DOI: http://doi.org/10.15446/ing.investig.120363

Research article/Civil and Sanitary Engineering

Assessment of Progressive Collapse Resistance in Reinforced Concrete Structures Using Machine Learning Models

Evaluación de la resistencia al colapso progresivo en estructuras de hormigón armado mediante modelos de aprendizaje automático

Efe Selman[1], Çağatay Berke Erdaş[2]

ABSTRACT

Ensuring structural integrity in buildings and infrastructure under extreme loading conditions represents a pivotal challenge in modern civil engineering. Exposure to natural disasters, accidental impacts, and deliberate attacks can result in the application of unprecedented stresses, which may ultimately lead to progressive collapse and catastrophic failures. While traditional analytical methods are reliable, they often prove inadequate in meeting the increasing demand for rapid and accurate assessments in complex scenarios. However, recent advances in computational tools, particularly machine learning (ML), offer a new approach to address these challenges. In this study, nonlinear static analyses of 250 reinforced concrete systems are conducted within pushdown procedures. Load factor and vertical drift capacities of systems are obtained and accepted as target outputs for ML based predictions. By leveraging data-driven models, it becomes possible to predict structural behavior under extreme conditions with greater precision and efficiency. This study builds on this emerging field, aiming to provide novel insights into collapse mechanisms and robustness through advanced machine learning techniques.

Keywords: progressive collapse, machine learning, reinforced concrete, nonlinear structural analysis

RESUMEN

Garantizar la integridad estructural de los edificios y las infraestructuras en condiciones de carga extremas representa un reto fundamental en la ingeniería civil moderna. La exposición de las estructuras a desastres naturales, impactos accidentales y ataques deliberados puede dar lugar a la aplicación de tensiones sin precedentes, que en última instancia pueden provocar un colapso progresivo y fallos catastróficos. Aunque los métodos analíticos tradicionales son fiables, a menudo resultan inadecuados para satisfacer la creciente demanda de evaluaciones rápidas y precisas en escenarios complejos. Sin embargo, los recientes avances en herramientas computacionales, en particular el aprendizaje automático, ofrecen un nuevo enfoque para abordar estos retos. En este estudio, se llevan a cabo análisis estáticos no lineales de 250 sistemas de hormigón armado mediante procedimientos de empuje hacia abajo. El factor de carga y las capacidades de deriva vertical de los sistemas se obtienen y aceptan como resultados objetivo para las predicciones basadas en técnicas de aprendizaje automático. Al aprovechar los modelos basados en datos, es posible predecir el comportamiento estructural en condiciones extremas con mayor precisión y eficiencia. Este estudio se basa en este campo emergente y tiene como objetivo proporcionar nuevos conocimientos sobre los mecanismos de colapso y la robustez mediante técnicas avanzadas de aprendizaje automático.

Palabras clave: colapso progresivo, aprendizaje automático, hormigón armado, análisis estructural no lineal

Received: May 14th 2025
Accepted: May 13th 2026

Introduction

Potential progressive collapse scenarios resulting from disproportionate failure consequences have received considerable attention in the specialized literature. The chain reaction mechanism leading to structural failure propagates due to the nonproportional local failure of individual elements. The concept of robustness is central to various analysis frameworks and it is defined as the insensitivity of a structural system local damage from extreme loads [1], [2], [3], [4], [5], [6], [7], [8], [9], [10], [11], [12]. Deterministic redundancy and reliability-based indices have been proposed to quantify robustness, and linear or nonlinear analysis methodologies are applied to characterize the potentially disproportionate deformation geometries of systems under the remaining load-carrying capacity [1], [2].

[1] PhD in Civil Engineering, Ege University, Türkiye. Affiliation: Lecturer, İzmir University of Economics, Türkiye. Email: efe.selman@ieu.edu.tr

[2] PhD in Computer Engineering, Baskent University, Türkiye. Affiliation: Lecturer, Baskent University, Türkiye. Email: berdas@baskent.edu.tr

Artificial intelligence (AI) applications have demonstrated strong performance in estimating structural response parameters and ultimate capacity in structural engineering. However, research on progressive collapse assessment using AI applications based on machine learning (ML) or deep learning (DL) models remains limited in scope and has not thoroughly addressed structural plasticity aspects. Although several attempts have been made [13], [14], [15], [16], [17], [18], the need for practical and effective assessments methods that explicitly account for the plastic mechanical behavior of systems remains unmet.

A review of existing literature reveals that a relatively small body of work addresses the identification of progressive collapse parameters using ML and DL applications, in contrast to the substantially larger volume of studies on AI methods applied to general structural and seismic performance parameters. Numerous applications exist for predicting structural stiffness parameters, and realistic performance curves, including force-deformation, shear-top displacement, and pushover curves, have been successfully estimated with ML algorithms. For example, base shear and corresponding lateral displacement values have been predicted using inputs such as spectral acceleration, natural period [17], [18], material and geometric properties [19], [20], [21], [22], [23], and nonlinear mechanical parameters [12], [20]. In the study by Hwang *et al.* [20], a pushover curve comprising up to 100 points was reproduced with a high degree of agreement with analytical results. The most preferred ML techniques in this domain include artificial neural network (ANN), random forest (RF), k-nearest neighbor algorithm (kNN), Naïve Bayes, relevance vector machine (RVM) and support vector machine (SVM). Prominent studies have also addressed the estimation of structural robustness and progressive collapse resistance, with the most frequently employed techniques being *k*NN, extreme gradient boosting (XGBoost), back propagation neural network (BPNN), and one-dimensional convolutional neural network (1D-CNN) [13], [14], [15], [24], [25], [26], [27], [28], [29], [30]. Inputs are generally drawn from material and geometrical properties, vertical period (T_v), and force- and deformation-based capacity parameters of cross-sections [13], [15], [24], [25], [26], [27], [28], [29], [30]. Close predictions have been reported for the dynamic increase factor [16], failure load levels [13], [16], [24], [29], failure modes [26], and ultimate progressive collapse resistance. It has been noted that the predictive reliability is higher when deformation-based input parameters are employed. Collectively, these studies demonstrate the effectiveness of ML algorithms are well suited to capturing both regimes, and structural optimization can be guided directly by nonlinearity parameters such as the deformation capacity of member ends [15], [26], [27], [28], [29], [30]. The transition from compressive arch action to tensile catenary action is the primary mechanism complicating the progressive collapse behavior of structural elements. ML algorithms are well suited to capturing the nonlinear characteristics of both regimes, enabling structural optimization to be guided directly by plasticity-based parameters such as deformation capacity of ends [15], [27], [28].

