy, model complexity, and interpretability. While enhanced predictive accuracy is a valuable outcome, it is essential to explore how these improvements translate into practical benefits in engineering practice. Questions remain about whether increased accuracy leads to more informed design choices, enhances safety, or provides cost savings in real-world applications. This paper aims to bridge this gap by investigating the practical implications of improved ML predictive models in the design and performance of structural systems within a reliability-based context.

Table 1
Percentage improvements in evaluation metrics for various structural element responses.

Structural Elements' Responses		Evaluation Metric				
		RMSE	R^2	*MAE*	*MAPE*	*COV*
Shear Strength	Beam [7,24–34]	32.8 % to 80.5 %	0.4 % to 99.4 %	35.3 % to 81.3 %	80.3 %	12.1 % to 84.2 %
	Slab [14,35–38]	54.7 % to 77.8 %	3.9 % to 21.6 %	–	32.3 % to 38.6 %	52.2 % to 67.1 %
	Column [4,39]	48.4 % to 62.6 %	31.4 % to 261.4 %	71.7 %	78.6 %	67.6 %
	Wall [10,40–42]	41.53 % to 80.87 %	6.5 % to 60.4 %	66.7 %	59.8 % to 79.8 %	38.2 %
	Joint [16,43–46]	−2.9 % to 93.7 %	66.7 % to 120.3 %	63.2 % to 99.9 %	46.9 % to 98.9 %	0.2 % to 90.7 %
Axial Strength	Column [47–53]	27.3 % to 90.2 %	0.4 % to 26.8 %	57.2 % to 77.9 %	57.1 % to 89.3 %	75.9 % to 90.4 %
Flexural and Torsional Strength	Beam [54–57]	18.6 % to 90.1 %	−0.4 % to 14.9 %	−129.6 %	–	21.2 % to 48.3 %
Deflection and rotational capacity	Beam [8,58]	73.9 %	0.2 %	−0.5 %	0.9 % to 70.5 %	–
Drift capacity	Column [2,3]	74.2 %	207.3 %	–	83.7 %	46.7 % to 61.0 %
	Wall [11]	41.0 %	38.1 %	–	–	43.3 %

In other fields where ML has been adopted in practice, there have been studies that have evaluated their benefits in practical terms. Reis et al. [59] showed that the extent to which ML applications are used to support internal organizational activities has a positive effect on financial performance. Additionally, Jannach and Jugovac [60] showed that ML-based recommender systems such as the ones used in e-commerce websites, have a positive effect on various business performance metrics including sales numbers and user engagement. Furthermore, Google reported that their App recommender system—a combination of neural networks and linear models—resulted in 3.9 % increase in app acquisitions for Google Play compared to a linear model [61]. In SE, where safety and performance of infrastructure are the primary goals, the implications of improving the accuracy of predictive models can be investigated using reliability-based principles.

1.2. Structural reliability

As a sub-discipline, structural reliability provides the language, tools, and models needed to evaluate and quantify the safety and overall performance of structures by explicitly accounting for various sources of uncertainty (e.g., [62,63]). In general, uncertainties can be categorized as aleatory or epistemic [64]. Aleatory uncertainty refers to the inherent randomness in a system that is irreducible. On the other hand, epistemic uncertainty represents the uncertainty due to the lack of knowledge which can be reduced if additional information is available. In seismic performance assessments, ground motion (record-to-record) and structural model (modeling) uncertainties are often considered. Record-to-record (RTR) uncertainty, which is a type of aleatory uncertainty, arises from the unique frequency characteristics of each ground motion. As a result, there is variance in the response of a structure subjected to a set of ground motions scaled to the same intensity. This uncertainty is typically considered in structural response estimation by utilizing a large number of records.

Modeling uncertainty is a form of epistemic uncertainty that can be classified into four categories [65]. Type I refers to uncertainties in the measurement of physical quantities (e.g., soil shear stiffness which can be directly measured). Since not all constitutive models use physical parameters that can be directly measured, Type II uncertainties consider the difference between the measured quantities and constitutive model parameters. Type III is used to describe the uncertainty associated with the selected constitutive model, which is typically theoretically derived

based on various assumptions (e.g., [66,67]) or empirically constructed based on observations (e.g., [68,69]). Finally, Type IV uncertainty is related to system level idealization (e.g., one- versus two-dimension analysis, neglecting soil-structure interaction).

Evaluating the effects of "low-level" uncertainties (Type I and Type II) is more common in prior research (e.g., [70–75]), whereas fewer studies have attempted to quantify the effects of "high-level" uncertainties (Type III and Type IV) (e.g., [76–78]). This could be because the data needed to quantify low-level uncertainty is generated by component-level physical experiments which are fairly common. In contrast, high-level uncertainty characterization requires data from subsystem or system-level experiments. High-level uncertainties are also known to be problem and system specific and therefore less generalizable than low-level uncertainties.

The reliability of a structural component can be defined as the probability that it meets one or more predefined requirements (e.g., strength capacity is not exceeded) over some specified time period. Structural component reliability is expressed in terms of a limit state function, g, which relates the resistance and load variables for specific state of interest. This limit state function is given by $g = R - Q$, where R is the resistance and Q is the load. In a general problem, R and Q may be functions of multiple basic random variables $\boldsymbol{X} = (X_1, X_2, \cdots, X_n)$. In this case, the limit state function is defined as $g(\boldsymbol{X}) = f(X_1, X_2, \cdots, X_n)$, where $f(.)$ is a function that relates R and Q with the basic random variables. Member failure occurs in the region where g is less than or equal to zero ($g(\boldsymbol{X}) \leq 0$). The probability of member failure is expressed as:

$$P_f = \mathbb{P}[g(\boldsymbol{X}) \leq 0] = \int_D f_{\boldsymbol{X}}(x_1, x_2, \cdots, x_n) dx_1 dx_2 \cdots dx_n \quad (1)$$

Where $f_{\boldsymbol{X}}(x)$ is the joint probability density function of $\mathbf{X}$ (the basic random variables) and D is the domain in which $g(\boldsymbol{X}) \leq 0$.

To address the difficulties in evaluating the probability of failure analytically, various approximate methods are often used such as first order reliability method (FORM), second order reliability method (SORM), and Monte Carlo simulation (MCS). In a design context, once the probability of failure has been established, an appropriate set of modifications to the system or component is made to enhance the reliability which in turn minimizes the risk of failure.

1.3. Objective

Despite the numerous SE studies that demonstrated the superior predictive performance of ML models, the implications of this improved accuracy, especially in a reliability-focused context, have not been addressed. To fill this research gap, this study presents a reliability-based investigation of the benefits provided by ML models in terms of their ability to improve predictive performance. Specifically, the enhanced predictive performance is quantified as a reduction in the epistemic uncertainty that is associated with the model. Ultimately, the goal is to investigate and quantify the effect of this epistemic uncertainty reduction on the "failure" probability of the considered system or component. The investigation focuses on the drift capacity of RC shear walls (RCSWs) with special boundary elements (SBEs). Section 2 provides an overview of the investigation methodology (Section 2.1) with a detailed explanation of each stage in the subsequent subsections. Section 2.2 provides a description of the ML-based drift capacity predictive model for RCSW-SBE that is used as the basis of the investigation. In Section 2.3, the RCSW-SBE automated design procedure is described. Section 2.4 discusses the reliability-based assessment that is performed using MCS. Section 3 performs a case study that considers 36 RCSW-SBE archetypes. The results, which are documented in Section 4, are separated based on performance assessment (Section 4.1) and design implications (Section 4.2). Section 5 summarizes the key findings, limitations, and suggestions for future related work.

2. Methodology

2.1. Overview of the methodology

The methodology for investigating the implications of improved predictive performance provided by ML models is illustrated in Fig. 1. The specific goal is to evaluate the effect of the reduction in epistemic uncertainty (due to increased predictive accuracy) that is provided by the ML models on seismic structural design and performance. The current study is focused on RCSWs with SBEs. However, the methodology is presented as being agnostic to the type of structural system and/or component.

The context of the proposed framework is the development of the drift capacity code requirements for RCSW-SBE structures. As such, the design implication is quantified in terms of the required wall thickness that is needed to meet the target reliability objective. With this frame of reference, the methodology begins with defining a set of archetype buildings (RCSWs in this case) that will be used in the evaluation. The archetype space used in the current study considers variations in the number of stories, aspect ratio, and seismic design category. Since the proposed context is code development (and not on a single building), the goal is to generalize the findings across a broad class of RCSW-SBE buildings. The target performance is specified as the maximum acceptable probability of exceeding the limit state of interest, which in this study, is the drift capacity of the RCSW-SBE.

For each archetype building, the lateral force resisting system is designed to satisfy the ACI 318–19 [80] provisions for special structural walls. For the reliability-based assessment, the limit state function is $g = R - D$, where D is the drift demand and R is the drift capacity. The drift capacity predictive model is used with j different COV values. Each COV represents a different level of uncertainty in the predictions. Dispersion in drift demand due to RTR variability is quantified by performing nonlinear response history analyses (NRHA). Dispersion in drift demand due to modeling uncertainty is also considered in the assessment. MCS is implemented to determine the probability of failure (P_f), which arises when the drift demand exceeds the drift capacity. For the reliability-based design, the archetypes are redesigned by incrementally increasing the wall thickness until the target probability of failure ($P_{f,t}$) is met. The effect of different drift capacity COV values on the reliability-based wall designs are compared. This comparison is made based on the wall thickness for each design that is required to meet the target reliability.

The wall thickness is a primary structural design variable due to its impact on both architectural and economic factors. Reducing wall thickness can provide benefits in space-sensitive projects. Thinner walls enhance aesthetics by allowing for more open spaces and more usable floor area. Additionally, reducing the wall thickness can lead to cost savings by decreasing material usage, lowering foundation requirements, and simplifying cladding installation, which further reduces labor costs. Overall, even small adjustments in wall thickness can impact both functionality and aesthetics. The extent to which these factors are influenced by changes in thickness depends on several variables, including the initial wall thickness, the building's intended use, the number of stories, and the specific functional requirements of the building.

2.2. Drift capacity model

As stated in the objective, this study aims to quantify the benefits of reduced uncertainty provided by ML models. The ML model considered in this study is the drift capacity prediction model for special structural walls based on the Extreme Gradient Boosting (XGBoost) algorithm [11]. The algorithm optimizes an objective function that combines loss and regularization terms:

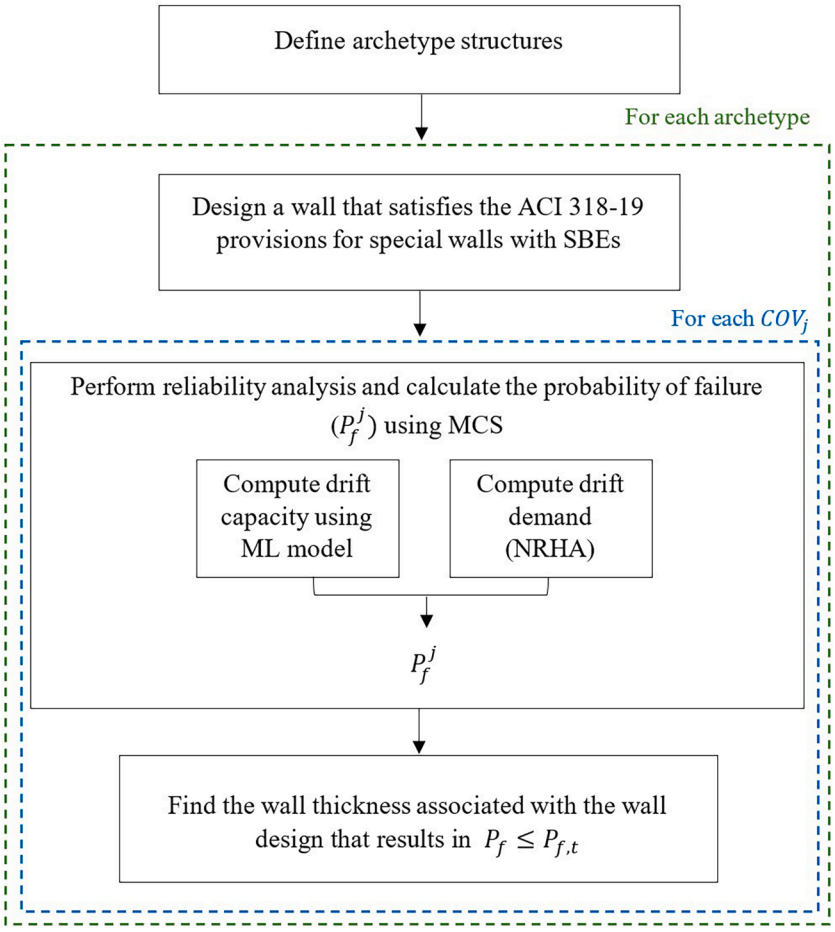

Fig. 1. Overview of the proposed methodology.