To date, only a limited number of ML-based studies have addressed the progressive collapse resistance of systems, and existing investigations do not incorporate the full set of nonlinear mechanical parameters of elements alongside system-level geometric characteristics. This study aims to accurately estimate the progressive collapse resistance of systems by providing an alternative procedure through an efficient implementation of ML algorithms.

The importance and originality of this study lie in its systematic exploration of the applicability of ML algorithms trained on nonlinear mechanical parameters for progressive collapse performance assessment. As noted above, the majority of prior studies have relied on linear performance parameters or have created ML models with a large number of output performance parameters. In contrast, the present study examines the relationships among dominant nonlinear performance parameters and evaluates the applicability of ML algorithms without requiring any additional structural analyses beyond those used to generate the training dataset..

Within this framework, the variations in force-deformation and moment versus rotation responses of elements in the vicinity of the removed column are the focal quantities of interest. Instead of performing time-consuming and complex analytical procedures in progressive collapse evaluations, the ML algorithms developed herein offer the opportunity to gain practicality and accurate results about robustness of elements. This approach enables the rapid evaluation of a large body of structural systems in the selected microregion and facilitates the effective integration of progressive collapse mitigation strategies into design practice. Specifically, this paper aims to estimate the load factors and vertical drift values by feeding regression-based ML models with a purpose-built progressive collapse dataset. The paper is organized into four sections: Section I provides the introduction; Section ii covers data generation, analysis details, and ML algorithm details; Section III presents and discusses the results of the procedure; and Section IV states the conclusions.

Methodology

Progressive collapse analysis and data generation

Nonlinear static analysis was performed in this study to obtain practical performance parameters, most notably the load factor (LF), which directly characterizes the progressive collapse resistance of structural systems, enabling ML algorithms to predict these quantities. In this framework, the progressive collapse performances of 250 reinforced concrete (RC) framed systems was evaluated through nonlinear static (pushdown) analysis in SAP2000 [31] under a corner column removal scenario at the first story.

Characteristic compressive strengths of concrete were varied between 30 to 50 MPa for each system, Fig. 1 shows the constitutive stress-strain relationships of concrete and reinforcing steel adopted in accordance with TS500 [32], ACI 318-19 [33] and Turkish seismic code (TSC-2018) [34].

For the compressive behavior of concrete, the Mander-Priestley-Park model was adopted in the context of the main codes [32], [33], [34]. For unconfined concrete, the peak of the stress-strain curve occurs at a stress equal to the unconfined cylinder compressive strength, with the corresponding strain taken to be 0.2%. The curve parameters are calculated from Equations (1), (2), (3) and (4). In these expressions, f_c denotes characteristic compressive strength of concrete, f_c is the compressive stress at any point on the curve, ε_c is the compressive strain at any point, and ε_{cc} is the peak compressive strain of the confined concrete. The subscript *c* denotes the confined model, while 0 denotes the unconfined condition, so that ε_{c0} is equal to 0.2%. (1) is valid up to $2\varepsilon_{c0}$, beyond which the curve descends linearly to the spalling strain of 0.5%. E_c and E_{sec} are the tangent and secant elastic moduli of concrete, respectively.

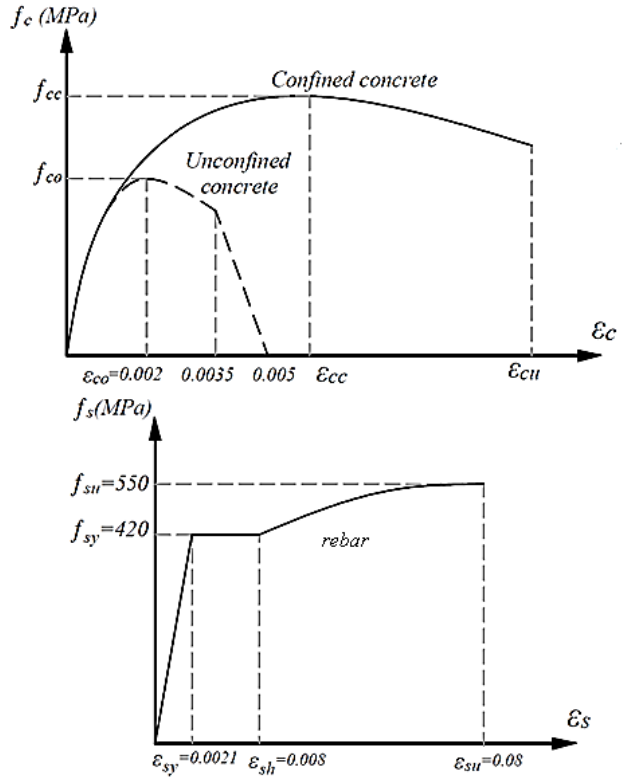


Fig. 1. Stress-strain relationships of concrete and reinforcement.
Source: Turkish seismic code (TSC-2018) [34].

$$f_c = \frac{f_c^{'} x.r}{r-1+x^r} \qquad (1)$$

$$x = \frac{\varepsilon_c}{\varepsilon_{cc}} \qquad (2)$$

$$r = \frac{E_c}{E_c - E_{\sec}} \qquad (3)$$

$$\varepsilon_{cc} = \varepsilon_{co}\left[1+5.\left(\frac{f_{cc}^{'}}{f_{c0}^{'}}-1\right)\right] \qquad (4)$$

For the tensile behavior of concrete, Chang-Mander model [32], [33], [34] was employed. In (5), f'_t denotes tensile strength of concrete while f_t is tensile stress at any point. The parameters n, x, and r are dimensionless ratios defined by (6) and (7). E_t is the tensile elastic modulus of concrete. ε_t is the compressive strain at any point, and ε_{tt} is the compressive strain corresponding to the tensile strength of concrete. The empirical constant r is recommended as 4 [33], [34].

$$f_t = f'_t \frac{nx}{1+(n-\frac{r}{r-1})x+\frac{x^r}{r-1}} \qquad (5)$$

$$n = \frac{E_t f_{c0}^{'}}{f_t^{'}} \qquad (6)$$

$$x = \frac{\varepsilon_t}{\varepsilon_{tt}} \qquad (7)$$

For the tensile and compressive stress-strain behavior of reinforcing steel, the Chang-Mander model [32], [33], [34] was implemented. The governing relationships are given in (8), (9), and (10), where f_s is the stress at any point of the rebar curve, f_{su} is the ultimate stress, f_{sy} is the yield stress, ε_s is the strain at any point, ε_{su} is the ultimate (peak) strain, ε_{sy} is the strain at the onset of the yield plateau, ε_{sh} is the strain at the end of the yield plateau prior to strain hardening, and E_s is the elastic modulus of steel.