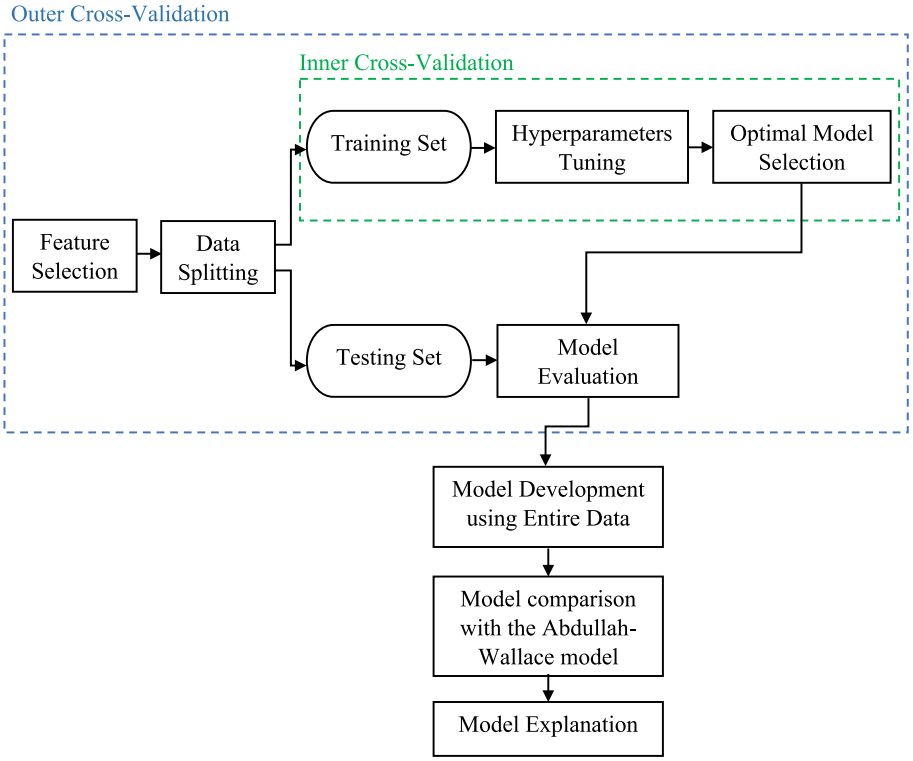

Fig. 2. Overview of the methodology used to develop the ML model.

$$L = \sum_i^n l(y_i, \hat{y}_i) + \sum_{k=1}^K \Omega(f_k) \quad (2)$$

Where $\sum_i^n l(y_i, \hat{y}_i)$ is the summation of the loss function value over the n datapoints in the training set, $\Omega(f_k)$ is the regularization term, and K is the number of trees. Decision trees sequentially constructed, with each correcting the errors of the previous ones. First- and second-order derivatives (gradients and Hessians) are used to minimize loss, and a greedy approach is used to identify optimal splits. Regularization is incorporated to prevent overfitting.

Fig. 2 depicts a schematic overview of the methodology used to develop and evaluate the model. The first step (feature selection) investigates the variables that affect wall drift capacity using statistical techniques, and the most influential features are chosen as inputs. The features included are (1) λ_b ($= l_w c/b^2$), where l_w is the wall length, c is the depth of the compression zone and b is the wall thickness. The λ_b parameter accounts for the slenderness of the cross section and the compression zone, (2) shear stress demand $\left(v_{max}/\sqrt{f'_c}\right)$, where v_{max} is the maximum shear stress and f'_c is the concrete compressive strength and (3) the configuration of boundary transverse reinforcement (CBTR), which is categorized into five types: overlapping hoops (OH), combination of a perimeter hoop and crossties with 90–135 degrees hooks (PH-90–135), combination of a perimeter hoop and crossties with 135–135 degrees hooks (PH-135–135), combination of a perimeter hoop and crossties with headed bars (PH-HB), and single hoop without intermediate legs of crossties (SH). Histograms for the input features and the drift capacity of the dataset are shown in Fig. 3. The data is then split using nested cross-validation where the inner layer is used to tune the hyperparameters, while the outer layer is used to evaluate the model performance.

After creating and independently evaluating the proposed model, its predictive performance is compared with that of the Abdullah and Wallace model [81,82]. This comparison is made by developing two distinct models based on the CBTR. These two models were developed solely for the purpose of an "all-else-equal" comparison with the Abdullah and Wallace model. Subsequently, the original XGBoost model, in which the CBTR is included as an input feature, was compared with the results obtained from the two distinct models. The comparison showed that the original model produced similar results to the separate models. The models were constructed using the same dataset as that used by Abdullah and Wallace [81,82].

The predictive performance of the XGBoost model, which includes the CBTR as an input feature, and the empirical equation developed by Abdullah and Wallace, were comparatively evaluated based on their accuracy. The *COV* of experimental-to-predicted drift capacity ratio resulting from the XGBoost and Abdullah and Wallace models is 0.085 and 0.15, respectively (an improvement of 43.3 %). Furthermore, the proposed model improved the *RMSE* by 41 % and R^2 by 38 %. Fig. 4 shows the experimental versus predicted drift capacity of the RC walls using the XGBoost model. Additional details regarding the development and performance evaluation of the drift capacity ML model can be found in Aladsani et al. [11].

In Fig. 5, the error—defined as the difference between experimental and predicted drift capacity—is plotted against each feature in the XGBoost model. Higher errors are primarily found at lower values of λ_b and $v_{max}/\sqrt{f'_c}$. Also, higher errors are observed in walls with OH, PH-90–135, and SH. Data points with absolute errors exceeding 0.5 % were further analyzed, revealing that overestimation of the observed value often occurred when the wall features produced higher experimental drift capacities. On the other hand, underestimation of the observed value mostly occurred in walls features that exhibit lower capacities. This suggests that additional factors, beyond the adopted input features, may influence the variations in the measured drift capacities.

These findings highlight the trade-offs involved in drift capacity

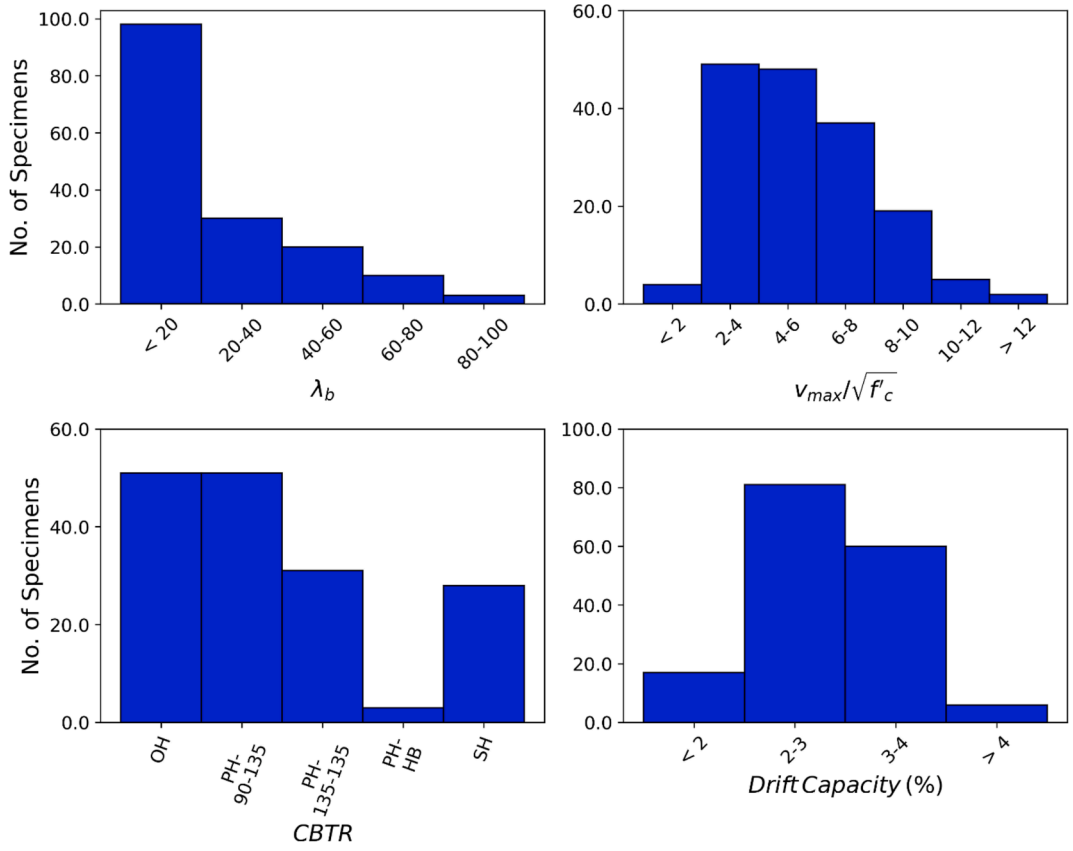

Fig. 3. Histograms of the input variables and drift capacity from the dataset.

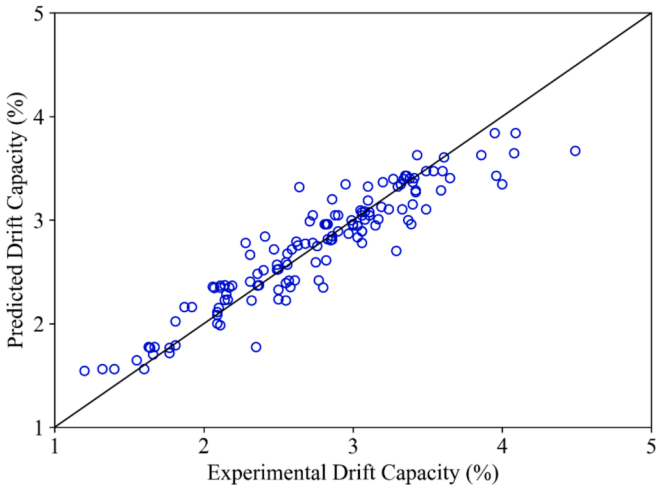

Fig. 4. Actual versus predicted drift capacity for the XGBoost model.

predictions. Overestimating the drift capacity can lead to designs with greater safety margins, potentially preventing structural failures. However, it may also result in unnecessary material usage and increased construction costs. Conversely, consistent underestimation can yield unsafe designs that fail to meet essential performance criteria, thereby elevating the risk of structural failure.