$$f_s = E_s.\varepsilon_s \quad \varepsilon_s < \varepsilon_{sy} \qquad (8)$$

$$f_s = f_{sy} \quad \varepsilon_{sy} \le \varepsilon_s \le \varepsilon_{sh} \qquad (9)$$

$$f_s = f_{su} - (f_{su} - f_{sy})\frac{(\varepsilon_{su}-\varepsilon_s)^2}{(\varepsilon_{su}-\varepsilon_{sh})^2} \quad \varepsilon_s > \varepsilon_{sh} \qquad (10)$$

The geometric layout of the structural systems is illustrated in Fig. 2. The framed systems consist of 5 to 7 bays in the x-direction, 5 to 8 bays in the y-direction, and 5 to 10 stories, with a uniform story height of 3 m. Columns cross-sections are 800 x 800 mm and beam cross-sections are 500 x 500 mm throughout. The longitudinal rebar ratio of columns ranges from 0.03 to 0.04, while that of beams is fixed at 0.018. The transverse reinforcement ratio is 8 x 10^{-3} for all members. The slab thickness is 18 cm, with a volumetric reinforcement ratio of 0.015. An equivalent viscous damping ratio of 5% was assumed throughout the analysis.

The collapse scenario for all systems consists of the removal of the corner column at the first story, as indicated by the dashed circle in Fig. 2. The LF of each system is obtained as a function of the final vertical drift at the location of the removed column. The load factor is the ratio of the applied load to the GSA load combination of 2(D + 0.25L) [1], [2, where D and L denote dead and live loads, respectively .The vertical drift δ_v is defined as the ratio of the relative vertical displacement between the end points of beams in the vicinity of the removed column to the span length. All analyses were conducted under displacement control, with structural instability serving as the main control criteria. The target vertical displacement value is 2.5% of the total building, as recommended in [1], [2].

In the pushdown analyses, M3 flexural hinges were assigned to beams, while PM2M3 interacting axial-flexural hinges were assigned to columns and shear walls. The nonlinearity hinge properties adopted in this study are illustrated in Fig. 3. For beams controlled by flexure [35], with no compression rebar, the moment-rotation parameters are a = 0.02, b = 0.04, c = 0.2, with performance levels defined as: Immediate occupancy (IO) = 0.005 rad., life safety (LS) = 0.01 rad, and collapse prevention (CP) = 0.02 rad. For columns controlled by flexure [35], with no compression rebar, the corresponding parameters are a = 0.015, b = 0.025, and c = 0.2, with IO = 0.003 rad, LS = 0.012 rad, and CP = 0.015 rad. Using concentrated plasticity approach, plastic hinges are located

at the initial and end points of elements of all beam and column elements.

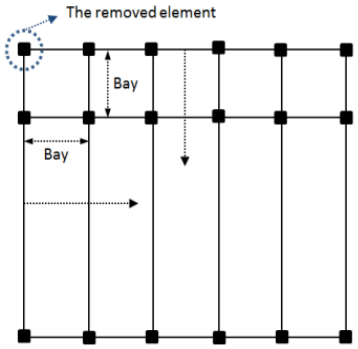


Fig. 2. Schematic layout of framed structural systems.
Source: Author.

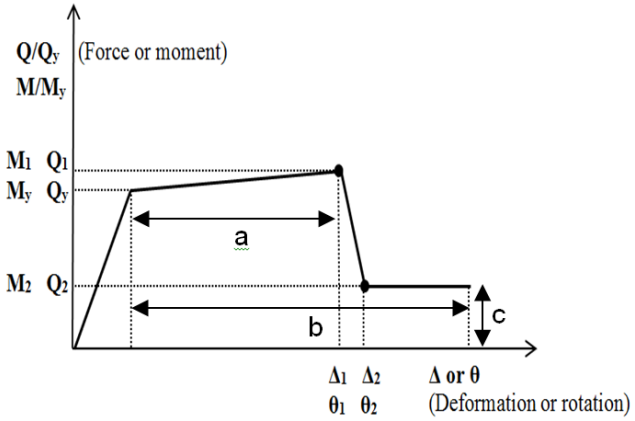


Figure 3. Nonlinearity relationships of hinging.
Source: Authors.

According to FEMA-356 [35], a plastic hinge length equal to half of the section depth is an acceptable approximation in the direction of loading, and this recommendation was adapted in the present study. Based on the plane-section hypothesis, the yield curvature φ_y is determined from (11).

$$\varphi_y = \frac{\varepsilon_{sy}}{\mathrm{d}-c_y} \qquad (11)$$

where ε_{sy} is the yield strain of the tensile longitudinal reinforcement, d is the distance from tensile longitudinal reinforcement to the extreme compression fiber, and c_y is the distance from the neutral axis to the extreme compression fiber at the yield level. The yield rotation ϑ_y is derived from the yield curvature as shown in Fig. 4. The ultimate curvature φ_u, ultimate rotation ϑ_u yield deformation Δ_y, and plastic deformation Δ_p are the main parameters of this relationship. The XTRACT program [36] was utilized to determine moment-rotation relationships of members, from which moment-curvature relationships are derived and incorporated into the hinge properties in SAP2000 [31]. For columns, the yield and ultimate moment values were updated using the three-dimensional interaction surface for the columns in SAP2000 [31].

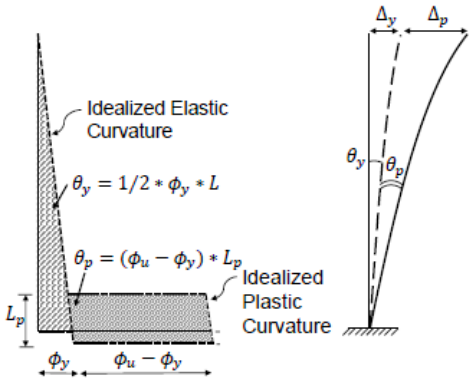


Figure 4. The schematic lay-out of framed structural systems
Source: [37]

Based on the fundamental characteristics of progressive collapse RC structures, the inputs variables are divided into two groups: descriptive parameters and primary nonlinear collapse parameters. Descriptive inputs include the number of bays and bay lengths in both x and y directions (n_x, n_y, l_x, l_y), total building height (H), ultimate strength of concrete (f_{cs}), and longitudinal reinforcement ratio of column (ρ_{lc}). The statistical properties of all input and output parameters of 250 building systems are listed in Table 1. Random sampling was employed to ensure representative coverage of realistic building configurations. The input parameters are within a certain range which allows variable output values. The complete analysis workflow is summarized schematically in Fig. 5.

The nonlinear collapse inputs reflect the highly nonlinear nature of progressive collapse behavior and include: maximum moment capacity (M_1), maximum force capacity (Q_1), residual moment (M_2), residual force (Q_2), deformation corresponding to the maximum force capacity (Δ_1), ultimate deformation (Δ_2), rotation corresponding to the maximum moment capacity (ϑ_1), ultimate rotation (ϑ_2), and the first-mode vertical vibration period (T_v). The main collapse parameters are read from the base section of the first story columns adjacent to the removed corner column, with maximum values recorded for each system, as listed in Table 1. The output variables are the load factor (LF) and the corresponding vertical drifts (δ_v) both obtained from the nonlinear static analysis in SAP2000 [31].

The correlation matrix of the complete dataset is presented in Fig. 6. The color gradation ranges from green (negative correlation) to purple (positive correlation), with white indicating negligible linear dependence between the variables. The heatmap confirms that both output variables are correlated with all input variables to varying degrees. The inherently complex nature of LF prevents uniformly strong correlations across all parameters; however, LF shows medium strong correlation with the nonlinear parameters as M_1, M_2, ϑ_2, Q_1, Q_2, Δ_1 and Δ_2. The dominant trend indicates that LF increases as the maximum moment capacity of column ends (M_1), increases, consistent with a compressive arch collapse mechanism. The rotation capacity at maximum moment (θ_1) shows a weak correlation with LF. As rotations increase beyond the peak moment, the collapse mechanism transitions to tensile catenary action in certain elements; consequently, a positive collinearity is observed between the residual moment (M_2) and LF, though at a lower magnitude than between M_1 and LF. As expected, the ultimate rotation capacity (ϑ_2) is positively associated with LF. For vertical drift, correlations are generally low at early load stages but strengthen significantly beyond force and moment capacity values,

vertical drift become strongly linked to the residual moment and shear capacities (M_2 and Q_2) where the positive collinearity can be adequately seen. This trend is also reflected in the corresponding rotation and deformation variables (Δ_2 and ϑ_2), which exhibit the highest positive correlation coefficients, reaching approximately 0.8.

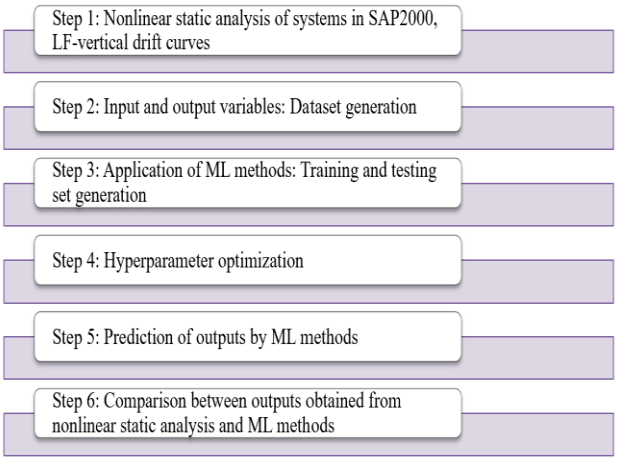


Figure 5. Workflow for progressive collapse model generation.
Source: Authors.

The first-mode vertical period (T_v) has limited influence on LF and δ_v; however, weak negative correlations are consistently observed at the lower degrees. Regarding the quantity of bays and bay lengths, the increase in the number of bays in either direction (n_x and n_y) generally promotes higher LF values, since greater structural redundancy enhances progressive collapse resistance. The range of n_y(5 to 8) is wider than that of n_x(5 to 7), which introduces additional plan asymmetry; consequently, the beneficial effect of redundancy on LF is more clearly reflected in the y-direction. The vertical drift δ_v shows a weak negative correlation with the number of bays in both directions, indicating that improved redundancy slightly decreases vertical drift values. Conversely, longer bay spans in both directions (l_x and l_y) have a negative impact on the progressive collapse resistance, yielding negative correlations with LF and positive correlations with δ_v at the medium levels.

Table 1. Dataset statistics

	Min.	Max.	Mean	Std.
n_x	5	7	5.33	0.65
n_y	5	8	6.125	0.93
l_x (cm)	400	450	418.6	18.32
l_y (cm)	400	700	550	89.62
ρ_{lc}	0.02	0.04	0.0302	0.005
H (cm)	1500	3100	2330	653.88
f_{cs} (MPa)	30	50	42	7.18
M_1	0.0135	0.027	0.0189	0.0024
ϑ_1	0.002016	0.0225	0.00346	0.00361
M_2	0.01	0.0918	0.01828	0.00733
ϑ_2	0.01	0.0341	0.01896	0.0046
Q_1	0.42	1.08	0.687	0.13
Δ_1	0.58	1.26	0.89	0.15
Q_2	0.02	0.89	0.45	0.294
Δ_2	0.399	1.28	0.84	0.19
T_V (s)	0.06	0.226	0.159	0.055
LF	0.892	1.122	0.995	0.068
δ_v	0.11	0.16	0.139	0.0136

Source: Authors

The total height of systems (H) presents an interesting dual effect: as H increases, the self-weight of the structure increases and progressive collapse resistance is negatively affected; however, the number of stories and the associated structural redundancy simultaneously increase. All adverse effects combined, the redundancy development become dominant for this dataset and the total height (H) positively correlates with LF at the lower levels, but the vertical drift rises dramatically up due to the increase of the total weight with a high degree of correlation. Parameters that enhance both robustness and general structural performance — namely concrete grade (f_{cs}) and longitudinal reinforcement ratio (ρ_{lc}) — yield favorable robustness characteristics. The medium strong correlation is achieved due to the multiparameter interaction during the collapse, yet the tendency confirms that higher concrete grades desirable and reinforcement rations produce more favorable LF values and reduced vertical drifts.