Given these considerations, it is crucial to prioritize the enhancement of the drift capacity prediction model by focusing on ranges that are near established failure thresholds. Specifically, attention should be directed toward drift capacities approaching the limits of acceptable performance, as these represent critical points where structural integrity is crucial. By identifying archetypes with drift capacities near these thresholds, it becomes possible to pinpoint the values most likely to influence failure probabilities. This targeted approach ensures that model improvements are concentrated where they can have the most significant impact on safety and overall structural performance.

2.3. Wall design

Because a large number of designs and analyses for the archetype space are required, a Python-based platform is developed based on the object-oriented programing paradigm and used to automate the design and NRHAs. Dubbed RCWall-SDA (Reinforced Concrete Wall Seismic Design and Analysis), the tool helped minimize the time and manual effort required to generate code-compliant wall designs and analyses.

The overall framework of the RCWall-SDA automated design process is shown in Fig. 6. The initial input parameters specified by the user include building arguments (floor plan dimensions, number of stories, story heights, number of walls, and wall location), loads (floor dead and live loads), site conditions (mapped spectral response acceleration parameter at a period of 1-s (S_1) and at short periods (S_s), site class, and long-period transition period (T_L)), concrete and steel parameters (concrete compressive strength (f'_c), concrete type, reinforcement yield strength (f_y), and areas of reinforcement bars), initial wall cross section dimensions, configuration of the boundary transverse reinforcement, and other miscellaneous parameters (redundancy factor, response modification coefficient, importance factor, and deflection amplification factor).

The design process starts with calculating the design spectral acceleration parameters at the short period (S_{DS}) and at 1 s period (S_{D1}) using the AccParam class. Then, the ELF class implements the ASCE 7–16 [79] equivalent lateral force (ELF) procedure and the vertical distribution of seismic forces along the height of the structure is determined based on the approximate fundamental period. The RCWall-SDA also includes a response spectrum analysis (RSA) class which can be used as an alternative to the ELF procedure. However, the ELF procedure is used in the current study. The ElasticAnaylsis class constructs an OpenSees [83] model of the wall and performs a linear elastic analysis to obtain the story drifts. The drifts are then scaled using the deflection amplification and importance factors, and compared with the drift limit specified in Table 12.12-1 of ASCE 7–16.

Using the outputs from the aforementioned three classes, the WallDesign class performs the design and analysis calculations specified in section 18.10 of ACI 318–19 to obtain an initial wall design. The PMInteraction class generates the axial-flexure interaction diagram, which is used to ensure that the combined axial and moment demands (P_u, M_u) for all load combinations do not exceed the axial and moment capacities ($\phi P_n, \phi M_n$). The need for special boundary elements at the ends of the structural wall can be evaluated using either a displacement-based design or stress-based design approach. The displacement-based approach can only be used for effectively continuous slender walls with a single critical section. Whereas the stress-based approach can be applied to any wall configuration. The RCWall-SDA tool implements the displacement-based approach to assess the need for SBEs. To optimize the design, the quantities of longitudinal and transverse reinforcement are reduced over the wall height (every two stories) as the demands decrease. To facilitate the automated design process, the total number of longitudinal reinforcement bars are chosen such that all layers have the same number of bars.

For all archetypes, the initial wall thickness is set to 12 *inch*. The design is evaluated to determine if it satisfies the provisions of ASCE 7–16 and ACI 318–19 for Special Structural Walls. If not, the thickness is increased by 1 *inch* and the design process is repeated until an adequate design is accomplished.

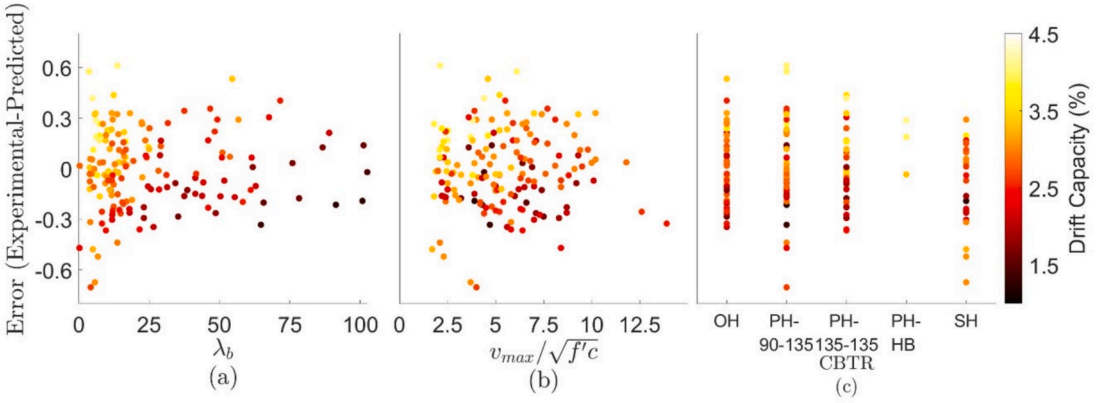

Fig. 5. Variation of the error relative to each feature in the XGBoost model: (a) λ_b, (b) $v_{max}/\sqrt{f'_c}$, and (c) configuration of boundary transverse reinforcement.

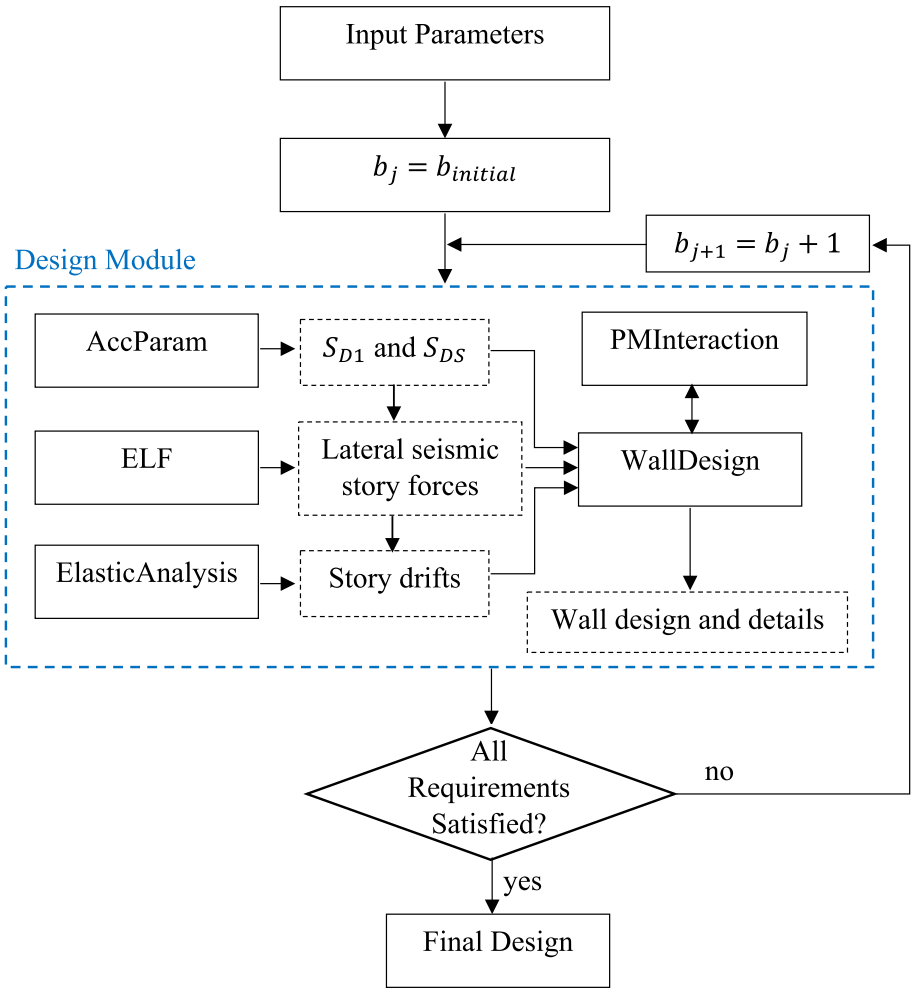

Fig. 6. Framework of the automated design process.

3. Case study

A case study is presented to quantify the implications of the reduced epistemic uncertainty in the drift capacity prediction model to the design and performance assessment of RCSW-SBEs. The RCWall-SDA tool is used to design a predefined set of archetypes and perform NRHA to determine the probability distribution of the drift demand. Using the relevant wall design variables, the XGBoost model is then used to estimate the drift capacity. By performing MCS, the performance and design of the RCSWs considering different drift capacity prediction accuracies are evaluated.

In the context of reliability-based design of RCSWs with SBEs, Abdullah and Wallace [82] introduced a drift demand-to-capacity ratio check based on the following equation: $\phi_d\left(\frac{\bar{\delta}_c}{h_w}\right) \geq \left(1.5\frac{\bar{\delta}_u}{h_w}\right)$, where ϕ_d is a reduction factor, $\left(\frac{\bar{\delta}_c}{h_w}\right)$ is the mean wall lateral drift capacity estimated using their predictive equation developed in (Abdullah and Wallace 2019) [81], and $\left(\frac{\bar{\delta}_u}{h_w}\right)$ is the mean roof drift demand. The 1.5 multiplier is used to convert the design earthquake (DE) drift demands to maximum considered earthquake (MCE) demands. The authors suggested using a 10 % probability of strength loss for DE-level shaking as a minimum criterion for collapse. Then, by assuming lognormal distributions and *COV* of 0.30 and 0.15 for the drift demand and capacity, respectively, a $\phi_d = 1$ is selected.

3.1. Archetypes

The floor plan selected for the archetypes (as shown in Fig. 7) is 150 ft x 75 ft with typical bay spans of 25 ft and story heights of 12 ft [82]. The developed archetypes have 2, 4, 8, and 12 stories. For a given number of stories, wall aspect ratios ranging from 2 to 4 are considered. The minimum wall length is taken to be 6 ft and only walls with length/thickness ratio $\left(\frac{l_w}{b}\right)$ greater than 6 are included. The walls are designed for the maximum and minimum spectral acceleration intensities (D_{max} and D_{min}) associated with Seismic Design Category (SDC) *D*. Considering the aforementioned variations, a total of 36 archetypes were developed

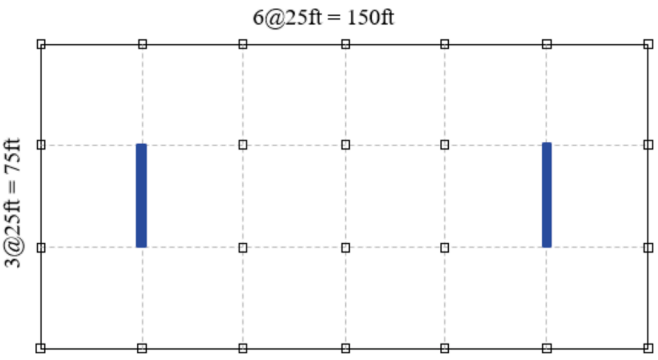

Fig. 7. Floor plan.

(Table 2).

Assuming residential occupancy for all buildings, risk category II is used, resulting in an importance factor $I_e = 1.0$. The response modification factor $R = 5$ and the drift amplification factor $C_d = 5$. The typical floor dead load is specified as 125 psf and 40 psf live load is considered. The design concrete compression strength is taken to be 6 ksi and the reinforcement yield stress is 60 ksi.