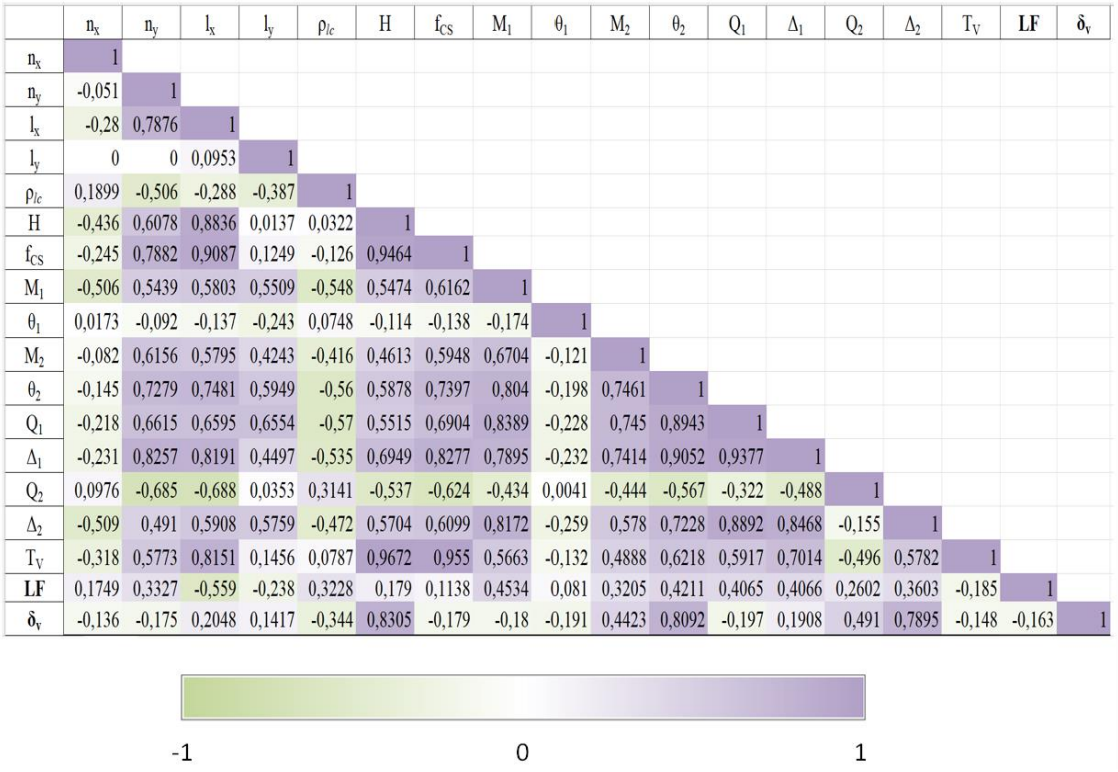
Unlike other studies, the present study aims to accurately predict structural performance using ML algorithms trained on non-linear mechanical parameters, rather than conducting time-consuming analyses.

Machine learning methods

Determining the collapse resistance of structural systems subjected to extreme loading conditions was formulated as a regression problem within the field of machine learning. Geometric and nonlinear collapse properties served as predictive inputs for model training, with the objective of accurately predicting structural behavior under load and effectively assessing the durability of systems.

The random forest (RF) algorithm is a powerful and flexible machine learning technique that is frequently employed in the context of data classification and regression problems. This method generates multiple decision trees and aggregates the independent predictions of each tree to generate a final prediction. The simultaneous application of random sampling and feature selection at each split renders the model resistant to over-learning while concomitantly enhancing its generalization capacity.

The k-nearest neighbor (*k*NN) algorithm is a parsimonious and intuitive approach employed in classification and regression problems. A prediction is made by identifying the k nearest training samples in the feature space and aggregating their target values. The calculation of the neighborhood distance is typically performed using metrics such as Euclidean, Manhattan, or Minkowski distance. While kNN is distinguished by its transparency, straightforward applicability, and explanatory structure, particularly in data-intensive domains, its predictive performance can degrade in high-dimensional datasets, being a potential drawback of the algorithm.



	n_x	n_y	l_x	l_y	ρ_{lc}	H	f_{CS}	M_1	θ_1	M_2	θ_2	Q_1	Δ_1	Q_2	Δ_2	T_V	**LF**	δ_v
n_x	1																	
n_y	-0,051	1																
l_x	-0,28	0,7876	1															
l_y	0	0	0,0953	1														
ρ_{lc}	0,1899	-0,506	-0,288	-0,387	1													
H	-0,436	0,6078	0,8836	0,0137	0,0322	1												
f_{CS}	-0,245	0,7882	0,9087	0,1249	-0,126	0,9464	1											
M_1	-0,506	0,5439	0,5803	0,5509	-0,548	0,5474	0,6162	1										
θ_1	0,0173	-0,092	-0,137	-0,243	0,0748	-0,114	-0,138	-0,174	1									
M_2	-0,082	0,6156	0,5795	0,4243	-0,416	0,4613	0,5948	0,6704	-0,121	1								
θ_2	-0,145	0,7279	0,7481	0,5949	-0,56	0,5878	0,7397	0,804	-0,198	0,7461	1							
Q_1	-0,218	0,6615	0,6595	0,6554	-0,57	0,5515	0,6904	0,8389	-0,228	0,745	0,8943	1						
Δ_1	-0,231	0,8257	0,8191	0,4497	-0,535	0,6949	0,8277	0,7895	-0,232	0,7414	0,9052	0,9377	1					
Q_2	0,0976	-0,685	-0,688	0,0353	0,3141	-0,537	-0,624	-0,434	0,0041	-0,444	-0,567	-0,322	-0,488	1				
Δ_2	-0,509	0,491	0,5908	0,5759	-0,472	0,5704	0,6099	0,8172	-0,259	0,578	0,7228	0,8892	0,8468	-0,155	1			
T_V	-0,318	0,5773	0,8151	0,1456	0,0787	0,9672	0,955	0,5663	-0,132	0,4888	0,6218	0,5917	0,7014	-0,496	0,5782	1		
LF	0,1749	0,3327	-0,559	-0,238	0,3228	0,179	0,1138	0,4534	0,081	0,3205	0,4211	0,4065	0,4066	0,2602	0,3603	-0,185	1	
δ_v	-0,136	-0,175	0,2048	0,1417	-0,344	0,8305	-0,179	-0,18	-0,191	0,4423	0,8092	-0,197	0,1908	0,491	0,7895	-0,148	-0,163	1

Figure 6. Correlation matrix of parameters in datasets.
Source: Authors.

The decision tree (DT) model is a machine learning algorithm that is widely used in the context of classification and regression problems. The model recursively partitions the dataset using a tree structure comprising decision points and performs the prediction operation on these partitions. Each node decides based on a feature, while branches represent possible outcomes and leaf nodes contain the predicted outcome. Decision trees provide an intuitive understanding and can be easily interpreted due to their visualizable structure. However, they are prone to overfitting and are thus usually supported by hyperparameter optimization or pruning methods.

Experimental setup and evaluation metrics

In all experiments that employed random forest and decision tree methods, the optimal features were identified as the basis for splitting, with entropy used for classification and MSE for regression as the feature selection metric. The maximum depth of the trees was unrestricted. For *k*NN, the number of neighbors was set to $k = 3$ with the Manhattan distance metric. In the Bayesian Ridge model, optimization was performed over a maximum of 300 iterations with a convergence tolerance of 1e-3; a Gamma distribution was utilized for the α and λ parameters, with the hyperparameters set at 1e-6, respectively. The model was trained to include the intercept term, and the data were not normalized. For the Gradient Boosting Regressor, MSE was adopted as the loss function, the learning rate was fixed at 0.1, the model incorporated a maximum of 100 trees, with a maximum depth of 3 for each tree, and a minimum of 1 sample was required for each leaf node; all features were considered, and no seed value was designated for randomness. In all experiments, a *k*-fold cross-validation procedure was applied. This entailed the division of the dataset into *k* randomly selected parts, with each part subsequently separated for testing while the remaining parts were used in the training phase. This process was repeated *k* times [38]. The evaluation criteria employed to assess the regression performance of the models were the correlation coefficient (R), coefficient of determination (R^2), Mean Absolute Error (MAE), Median Absolute Error (MedAE), and Mean Squared Error (MSE) [39].

The Bayesian Ridge regression model applies Bayes' theorem to the linear regression problem by integrating an L2 regularization term with probability distributions over the model parameters, yielding stable and well-regularized estimates. The model offers effective results, particularly in datasets comprising a limited number of samples or exhibiting multiple linear dependencies. Hyperparameters are optimized automatically during training to prevent overfitting, and the probabilistic formulation provides uncertainty estimates alongside point predictions.

The gradient boosting regressor (GBR) is a ML algorithm in which weak learners (typically decision trees) are trained sequentially, with each stage aiming to reduce the errors of the previous model. Each weak learner is trained based on its error gradient, resulting in a strong ensemble model that exhibits high accuracy and generalization ability. Regularization through shrinkage (learning rate) and subsampling further mitigates overfitting, making GBR particularly effective for complex data structures and variable relationships.

The correlation coefficient (*R*) quantifies the linear agreement between predicted values (X) and ground-truth values (Y) and is defined in (12), where $Cov(X, Y)$ is the covariance between predicted and true variables, and σ_x and σ_y are their respective standards. The coefficient of determination (R^2) is analogous to *R*,

measuring the explanatory power of a model. However, it offers a more balanced assessment by penalizing model complexity. Mean absolute error (MAE) represents the average of the prediction errors, with lower values indicating superior performance, and is defined in Equation (13), where $\hat{y_i}$ and y_i predicted and actual values, respectively, and $n$is the number of samples.

$$R = \frac{Cov(X,Y)}{\sigma_x \ \ \sigma_y} \tag{12}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y_i}| \tag{13}$$

The MedAE provides the median of the prediction errors, offering robustness to outliers, while MSE calculates the mean of the squares of the errors, and thus penalizes large errors more. Equation (14) illustrates the MedAE formulation, whereas Equation (15) depicts the MSE formulation.

$$\text{MedAE} = (|y_i - \hat{y_i}|) \tag{14}$$

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y_i})^2 \tag{15}$$

Results and discussions

The results obtained from both nonlinear static analysis and the ML models are discussed below, with primary emphasis on the role of plastic hinge nonlinearity parameters (*M-ϑ* and *Q-Δ*) in governing progressive collapse resistance.

The lateral restraints provided by adjacent structural elements ensure compressive arch mechanism which induces compression at the ends, and this provides additional vertical load capacity at the lower load stages. The flexural and compressive arching capacities are surely influenced by the moment capacity and total curvature characteristics of elements. As long as the vertical deflection at one end of an element remains below the section depth, higher ultimate moment capacity and satisfying corresponding rotation values lead to robust systems which have greater LF. Under compressive arch action, elements with greater moment capacity at their ends are more effective in redistributing and sustaining lowe displacement values of the element ends. Therefore, LF can be closely related to the ultimate moment capacity particularly for systems which go through compressive arch mechanism.

When the lateral restraints at element ends develop tensile forces and the vertical deflection at one end exceeds the section depth, the major parameters become the ultimate rotation and the residual moment capacity. The greater ultimate rotation capacity of beam ends creates additional capacity to the elements to cope with increasing vertical displacement values. Accordingly, the ultimate rotation capacity of member ends with corresponding residual moment capacity positively support LF and δ_v while the tensile catenary mechanism is active.

The axial load-carrying of a member is directly related to the length of plastic hinging zone, and the axial load level increases the plastic hinge length of columns regardless of whether compressive arch or tensile catenary is the active mechanism. Axial strain concentrated on the damaged zones can be effectively controlled by the axial elongation capacity of members. More strong axial capacity values minimize plastic curvature values arranging plastic deformation distribution and positively contribute to the robustness of elements. The obtained outputs prove it by the strong collinearity between (*F-Δ*) and LF and vertical δ_v.

Five ML regression models — random forest (RF), *k*NN, decision tree (DT), Bayesian Ridge (BR), and gradient boosting regressor (GBR) — were evaluated for the prediction of LF and δ_v, which are the two primary progressive collapse performance parameters. In this context, the regression performance obtained for the LF problem with the models is shown in Table 2, while the results obtained for drift are given in Table 3.

Table 2. Regression results for LF

	R	R^2	MAE	MedAE	MSE
Random Forest	0.9925	0.9850	0.0046	0.0018	0.0477
*k*NN	0.8796	0.7737	0.0780	0.0634	0.0847
Decision Tree	0.9929	0.9859	0.0036	0.0010	0.0448
Bayesian Ridge	0.9674	0.9359	0.0127	0.0299	0.0603
Gradient Boosting Regressor	0.9905	0.9810	0.0061	0.0033	0.0522

Source: Authors

Table 3. Regression results for vertical drift

	R	R^2	MAE	MedAE	MSE
Random Forest	0.9932	0.9864	0.0008	0.0002	0.0063
*k*NN	0.9458	0.8946	0.0032	0.0033	0.0132
Decision Tree	0.9875	0.9359	0.0008	0.0014	0.0115
Bayesian Ridge	0.9565	0.9149	0.0032	0.0031	0.0107
Gradient Boosting Regressor	0.9942	0.9884	0.0008	0.0004	0.0045

Source: Authors

Table 2 illustrates the performance of various machine learning algorithms in regression tasks for estimating the LF value. The decision tree and random forest models were identified as the most successful in predicting the target variable, exhibiting the highest *R* (0.9929 and 0.9925) and R^2 (0.9859 and 0.9850) values. These models also demonstrate that errors are minimized at both the general level and in the median context, with low MAE and MedAE values; nevertheless, the gradient boosting regressor exhibited a marginally superior outcome in terms of the MSE value. The *k*NN and Bayesian Ridge algorithms demonstrated comparatively diminished performance. Although *k*NN exhibits relatively inferior performance about *R* (0.8796) and R^2 (0.7737) values, it also displays higher values in error metrics than other models. Similarly, the Bayesian Ridge model demonstrated success regarding *R* (0.9674) and R^2 (0.9359) but was consistently outperformed by the ensemble-based methods. In general, random forest and gradient boosting regressor were identified as the most effective options for regression problems, with high accuracy and low error metrics. While other algorithms may be preferred in certain cases, the success of these two models is evident.