3.2. Ground motion selection

Two sites, one in Los Angeles (33.7917, −118.1927) and the other in Fresno (36.7357, −119.6784), are used for the ground motion selection, which correspond to the D_{max} and D_{min} SDCs, respectively. The sites were selected so that the spectral acceleration at short periods (S_S) and at a period of one second (S_1) were consistent with those specified in FEMA P-695 [84].

For each site, seismic hazard deaggregation is performed using the Unified Hazard Tool provided by the United States Geological Survey (USGS) [85] to obtain the expected characteristics (mean magnitude, source-to-site distance, and epsilon) of the ground motions for the DE (10 % probability of exceedance in 50 years) hazard level. With the goal of using the same set of ground motions for all archetypes associated with a given SDC, the deaggregation is conducted at the mean ASCE 7–16 fundamental period of the buildings (0.7 s). Once the target spectrum and its standard deviation are calculated using the Campbell and Bozorgnia model [86], a set of 40 ground motions is selected from the NGA-West2 database and scaled such that the mean spectra and standard deviation follow those of the target.

Fig. 8a and 8b show the individual and mean spectra of the selected ground motions along with the target spectrum and the 95 % confidence intervals for SDC D_{max} and D_{min}, respectively. The standard deviations of the spectral values for the target and selected records for SDC D_{max} and D_{min} are compared in Fig. 9a and 9b, respectively.

Table 2
Shear wall archetypes.

Archetype ID	No. of Stories	SDC	Aspect Ratio
1	2	D_{max}	2
2			2.5
3	4		2
4			2.5
5			3
6			3.5
7	8		2
8			2.5
9			3
10			3.5
11			4
12	12		2
13			2.5
14			3
15			3.5
16			4
17	2	D_{min}	2
18			2.5
19			3
20			3.5
21			4
22	4		2
23			2.5
24			3
25			3.5
26			4
27	8		2
28			2.5
29			3
30			3.5
31			4
32	12		2
33			2.5
34			3
35			3.5
36			4

3.3. Nonlinear modeling

A numerical model of each archetype is generated in OpenSees. Considering the symmetry of the archetypal structure, a two-dimensional model is created consisting of one wall. An illustration of the model is presented in Fig. 10. The seismic mass is lumped at the wall nodes at each story level, and gravity loads are assigned according to the corresponding tributary areas at the same nodes. P-delta effects that consider loads from the gravity system, which is not explicitly modeled, are considered by including a leaning column. A linear elastic bar with zero lateral stiffness is used to represent the leaning column. Horizontal and axially rigid truss elements are used to connect the leaning column to the wall. Rayleigh damping at 2 % is specified at the first and third modes of vibration.

The behavior of the wall is simulated using the Shear-Flexural Interaction Multi-Vertical Line Element Model (SFI-MVLEM) element [87,88]. Each wall is modeled using two equal-length elements per story along the building height, as shown in Fig. 10. Under cyclic loads, the SFI-MVLEM captures the interaction between axial/flexural and shear responses in RC structural walls. The Fixed-Strut-Angle-Model (FSAM) is used to describe a two-dimensional constitutive RC panel behavior at each macro-fiber of the Multiple-Vertical Line-Element-Model (MVLEM) in the SFI-MVLEM element. Eight macro-fibers are used to represent the wall cross section in the horizontal direction, with four outer macro-fibers representing the confined wall boundaries and the other four representing the unconfined wall web.

The uniaxial hysteretic constitutive model for concrete (Concrete02 in OpenSees) proposed by Yassin [90] is used to simulate the hysteretic behavior of confined and unconfined concrete. The Concrete02 material model was chosen because it is computationally efficient, numerically robust, and comparatively simple. It also captures important concrete behavioral characteristics such as tension stiffening and hysteretic stiffness degradation.

To simulate the behavior of the reinforcement, the uniaxial constitutive nonlinear hysteretic material model for steel (SteelMPF in OpenSees) proposed by Menegotto and Pinto [91] and extended by Filippou et al. [92] is used. The SteelMPF material model includes isotropic strain hardening effects. The material model also permits different yield stresses and strain hardening ratios for compression and tension, and considers the degradation of the cyclic curvature parameter R for strain reversals in both pre- and post-yielding regions of the hysteretic stress–strain behavior for better yield capacity predictions. It is worth noting that the stress overshooting effect upon partial unloading has been corrected in SteelMPF.

3.4. Uncertainty characterization for structural model Parameters

As previously stated, the objective of this study is to investigate the ramifications of the reduction in epistemic uncertainty in the RCSW drift capacity prediction, in terms of reliability. To account for the uncertainty in the predictive model, an error factor (experimental-to-predicted ratio) is used. The error factor is a scaling factor that is applied to the estimated capacity to incorporate the potential deviation between the model prediction and the actual capacity. A histogram of the error associated with the XGBoost model is shown in Fig. 11 along with the density curve of normal distribution having mean and standard error values computed from the empirical data. The Shapiro-Wilk test is conducted to assess the normality of the error factor. The null hypothesis for the Shapiro-Wilk test is that the data is normally distributed, therefore, if the p-value associated with the test is greater than a predefined significance level (e.g., 0.05), the data can be considered normally distributed. The p-value obtained = 0.65, therefore, the model error is assumed to follow a normal distribution with mean = 0.999 and *COV*

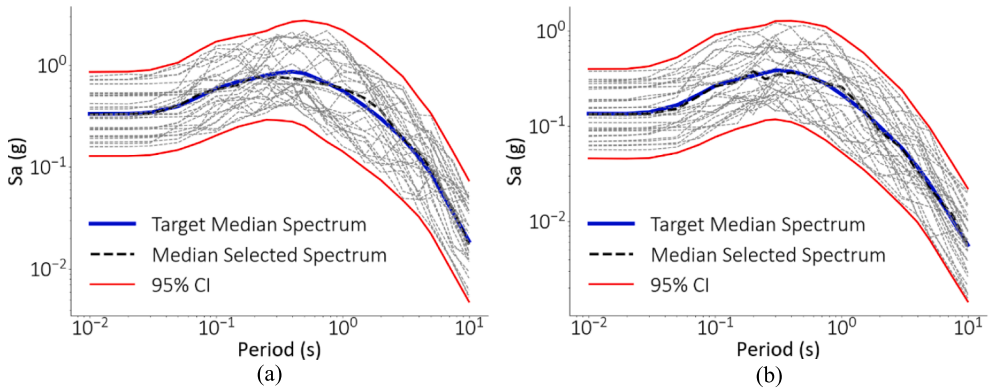

Fig. 8. Spectra for selected suites of ground motions and their respective targets for a)D_{max} and b)D_{min}

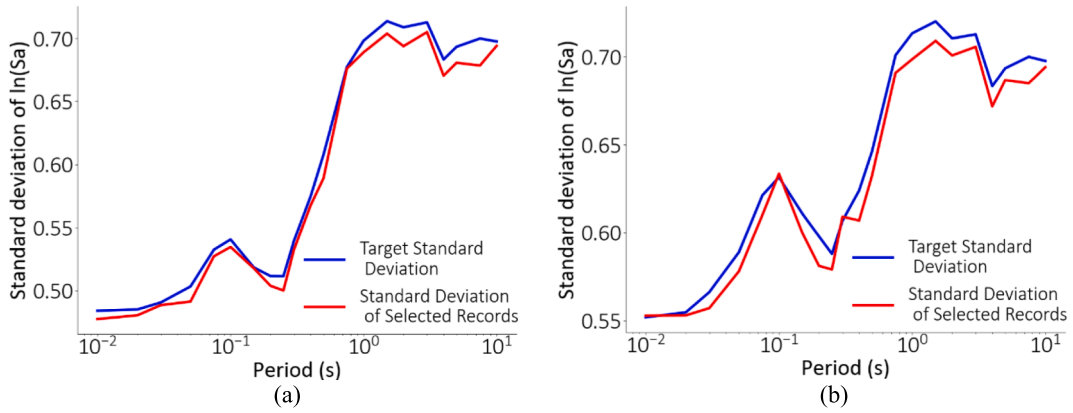

Fig. 9. Standard deviations of the spectral values for the target and selected ground motions for a)D_{max} and b)D_{min}

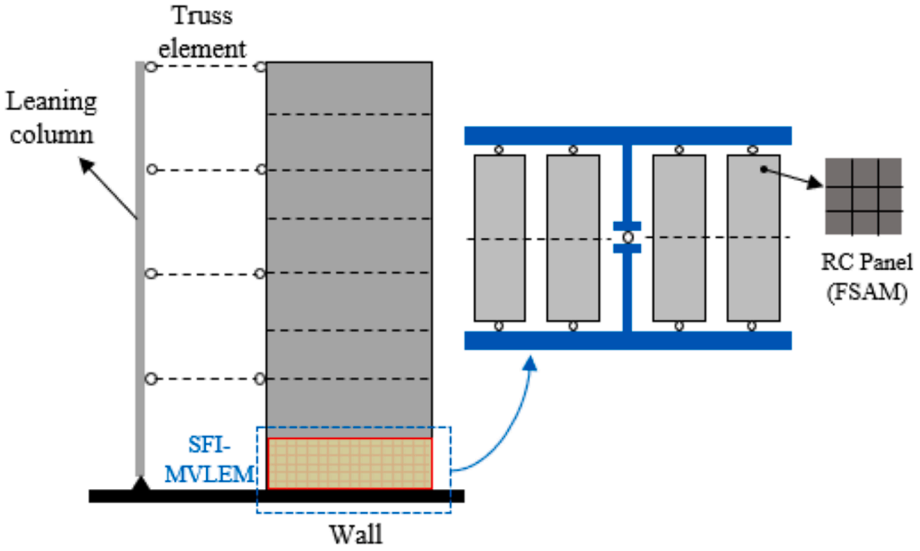

Fig. 10. Elevation view of the structure and SFI-MVLEM element (adapted from Kolozvari et al. [89])

values = {0.085,0.15,0.2,0.25,0.3}. The 0.085 and 0.15 correspond to the XGBoost and linear models, respectively, whereas the other values are included to evaluate the effect of more significant accuracy improvements (e.g., through further enhancements of the ML model). For context, the improvement in *COV* resulting from the XGBoost model is 43.3 %, which falls within the middle of the range reported in Table 1 (0.22 % – 90.7 %). Additionally, the highest improvement considered (from 0.3 to 0.085) is 71.7 %, which also falls within that range. The 43.3 % improvement in *COV* corresponds to a 38 % improvement in R^2 and a 41 % improvement in *RMSE*.

The drift demand can be obtained using either linear static analysis (e.g., ELF) (with the appropriate amplification factors) or NRHA. Since NRHA is the more reliable of the two, it is used in this paper. First, for simplicity and to reduce the computational expense, the drift demand

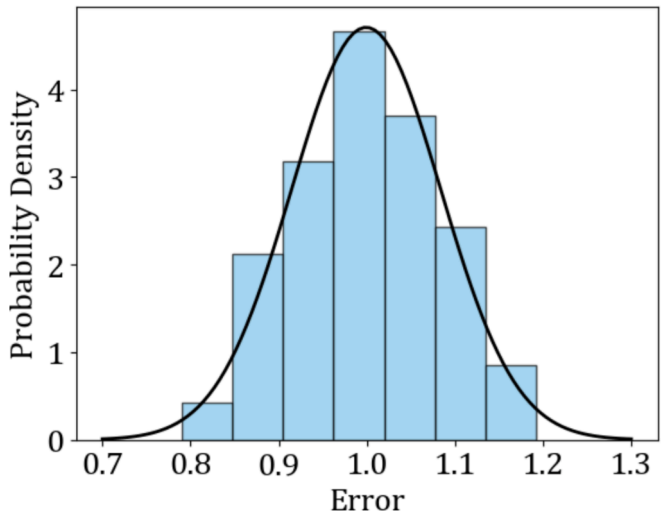

Fig. 11. Histogram showing the distribution of the XGBoost model prediction error.

probability distribution is computed by considering only RTR uncertainty. For each archetype, the drift demand is assumed to follow a lognormal distribution with the dispersion resulting from NRHAs using the 40 ground motions. Next, both RTR and modeling uncertainties are considered by using results from Kim [93] that reports RCSW roof drift dispersion values at DE hazard level. In Kim [93], the total *COV* of roof drift which incorporates RTR and model parameter/design uncertainties (includes concrete compressive strengths, reinforcing steel yield strength, mass, dead and live loads, damping, and shear wall boundary element design) is 0.39, while the RTR *COV* of roof drift is 0.36. The percentage increase in *COV* (8.33 %) is used to amplify the RTR dispersion of the drift demand to get the total dispersion for each archetype.

3.5. Number of monte carlo simulations

In order to determine the probability of failure, 15,000 simulations are performed. An iterative assessment showed that this number of simulations was adequate to reduce the variance in the P_f estimate to an acceptably low level.

The obtained probability of failure is compared to a preselected target probability of failure to assess the performance. Furthermore, the effect of improved model accuracy on the design is quantified by comparing the required wall thicknesses corresponding to designs that yield a probability of failure less than that of the specified target.

4. Results & discussion

4.1. Effect of increased accuracy on performance

The reliability assessment was conducted for two scenarios: one considering only the RTR uncertainty in drift demand (shown in Fig. 12), and the other incorporating the total dispersion (RTR and modeling) in drift demand (shown in Fig. 15). Recall that the improvement in predictive performance obtained from the ML model is interpreted as a reduction in epistemic uncertainty associated with the drift capacity. By comparing the results for the two cases (RTR versus RTR + modeling uncertainty), we hope to gain insight into whether the implications of improved predictive performance on the structural design and performance is affected by the type(s) of uncertainty that is considered. Fig. 12a and 12b display boxplots that visualize the distribution of the probability of failure when only the RTR uncertainty in the drift demand is considered for the different archetypes, and SDC D_{max} and D_{min}, respectively. Each boxplot corresponds to a *COV* value. The boxplot is a visual representation of a given data set using five summary metrics. The metrics include the minimum, first quartile (Q1), median, third quartile (Q3), and maximum. The median is the middle value of the data set. Q1 is the middle value between the lowest value and the median, and Q3 is the middle value between the median and the highest value. The minimum and maximum are calculated as Q1-1.5*IQR and Q3 + 1.5*IQR, respectively, where IQR (interquartile range) is the distance between the Q1 and Q3. The outliers are the datapoints located outside the whiskers of the boxplot.

For SDC D_{max}, the probability of failure increases with increased *COV*, as expected. The target probability of failure is taken as 10 % as recommended by Abdullah and Wallace (2020) [82]. Archetypes at the higher end of the P_f range typically have fewer stories, as illustrated in Fig. 13. Additionally, there is a weak negative correlation between the P_f and wall thickness, as shown in Fig. 13a. Moreover, these archetypes at the higher end of the P_f range are characterized by shorter wall lengths and lower $\frac{l_w}{b}$ values, as depicted in Fig. 13b and 13c. The reduction in length leads to lower wall stiffness, consequently increasing the drift demand. Furthermore, lower $\frac{l_w}{b}$ values result in higher drift capacity. This association is driven by the fact that higher $\lambda_b = l_w c/b^2$ values contribute to increased drift capacity. Despite the higher drift capacity exhibited by these archetypes, their drift demand dispersion is higher which leads to an increased probability of failure.

For SDC D_{min}, the probability of failure also increases with higher *COV* values. However, the variation in the P_f is comparatively smaller. Furthermore, the P_f is below 10 % across all *COV* values, suggesting that the improved accuracy of the ML model is more likely to yield tangible benefits in regions of very high seismicity. The outlier is identified as the

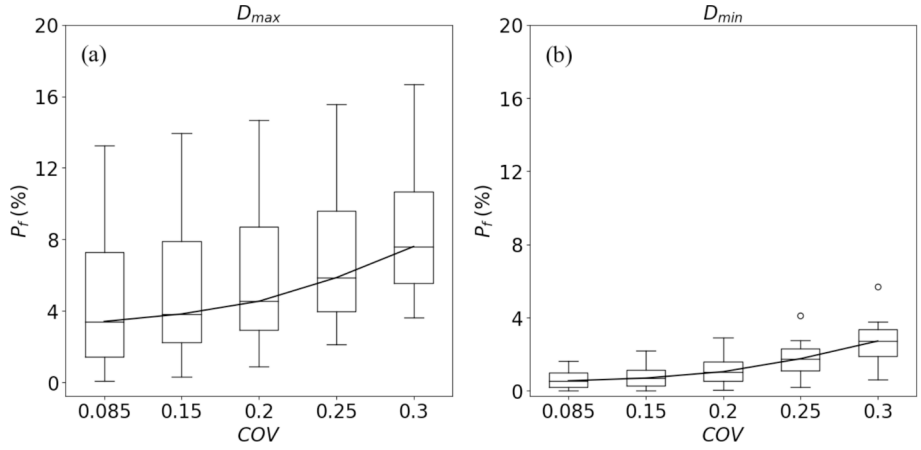

Fig. 12. Boxplots depicting the P_f considering RTR dispersion of drift demand under varying *COV* values for a) D_{max} and b)D_{min}

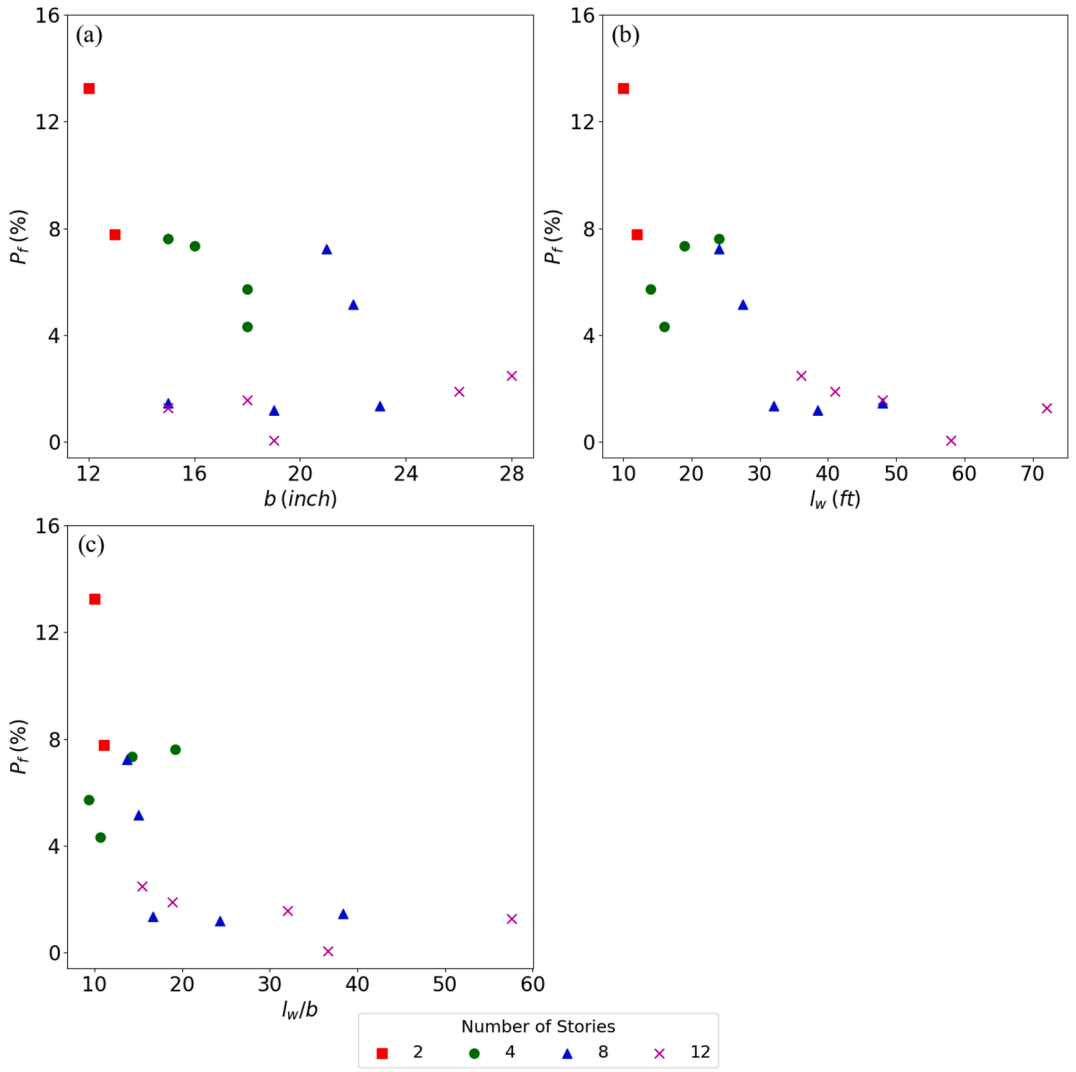

Fig. 13. Effect of archetype characteristics on the probability of failure for D_{max}: a) wall thickness, b) wall length, and c) $\frac{l_w}{b}$ ratio.

archetype with the smallest difference between capacity and demand. The majority of archetypes under SDC D_{min} have equal wall thickness. This equality makes it challenging to directly observe its impact on the P_f, as shown in Fig. 14a. Additionally, as seen in Fig. 14b, there is no discernible relationship between the wall length and the P_f. Consequently, $\frac{l_w}{b}$ also does not exhibit a correlation with the P_f, as depicted in Fig. 14c. Note that the results shown in Figs. 13 and 14 were obtained using $COV = 0.085$, however, similar conclusions were drawn for the remaining COV values.

The findings for the case where the total dispersion (RTR and modeling uncertainty) in the drift demand is considered (Fig. 15a and 15b for SDC D_{max} and D_{min}, respectively) are comparable to when only RTR uncertainty is included. The probability of failure increases slightly compared with the ones resulting from RTR variability only, however the increase is insignificant. More importantly, the rate at which P_f increases with an increase in COV is comparable for the two uncertainty cases. Note that the outlier in Fig. 15a corresponds to the archetype characterized by the thinnest wall and shortest length.

For most archetypes, the inclusion of modeling uncertainty does not significantly alter the influence of improved predictive accuracy on reliability. However, for two archetypes, relative to when only RTR uncertainty is considered, the added presence of modeling uncertainty increases the tangible effect of improved ML model accuracy on reliability. This suggests that the inclusion of modeling uncertainty could amplify the impact of ML model accuracy improvements on the reliability of structures.

In Fig. 16a, the ratio of the median probability of failure with and without modeling uncertainty $\left(\widetilde{P}_{f,Total}/\widetilde{P}_{f,RTR}\right)$ for each COV value is depicted. Correspondingly, Fig. 16b illustrates the ratio of the COV of probability of failure with and without modeling uncertainty $\left(COV_{P_{f,Total}}/COV_{P_{f,RTR}}\right)$. In general, the impact of modeling uncertainty decreases as the predictive model uncertainty increases. Additionally, modeling uncertainty has a more pronounced effect on the median P_f in regions of lower seismicity, while its impact on the dispersion of the P_f is stronger in regions of higher seismicity.