Table 3 presents the results obtained for the estimation of the drift value. In this context, the gradient boosting regressor and random forest models have the highest accuracy for δ_v prediction, with R values of 0.9942 and 0.9932 and R^2 values of 0.9884 and

0.9864, respectively. Both models demonstrated an ability to minimize errors, as evidenced by their low MAE and MedAE values. Although the decision tree model yielded slightly lower R and R^2 values (0.9875 and 0.9359) in comparison to the other models, it merits attention due to its exceptionally low MSE value. This indicates that the margin of error in the model's estimates is remarkably minimal. The *k*NN model again produced the lowest (0.9458 and 0.8946) in comparison to the other models. Overall, the gradient boosting regressor and random forest models yielded the most favorable results, while the decision tree model demonstrated notable success with its exceptionally low error values.

Conclusions

This study has identified any successful predictive models using ML approaches to effectively and practically assess the progressive collapse resistance parameters of systems which are LF and vertical drift values. Nonlinear static pushdown analyses were conducted in a displacement-controlled manner for 250 reinforced concrete framed systems. For the quantitative assessment of the robustness for resisting progressive collapse of systems, load factors and corresponding vertical drifts are obtained for each system and accepted as main outputs for ML algorithms. The inputs consist of main structural and nonlinear plastic hinging parameters as unfavorable force-deformation and moment-rotation relationships of ends in the removed element region. Unlike many prior studies, this work has accurately examined the role of effective nonlinear performance parameters on structural behavior and demonstrated the applicability of ML algorithms across a borad range of system configurations, entirely bypassing the need for repeated time-consuming structural analysis methods.

Among the five models evaluated, ensemble-based methods (random forest and gradient boosting regressor) exhibiting superior performance compared to other techniques. The results demonstrate the effectiveness of ML in estimating collapse resistance and drift values, offering a significant reduction in computational effort compared to traditional methods. The incorporation of nonlinear mechanical parameters into the models was shown to enhance their predictive power.

This study successfully integrates ML techniques into the evaluation of progressive collapse resistance in RC structures. By leveraging ML, engineers can efficiently predict critical parameters, facilitating risk mitigation strategies. Future work will expand the dataset and explore the integration of additional structural parameters.

Acknowledgements

The authors are grateful for the computational resources made available by İzmir University of Economics for this study.

Author contributions

All authors contributed equally to the conceptualization of the research, background research, workflow development, assessments, and preparation of the original draft.

Conflicts of interest

No potential conflict of interest is reported by the authors.

Data availability

All data, analyses, and results of this study are presented within the article as submitted.

Statement on artificial intelligence

The authors did not use IAG. The authors take full responsibility for the contents of this publication

References

[1]. General Service Administration, *Progressive Collapse Analysis and Design Guidelines for New Federal Office Buildings and Major Modernization Projects*. GSA 2016. Washington, USA, Jan. 2016.

[2]. Department of Defense (DoD), *Design of Buildings to Resist Progressive Collapse*. Unified Facilities Criteria UFC 4-023-03, Washington, USA, Jan. 2024.

[3]. American Society of Civil Engineers, *Minimum Design Loads for Buildings and Other Structures*, ASCE 7-16, Reston, VA, USA, 2016.

[4]. F. Petrone, L. Shan, S.K. Kunnath, "Modeling of RC frame buildings for progressive collapse analysis," *Int. J. Concr. Struct. Mater*, vol.10, no.1, pp. 1–13, Mar. 2016, https://doi.org/10.1007/s40069-016-0126-y.

[5]. J. Kim and M. Jung, "Progressive collapse resisting capacity of modular mega frame buildings," *Struct. Des. Tall Spec. Build.*, vol. 22, no. 6, pp. 471–484, Apr. 2013, https://doi.org/10.1002/tal.697

[6]. L. Shan, F. Petrone, and S. Kunnath, "Robustness of RC buildings to progressive collapse: Influence of building height," *Eng. Struct.,* vol.183, pp. 690–701, Mar. 2019, https://doi.org/10.1016/j.engstruct.2019.01.052

[7]. L. Shan, "Progressive collapse simulation of multi-story reinforced concrete buildings," PhD dissertation, Dept. Civ. Environ. Eng., Univ. of California, Berkeley, CA, USA, 2017.

[8]. Y. Cheng, J. Liu, L. Sun, Z. Xiao, and C. Ou, "Dynamic performance of multi-column removal framed structures subjected to impact and heaped loads," *Struct. Eng. Int.*, vol.1, pp. 1-13, Jan. 2024. https://doi.org/10.1080/10168664.2023.2271500

[9]. H. Xiao and B. Hedegaard, "Flexural, compressive arch, and catenary mechanisms in pseudostatic progressive collapse analysis," *J. Perform. Constr. Facil.*, vol. 32, no. 1, Art. no. 04017115, Feb. 2018. https://doi.org/10.1061/(ASCE)CF.1943-5509.0001110

[10]. Y. Bao and K. H. Tan, "Analytical model of compressive arch action for one-way RC beams and two-way framed substructures under column removal scenarios," *Eng. Struct.*, vol. 304, 117694, Jun. 2024. https://doi.org/10.1016/j.engstruct.2024.117694

[11]. J. F. Cheng, Q. Gu, K. Qian, F. Fu, and X. F. Deng, "Dynamic behavior of RC frames against progressive collapse subjected to loss of a corner column scenario," *J. Build. Eng.*, vol. 86, Art. no. 108872, Feb. 2024. https://doi.org/10.1016/j.jobe.2024.108872

[12]. F. Kiakojouri, E. Zeinali, J.M. Adam, and V. De Biagi, "Experimental studies on the progressive collapse of building structures: A review and discussion on dynamic column removal

techniques," *Structures*, vol. 57, Art. no. 105059, Nov. 2023. https://doi.org/10.1016/j.istruc.2023.105059

[13].C. Angarita, C. Montes, and O. Arroyo, "Machine learning–based approach for predicting pushover curves of low-rise reinforced concrete frame buildings," *Structures*, vol. 70, Art. no. 107694, Dec. 2024. https://doi.org/10.1016/j.istruc.2024.107694

[14].S. Wang, X. Cheng, Y. Li, X. Song, R. Guo, H. Zhang, and Z. Liang, "Rapid visual simulation of the progressive collapse of regular reinforced concrete frame structures based on machine learning and physics engine," *Eng. Struct.*, vol. 286, Art. no. 116129, Jul. 2023. https://doi.org/10.1016/j.engstruct.2023.116129.