As mentioned in Section 2.2, the COV of the ML model and the linear model are 0.085 and 0.15, respectively. Based on this work, it appears that the XGBoost drift capacity model would not have a tangible impact on the reliability of the structure. The median and dispersion of the probability of failure for both COV values are similar. For the ML model to have a tangible impact, reduction in COV value by 60–68 % or higher is needed. While Table 1 indicates that the maximum improvement in COV provided by an ML model for drift capacity is 61 %, it is evident that for other types of components and structural responses, ML models have demonstrated the capability to enhance the COV beyond this threshold. Thus, there could be a potential to further improve ML models for

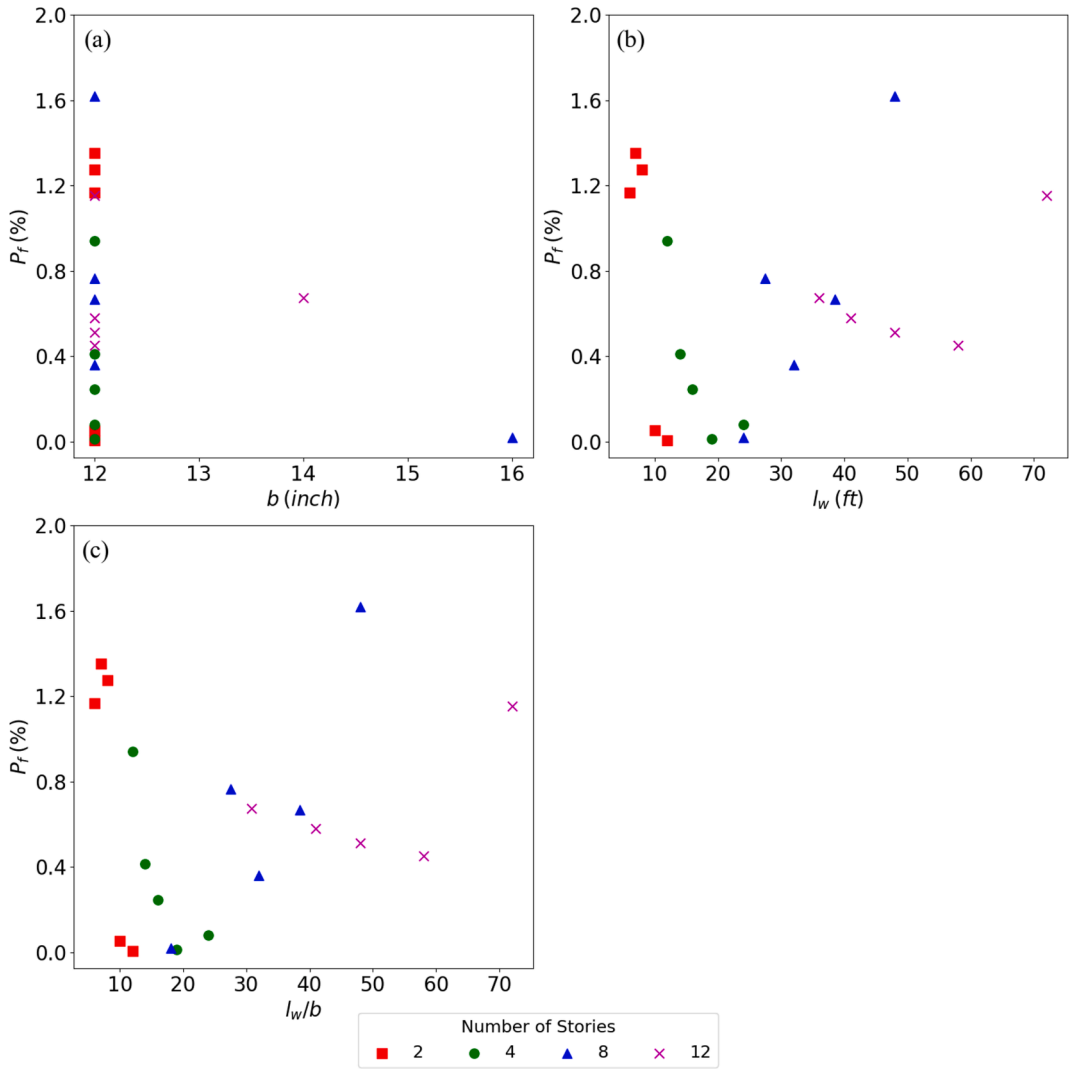

Fig. 14. Effect of archetype characteristics on the probability of failure for D_{min}: a) wall thickness, b) wall length, and c) $\frac{l_w}{b}$ ratio.

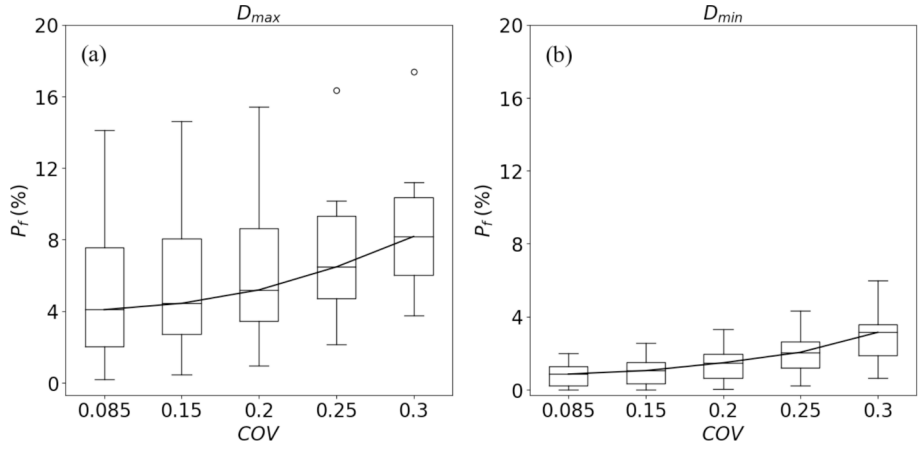

Fig. 15. Boxplots depicting the P_f considering total dispersion of drift demand under varying COV values for a) D_{max} and b)D_{min}

predicting drift capacity beyond the observed level.

The results in Figs. 12 and 15 indicate that the accuracy of the prediction model has direct implications to structural reliability. Specifically, a less accurate predictive model has a higher level of uncertainty, which overestimates the probabilities of failure, and thus leads to more conservative designs. The next subsection seeks to quantify the impact on the shear wall design.

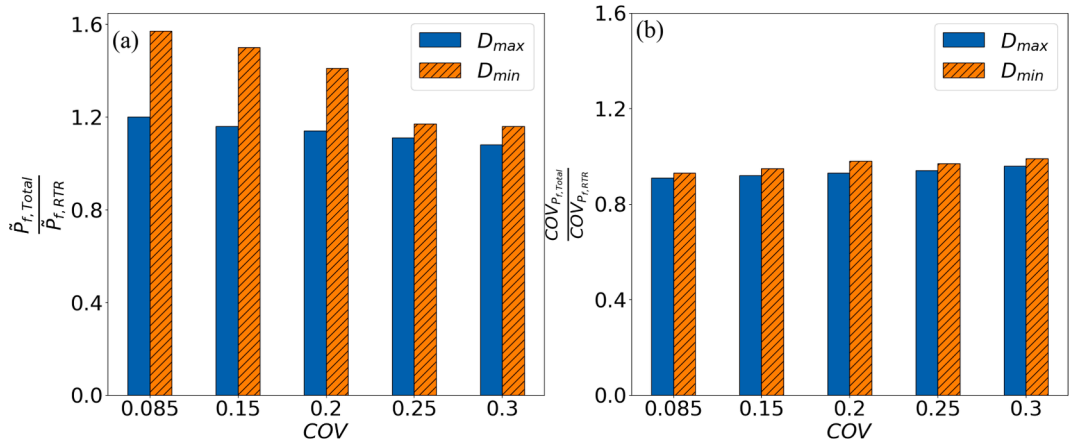

Fig. 16. Bar charts showing effect of modeling uncertainty in terms of a) ratio of median P_f and b) ratio of COV of P_f

4.2. Effect of increased accuracy on design

To quantify the effect of increased predictive model accuracy on reliability-based design, a target probability of failure is first established. Then, for each structure, the minimum wall thickness that results in a probability of failure that is less than or equal to the target is obtained. The impact of improved model accuracy is examined by comparing the required wall thicknesses for designs that meet the target.

The effect of predictive model accuracy on design is assessed and the results are shown in Table 3 for the case where only the RTR uncertainty in the drift demands is considered. The target probability of failure used to demonstrate the methodology is 10 %. Out of the 36 archetypes, 31 have a probability of failure less than 10 % for all COV values. On the other hand, 5 archetypes were redesigned to reach the target reliability. The results from these 5 archetypes showed that a change in COV from 0.15 to 0.085 did not impact wall thicknesses, while a change from 0.2 or 0.25 to 0.085 resulted in a 0–1 inch change of wall thicknesses. A more significant COV change from 0.3 to 0.085 led to 1–3 inch change in wall thickness.

In the case where both RTR and modeling uncertainty are considered for the drift demand dispersion, six archetypes needed to be redesigned to meet the target probability, as shown in Table 4. Among these six archetypes, the wall thickness remained consistent for COV values of 0.085 and 0.15, except for archetype 2, which required an increased thickness of 1 *inch*. Changing the COV from 0.2 or 0.25 to 0.085 resulted in a 0–1 *inch* change in wall thickness, while a COV change from 0.3 to 0.085 led to a 1–3 *inch* change in wall thickness. With the exception of two archetypes, the results from both cases are similar. For the aforementioned two archetypes, incorporating modeling uncertainty results in a small increase in the extent to which ML model accuracy affects in the structural design. This stands in contrast to the case where only RTR uncertainty is considered, where a more substantial improvement in ML model accuracy is required to induce a change in the design. It is important to note that this does not imply that a small change in COV always results in similar target reliability-compliant designs.

Table 3
Required wall thickness to reach target reliability considering only RTR dispersion of drift demand.

Archetype	Required wall thickness (*inch*)				
	COV = 0.085	COV = 0.15	COV = 0.2	COV = 0.25	COV = 0.3
1	13	13	13	13	14
2	16	16	17	17	19
3	15	15	15	15	16
4	16	16	16	16	17
11	21	21	21	21	23

Table 4
Required wall thickness to reach target reliability considering total dispersion of drift demand.

Archetype	Required wall thickness (*inch*)				
	COV = 0.085	COV = 0.15	COV = 0.2	COV = 0.25	COV = 0.3
1	13	13	13	13	14
2	16	17	17	17	19
3	15	15	15	16	16
4	16	16	16	16	17
6	18	18	18	18	19
11	21	21	21	21	23

5. Conclusion

This study used a reliability-based approach to assess the benefits of reduced epistemic uncertainty provided by machine learning (ML) models in terms of seismic design and performance. Specifically, an extreme gradient boosting (XGBoost) based drift capacity model for reinforced concrete shear walls (RCSWs) with special boundary elements (SBEs) was employed to investigate its advantages in a structural reliability context. However, the framework could be easily adjusted to include any ML-based model and other structural response parameters. The methodology was applied to a set of 36 archetype structures. The wall design process is automated using a python-based design tool that satisfies ACI 318–19 provisions for special structural walls. Monte Carlo Simulation (MCS) was utilized for the reliability analysis. The uncertainties in the drift demand and drift capacity were considered in the current study. Five COV values, indicative of the predictive accuracy of the drift capacity model, were included, with the lower two corresponding to the XGBoost and linear models.

The results of the reliability-based assessment showed that prediction models with different uncertainty levels generate different structural reliability outcomes. An increase in the error or uncertainty in a predictive model leads to overestimation of the probability of failure. However, both the linear and XGBoost models demonstrate similar levels of reliability, emphasizing the requirement for a more significant improvement in predictive accuracy to have an impact on reliability. The results also showed that enhanced accuracy of the predictive model is more likely to provide tangible advantages in areas characterized by extremely high seismic activity. The effect of improved accuracy on reliability-based design was also evaluated. In archetypes requiring redesign to achieve the target reliability, differences in wall thickness were primarily observed between the highest (0.3) and lowest (0.085) values of COV. Since a reduction in the error or uncertainty of a predictive model reduces overestimation of the probability of failure, exceedingly conservative designs can be avoided.

It is essential to emphasize that the findings derived from this study cannot be generalized beyond the specific archetypes under consideration. The nuances of structural behavior, uncertainty, and design are inherently influenced by the unique characteristics of each archetype. Consequently, further studies across a more extensive spectrum of structures are recommended to enhance the reliability and applicability of the findings.

The source code for the nonlinear modeling of RCSWs established using OpenSees is publicly available in a GitHub repository. [94].

CRediT authorship contribution statement

Muneera A. Aladsani: Writing – original draft, Methodology, Investigation, Formal analysis, Conceptualization. **Henry V. Burton:** Writing – review & editing, Validation, Supervision, Methodology, Investigation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

References

[1] Luo H, Paal SG. Machine Learning-Based Backbone Curve Model of Reinforced Concrete Columns Subjected to Cyclic Loading Reversals. J Comput Civ Eng 2018; 32. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000787.

[2] Luo H, Paal SG. A locally weighted machine learning model for generalized prediction of drift capacity in seismic vulnerability assessments. Comput-Aid Civ Infrastruct Eng 2019;34:935–50. https://doi.org/10.1111/mice.12456.

[3] Inel M. Modeling ultimate deformation capacity of RC columns using artificial neural networks. Eng Struct 2007;29:329–35. https://doi.org/10.1016/j.engstruct.2006.05.001.

[4] Feng DC, Liu ZT, Wang XD, Jiang ZM, Liang SX. Failure mode classification and bearing capacity prediction for reinforced concrete columns based on ensemble machine learning algorithm. Adv Eng. Inform 2020;45. https://doi.org/10.1016/j.aei.2020.101126.

[5] Mangalathu S, Jeon JS. Machine Learning-Based Failure Mode Recognition of Circular Reinforced Concrete Bridge Columns: Comparative Study. J Struct Eng (United States) 2019;145. https://doi.org/10.1061/(ASCE)ST.1943-541X.0002402.

[6] Zhang J, Sun Y, Li G, Wang Y, Sun J, Li J. Machine-learning-assisted shear strength prediction of reinforced concrete beams with and without stirrups. Eng Comput 2020. https://doi.org/10.1007/s00366-020-01076-x.

[7] Prayogo D, Cheng MY, Wu YW, Tran DH. Combining machine learning models via adaptive ensemble weighting for prediction of shear capacity of reinforced-concrete deep beams. Eng Comput 2020;36:1135–53. https://doi.org/10.1007/s00366-019-00753-w.

[8] Pham A-D, Ngo N-T, Nguyen T-K. Machine learning for predicting long-term deflections in reinforce concrete flexural structures. J Comput Des Eng 2020;7: 95–106. https://doi.org/10.1093/jcde/qwaa010.

[9] Moradi MJ, Hariri-Ardebili MA. Developing a library of shear walls database and the neural network based predictive meta-model. Applied Sciences (Switzerland) 2019;9. https://doi.org/10.3390/app9122562.

[10] Chen XL, Fu JP, Yao JL, Gan JF. Prediction of shear strength for squat RC walls using a hybrid ANN–PSO model. Eng Comput 2018;34:367–83. https://doi.org/10.1007/s00366-017-0547-5.

[11] Aladsani MA, Burton H, Abdullah SA, Wallace JW. Explainable Machine Learning Model for Predicting Drift Capacity of Reinforced Concrete Walls. ACI Struct J 2022;119. https://doi.org/10.14359/51734484.

[12] Mangalathu S, Jang H, Hwang SH, Jeon JS. Data-driven machine-learning-based seismic failure mode identification of reinforced concrete shear walls. Eng Struct 2020;208. https://doi.org/10.1016/j.engstruct.2020.110331.

[13] Almustafa MK, Nehdi ML. Machine learning model for predicting structural response of RC slabs exposed to blast loading. Eng Struct 2020;221. https://doi.org/10.1016/j.engstruct.2020.111109.

[14] Vu DT, Hoang ND. Punching shear capacity estimation of FRP-reinforced concrete slabs using a hybrid machine learning approach. Struct Infrastruct Eng 2016;12: 1153–61. https://doi.org/10.1080/15732479.2015.1086386.

[15] Alwanas AAH, Al-Musawi AA, Salih SQ, Tao H, Ali M, Yaseen ZM. Load-carrying capacity and mode failure simulation of beam-column joint connection: Application of self-tuning machine learning model. Eng Struct 2019;194:220–9. https://doi.org/10.1016/j.engstruct.2019.05.048.

[16] Mangalathu S, Jeon JS. Classification of failure mode and prediction of shear strength for reinforced concrete beam-column joints using machine learning techniques. Eng Struct 2018;160:85–94. https://doi.org/10.1016/j.engstruct.2018.01.008.

[17] Jeon JS, Shafieezadeh A, Desroches R. Statistical models for shear strength of RC beam-column joints using machine-learning techniques. Earthq Eng Struct Dyn 2014;43:2075–95. https://doi.org/10.1002/eqe.2437.

[18] Sun H, Burton HV, Huang H. Machine learning applications for building structural design and performance assessment: State-of-the-art review. J Build Eng 2021;33. https://doi.org/10.1016/j.jobe.2020.101816.

[19] Thai HT. Machine learning for structural engineering: A state-of-the-art review. Struct 2022;38. https://doi.org/10.1016/j.istruc.2022.02.003.

[20] Wang X, Mazumder RK, Salarieh B, Salman AM, Shafieezadeh A, Li Y. Machine Learning for Risk and Resilience Assessment in Structural Engineering: Progress and Future Trends. J Struct Eng 2022;148. https://doi.org/10.1061/(asce)st.1943-541x.0003392.

[21] Tapeh ATG, Naser MZ. Artificial Intelligence, Machine Learning, and Deep Learning in Structural Engineering: A Scientometrics Review of Trends and Best Practices. Arch Comput Methods Eng 2023;30. https://doi.org/10.1007/s11831-022-09793-w.

[22] Van Lent M, Fisher W, Mancuso M. An explainable artificial intelligence system for small-unit tactical behavior. Proceedings of the National Conference on Artificial Intelligence. 2004.

[23] Linardatos P, Papastefanopoulos V, Kotsiantis S. Explainable ai: A review of machine learning interpretability methods. Entropy 2021;23. https://doi.org/10.3390/e23010018.

[24] Mansour MY, Dicleli M, Lee JY, Zhang J. Predicting the shear strength of reinforced concrete beams using artificial neural networks. Eng Struct 2004;26:781–99. https://doi.org/10.1016/j.engstruct.2004.01.011.

[25] El Chabib H, Nehdi M, Saïd A. Predicting the effect of stirrups on shear strength of reinforced normal-strength concrete (NSC) and high-strength concrete (HSC) slender beams using artificial intelligence. Can. J Civ Eng 2006;33. https://doi.org/10.1139/L06-033.

[26] Abdalla JA, Elsanosi A, Abdelwahab A. Modeling and simulation of shear resistance of R/C beams using artificial neural network. J Franklin Inst 2007;344. https://doi.org/10.1016/j.jfranklin.2005.12.005.

[27] Fu B, Feng DC. A machine learning-based time-dependent shear strength model for corroded reinforced concrete beams. J Build Eng 2021;36. https://doi.org/10.1016/j.jobe.2020.102118.

[28] Feng DC, Wang WJ, Mangalathu S, Hu G, Wu T. Implementing ensemble learning methods to predict the shear strength of RC deep beams with/without web reinforcements. Eng Struct 2021;235. https://doi.org/10.1016/j.engstruct.2021.111979.

[29] Elsanadedy HM, Abbas H, Al-Salloum YA, Almusallam TH. Shear strength prediction of HSC slender beams without web reinforcement. Mater Struct 2016; 49. https://doi.org/10.1617/s11527-015-0752-x.

[30] Olalusi OB, Awoyera PO. Shear capacity prediction of slender reinforced concrete structures with steel fibers using machine learning. Eng Struct 2021;227. https://doi.org/10.1016/j.engstruct.2020.111470.

[31] Rahman J, Ahmed KS, Khan NI, Islam K, Mangalathu S. Data-driven shear strength prediction of steel fiber reinforced concrete beams using machine learning approach. Eng Struct 2021;233. https://doi.org/10.1016/j.engstruct.2020.111743.

[32] Abuodeh OR, Abdalla JA, Hawileh RA. Prediction of shear strength and behavior of RC beams strengthened with externally bonded FRP sheets using machine learning techniques. Compos Struct 2020;234. https://doi.org/10.1016/j.compstruct.2019.111698.

[33] Ly HB, Le TT, Vu HLT, Tran VQ, Le LM, Pham BT. Computational hybrid machine learning based prediction of shear capacity for steel fiber reinforced concrete beams. Sustainability 2020;12. https://doi.org/10.3390/su12072709.

[34] Sarveghadi M, Gandomi AH, Bolandi H, Alavi AH. Development of prediction models for shear strength of SFRCB using a machine learning approach. Neural Comput Appl 2019;31:2085–94. https://doi.org/10.1007/s00521-015-1997-6.

[35] Mangalathu S, Shin H, Choi E, Jeon JS. Explainable machine learning models for punching shear strength estimation of flat slabs without transverse reinforcement. J Build Eng 2021;39. https://doi.org/10.1016/j.jobe.2021.102300.

[36] Gandomi AH, Roke DA. Assessment of artificial neural network and genetic programming as predictive tools. Adv Eng Softw 2015;88. https://doi.org/10.1016/j.advengsoft.2015.05.007.

[37] Nguyen HD, Truong GT, Shin M. Development of extreme gradient boosting model for prediction of punching shear resistance of r/c interior slabs. Eng Struct 2021; 235. https://doi.org/10.1016/j.engstruct.2021.112067.

[38] Tran VL, Kim SE. A practical ANN model for predicting the PSS of two-way reinforced concrete slabs. Eng Comput 2021;37. https://doi.org/10.1007/s00366-020-00944-w.

[39] Caglar N. Neural network based approach for determining the shear strength of circular reinforced concrete columns. Constr Build Mater 2009;23. https://doi.org/10.1016/j.conbuildmat.2009.06.002.

[40] Zhou Q, Zhu F, Yang X, Wang F, Chi B, Zhang Z. Shear capacity estimation of fully grouted reinforced concrete masonry walls using neural network and adaptive neuro-fuzzy inference system models. Constr Build Mater 2017;153. https://doi.org/10.1016/j.conbuildmat.2017.07.171.

[41] Nguyen DD, Tran VL, Ha DH, Nguyen VQ, Lee TH. A machine learning-based formulation for predicting shear capacity of squat flanged RC walls. Struct 2021; 29. https://doi.org/10.1016/j.istruc.2020.12.054.
[42] Keshtegar B, Nehdi ML, Trung NT, Kolahchi R. Predicting load capacity of shear walls using SVR–RSM model. Appl. Soft Comput 2021;112. https://doi.org/10.1016/j.asoc.2021.107739.
[43] Kotsovou GM, Cotsovos DM, Lagaros ND. Assessment of RC exterior beam-column Joints based on artificial neural networks and other methods. Eng Struct 2017;144. https://doi.org/10.1016/j.engstruct.2017.04.048.
[44] Xu JG, Chen SZ, Xu WJ, Sen SZ. Concrete-to-concrete interface shear strength prediction based on explainable extreme gradient boosting approach. Constr Build Mater 2021;308. https://doi.org/10.1016/j.conbuildmat.2021.125088.
[45] Liu T, Wang Z, Zeng J, Wang J. Machine-learning-based models to predict shear transfer strength of concrete joints. Eng Struct 2021;249. https://doi.org/10.1016/j.engstruct.2021.113253.
[46] Sarothi SZ, Ahmed KS, Khan NI, Ahmed A, Nehdi ML. Predicting bearing capacity of double shear bolted connections using machine learning. Eng Struct 2022;251. https://doi.org/10.1016/j.engstruct.2021.113497.
[47] Tran VL, Kim SE. Efficiency of three advanced data-driven models for predicting axial compression capacity of CFDST columns. Thin-Walled Struct 2020;152. https://doi.org/10.1016/j.tws.2020.106744.
[48] Tran VL, Thai DK, Nguyen DD. Practical artificial neural network tool for predicting the axial compression capacity of circular concrete-filled steel tube columns with ultra-high-strength concrete. Thin-Walled Struct 2020;151. https://doi.org/10.1016/j.tws.2020.106720.
[49] Ho NX, Le TT. Effects of variability in experimental database on machine-learning-based prediction of ultimate load of circular concrete-filled steel tubes. Meas 2021; 176. https://doi.org/10.1016/j.measurement.2021.109198.
[50] Lee S, Vo TP, Thai HT, Lee J, Patel V. Strength prediction of concrete-filled steel tubular columns using Categorical Gradient Boosting algorithm. Eng Struct 2021; 238. https://doi.org/10.1016/j.engstruct.2021.112109.
[51] Le TT, Asteris PG, Lemonis ME. Prediction of axial load capacity of rectangular concrete-filled steel tube columns using machine learning techniques. Eng Comput 2022;38. https://doi.org/10.1007/s00366-021-01461-0.
[52] Le TT. Practical Hybrid Machine Learning Approach for Estimation of Ultimate Load of Elliptical Concrete-Filled Steel Tubular Columns under Axial Loading. Adv. Civ Eng 2020; 2020.. https://doi.org/10.1155/2020/8832522.
[53] Lyu F, Fan X, Ding F, Chen Z. Prediction of the axial compressive strength of circular concrete-filled steel tube columns using sine cosine algorithm-support vector regression. Compos Struct 2021;273. https://doi.org/10.1016/j.compstruct.2021.114282.
[54] Hanoon AN, Al Zand AW, Yaseen ZM. Designing new hybrid artificial intelligence model for CFST beam flexural performance prediction. Eng Comput 2022;38. https://doi.org/10.1007/s00366-021-01325-7.
[55] Basarir H, Elchalakani M, Karrech A. The prediction of ultimate pure bending moment of concrete-filled steel tubes by adaptive neuro-fuzzy inference system (ANFIS). Neural Comput Appl 2019;31. https://doi.org/10.1007/s00521-017-3108-3.
[56] Haroon M, Koo S, Shin D, Kim C. Torsional behavior evaluation of reinforced concrete beams using artificial neural network. Appl Sci 2021;11. https://doi.org/10.3390/app11104465.
[57] Arslan MH. Predicting of torsional strength of RC beams by using different artificial neural network algorithms and building codes. Adv Eng Softw 2010;41. https://doi.org/10.1016/j.advengsoft.2010.05.009.
[58] D'Aniello M, Güneyisi EM, Landolfo R, Mermerdaş K. Analytical prediction of available rotation capacity of cold-formed rectangular and square hollow section beams. Thin-Walled Struct 2014;77. https://doi.org/10.1016/j.tws.2013.09.015.
[59] Reis C, Ruivo P, Oliveira T, Faroleiro P. Assessing the drivers of machine learning business value. J Bus Res 2020;117. https://doi.org/10.1016/j.jbusres.2020.05.053.
[60] Jannach D, Jugovac M. Measuring the business value of recommender systems. ACM Trans Manag. Inf Syst 2019;10. https://doi.org/10.1145/3370082.
[61] Cheng HT, Koc L, Harmsen J, Shaked T, Chandra T, Aradhye H, et al. Wide & deep learning for recommender systems. ACM International Conference Proceeding Series 2016;vol. 15-:September-2016. https://doi.org/10.1145/2988450.2988454.
[62] Wen YK. Reliability and performance-based design. Struct Saf 2001;23. https://doi.org/10.1016/S0167-4730(02)00011-5.
[63] Celik OC, Ellingwood BR. Seismic fragilities for non-ductile reinforced concrete frames - Role of aleatoric and epistemic uncertainties. Struct Saf 2010;32. https://doi.org/10.1016/j.strusafe.2009.04.003.
[64] Der Kiureghian A, Ditlevsen O. Aleatory or epistemic? Does it matter? Struct Saf 2009;31. https://doi.org/10.1016/j.strusafe.2008.06.020.
[65] Bradley BA. A critical examination of seismic response uncertainty analysis in earthquake engineering. Earthq Eng Struct Dyn 2013;42. https://doi.org/10.1002/eqe.2331.
[66] Cubrinovski M, Ishihara K. State concept and modified elastoplasticity for sand modelling. Soil Found 1998;38. https://doi.org/10.3208/sandf.38.4_213.
[67] Yang Z, Elgamal A, Parra E. Computational Model for Cyclic Mobility and Associated Shear Deformation. J Geotech Geoenviron Eng 2003;129. doi: 10.1061/(asce)1090-0241(2003)129:12(1119).
[68] Otani SSAKE. A computer program for inelastic response of r/c frames to earthquakes. Ill Univ Dep Civ Eng Struct Res Ser 1974.
[69] Ibarra LF, Krawinkler H. Global Collapse of Frame Structures under Seismic Excitations. 2005.
[70] Dolsek M. Incremental dynamic analysis with consideration of modeling uncertainties. Earthq Eng Struct Dyn 2009;38. https://doi.org/10.1002/eqe.869.
[71] Liel AB, Haselton CB, Deierlein GG, Baker JW. Incorporating modeling uncertainties in the assessment of seismic collapse risk of buildings. Struct Saf 2009;31. https://doi.org/10.1016/j.strusafe.2008.06.002.
[72] Yin YJ, Li Y. Seismic collapse risk of light-frame wood construction considering aleatoric and epistemic uncertainties. Struct Saf 2010;32. https://doi.org/10.1016/j.strusafe.2010.03.012.
[73] Gokkaya BU, Baker JW, Deierlein GG. Quantifying the impacts of modeling uncertainties on the seismic drift demands and collapse risk of buildings with implications on seismic design checks. Earthq Eng Struct Dyn, vol. 45, 2016. doi: 10.1002/eqe.2740.
[74] Asgarian B, Ordoubadi B. Effects of structural uncertainties on seismic performance of steel moment resisting frames. J Constr Steel Res 2016;120. https://doi.org/10.1016/j.jcsr.2015.12.031.
[75] Haselton CB, Deierlein GG. Assessing seismic collapse safety of modern reinforced concrete moment-frame buildings. Berkeley, CA: Pacific Earthquake Engineering Research Center 2008. PEER Report 2007/08.
[76] Aslani H. Probabilistic earthquake loss estimation and loss disaggregation in buildings. Stanford University; 2005. PhD Thesis.
[77] Browning JA, Li YR, Lynn A, Moehle JP. Performance Assessment for a Reinforced Concrete Frame Building. Earthq Spectra 2000;16. https://doi.org/10.1193/1.1586126.
[78] Sattar S, Weigand JM, Wong KKF. Quantification of uncertainties in the response of beam-columns in steel moment frames. 11th National Conference on Earthquake Engineering 2018, NCEE 2018: Integrating Science, Engineering, and Policy, vol. 11, 2018.
[79] Asce. Minimum design loads for buildings and other structures. ASCE standard ASCE/SEI 2016:7–16.
[80] ACI Committee 318. Building Code Requirements for Structural Concrete (ACI 318-19) and Commentary (ACI 318R-19). 2019.
[81] Abdullah SA, Wallace JW. Drift capacity of reinforced concrete structural walls with special boundary elements. ACI Struct J 2019;116:183–94. https://doi.org/10.14359/51710864.
[82] Abdullah SA, Wallace JW. Reliability-based design methodology for reinforced concrete structural walls with special boundary elements. ACI Struct J 2020;117. https://doi.org/10.14359/51721375.
[83] McKenna F, Scott MH, Fenves GL. Nonlinear Finite-Element Analysis Software Architecture Using Object Composition. J Comput Civ Engi 2010;24. https://doi.org/10.1061/(asce)cp.1943-5487.0000002.
[84] Atc. Quantification of building seismic performance factors. US Department of Homeland Security. FEMA 2009.
[85] United States Geological Survey (USGS). Unified Hazard Tool. URL https://earthquake.usgs.gov/hazards/interactive/.
[86] Campbell KW, Bozorgnia Y. NGA-West2 ground motion model for the average horizontal components of PGA, PGV, and 5% damped linear acceleration response spectra. Earthq Spectra 2014;30. https://doi.org/10.1193/062913EQS175M.
[87] Kolozvari K, Orakcal K, Wallace JW. Modeling of Cyclic Shear-Flexure Interaction in Reinforced Concrete Structural Walls. I: Theory. J Struct Eng 2015;141. https://doi.org/10.1061/(asce)st.1943-541x.0001059.
[88] Kolozvari K, Tran TA, Orakcal K, Wallace JW. Modeling of Cyclic Shear-Flexure Interaction in Reinforced Concrete Structural Walls. II: Experimental Validation. J Struct Eng 2015;141. https://doi.org/10.1061/(asce)st.1943-541x.0001083.
[89] Kolozvari K, Orakcal K, Wallace JW. Shear-flexure interaction modeling for reinforced concrete structural walls and columns under reversed cyclic loading. PEER Report 2015/12. 2015..
[90] Yassin MH. Nonlinear Analysis of Prestressed Concrete Structures under Monotonic and Cycling Loads. Berkeley: University of California; 1994. PhD dissertation.
[91] Menegotto M, Pinto PE. Method of analysis for cyclically loaded reinforced concrete plane frames including changes in geometry and non-elastic behavior of elements under combined normal force and bending. Proceedings, IABSE. Symposium 1973.
[92] Filippou FC, Popov EP, Bertero VV. Effects of Bond Deterioration on Hysteretic Behaviour of Reinforced Concrete Joints. Earthquake Engineering Research Center 1983.
[93] Kim S. Reliability of Structural Wall Shear Design for Tall Reinforced Concrete Core Wall Buildings. Los Angeles: University of California; 2016. PhD dissertation.
[94] Aladsani M. Reliability-Based Quantification of the Benefits of Machine Learning Predictive Models in Seismic Structural Design and Performance Assessment. 2024. https://github.com/aladsani/RCSW_NRHA.