[15].S. Wang, X. Cheng, Y. Li, X. Yang, H. Zhang, R. Guo, and Z. Liang, "Assessing progressive collapse regions of reinforced concrete frame structures using graph convolutional networks," *Eng. Struct.*, vol. 322, Art. no. 119076, Jan. 2025, https://doi.org/10.1016/j.engstruct.2024.119076

[16].Y. F. Zhu, Y. Yao, Y. Huang, C.H. Chen, H.Y. Zhang, and Z. Huang, "Machine learning applications for assessment of dynamic progressive collapse of steel moment frames," *Structures*, vol. 36, pp. 927–934, Feb. 2022. https://doi.org/10.1016/j.istruc.2021.12.067

[17].H. D. Nguyen, N.D. Dao, and M. Shin, "Machine learning-based prediction for maximum displacement of seismic isolation systems," *J. Build. Eng.,* vol. 51, Art. no. 104251, Jul. 2022. https://doi.org/10.1016/j.jobe.2022.104251

[18].H. D. Nguyen, J.M. LaFave, Y.J. Lee, and M. Shin, "Rapid seismic damage-state assessment of steel moment frames using machine learning," *Eng. Struct.,* vol. 252, Art. no. 113737, Feb. 2022. https://doi.org/10.1016/j.engstruct.2021.113737

[19].E. Junda, C. Malaga-Chuquitaype, and K. Chawgien, "Interpretable machine learning models for the estimation of seismic drifts in CLT buildings," *J. Build. Eng.,* vol. 70, Art. no. 106365, Jul. 2023. https://doi.org/10.1016/j.jobe.2023.106365

[20].S. H. Hwang, S. Mangalathu, J. Shin, and J. S. Jeon, "Machine learning-based approaches for seismic demand and collapse of ductile reinforced concrete building frames," *J. Build. Eng.,* vol. 34, Art. no. 101905, Feb. 2021. https://doi.org/10.1016/j.jobe.2020.101905

[21].H. Luo and S.G. Paal, "Machine learning–based backbone curve model of reinforced concrete columns subjected to cyclic loading reversals," *J. Comput. Civ. Eng.,* vol. 32, no. 5, Art. no. 04018042, Sep. 2018. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000787

[22].H. Pak and S.G. Paal, "Evaluation of transfer learning models for predicting the lateral strength of reinforced concrete columns," *Eng. Struct.,* vol. 266, Art. no. 114579, Sep. 2022. https://doi.org/10.1016/j.engstruct.2022.114579

[23].J. M. Rodríguez, A.M. Esteban, M.V. Cruz, Z.B. Blanco, M.L. Verjel, E.R. Sanchez, and J.M. Estevao, "Fast seismic assessment of built urban areas with the accuracy of mechanical methods using a feedforward neural network," *Sustainability*, vol. 14, no. 9, Art. no. 5274, Apr. 2022. https://doi.org/10.3390/su14095274

[24].Y. Gan, J. Chen, Y. Li, and Z. Xu, "Prediction of progressive collapse resistance of RC frames using deep and cross network model," *Structures*, vol. 51, pp. 800–813, May 2023. https://doi.org/10.1016/j.istruc.2023.03.087

[25].M. J. Esfandiari, H. Haghighi, and G. Urgessa, "Machine learning-based optimum reinforced concrete design for progressive collapse," *Electron. J. Struct. Eng.,* vol. 23, no. 1, pp. 1–8, Mar. 2023. https://doi.org/10.56748/ejse.233642

[26].F. Fu, "Fire induced progressive collapse potential assessment of steel framed buildings using machine learning," J. *Constr. Steel Res.,* vol. 166, Art. no. 105918, Mar. 2020. https://doi.org/10.1016/j.jcsr.2019.105918

[27].M. J. Esfandiari and G.S. Urgessa, "Progressive collapse design of reinforced concrete frames using structural optimization and machine learning," *Structures*, vol. 28, pp. 1252–1264, Dec. 2020. https://doi.org/10.1016/j.istruc.2020.09.039

[28].J. Rezania, M. Hamian, and A. Rasekhi, "Investigating the Application of Artificial Intelligence in Civil Engineering and Progressive Collapse," *Civil Proj. J.*, vol. 5, pp. 11–22, Dec. 2023. https://doi.org/10.22034/cpj.2023.426006.1233

[29].D. Li and C. Zhen, "Progressive collapse response and ultimate strength evaluation of stiffened plates with welding residual stress under combined biaxial cyclic loads and lateral pressure," *Mar. Struct.*, vol. 99, Art. no. 103703, Jan. 2025. https://doi.org/10.1016/j.marstruc.2024.103703

[30].A. G. Padilha, L.F. Fadel, L.R. Holdorf, and A.T. Beck, "Reliability assessment of guyed transmission towers through active learning metamodeling and progressive collapse simulation," *Struct. Infrastruct. Eng.,* vol. 20, no. 5, pp. 715–729, May 2024. https://doi.org/10.1080/15732479.2022.2122516

[31].Computer and Structures Inc., *SAP2000 Structural Analysis Program*, ver. 23, Berkeley, CA, USA, 2024.

[32].Turkish Standards Institution, *TS-500(2000), Requirements for Design and Construction of Reinforced Concrete Structures,* Ankara, Türkiye, Feb. 2000.

[33]. ACI Committee 318, *Building Code Requirements for Structural Concrete*, ACI 318-19, American Concrete Institute, Farmington Hills, MI, USA, 2019.

[34].Turkish Disaster and Emergency Management Authority, *Turkish Building Earthquake Code*, TBEC-2018, Ankara, Türkiye, Mar. 2018.

[35].Federal Emergency Management Agency, *Prestandard and Commentary for the Seismic Rehabilitation of Buildings*, FEMA 356, Washington, DC, USA, Nov. 2000.

[36].Imbsen & Associates, *XTRACT Cross-Section Analysis Program*, ver. 3.0.8, Rancho Cordova, CA, USA, 2025.

[37].M. Pokhrel and M.J. Bandelt, "Plastic Hinge Behavior and Rotation Capacity in Reinforced Ductile Concrete Flexural Members," *Eng. Struct.,* vol. 200, Art. no. 109699, Dec. 2019. https://doi.org/10.1016/j.engstruct.2019.109699

[38].Ç.B. Erdaş and E. Sümer, "A deep learning method to detect parkinson's disease from MRI slices," *SN Comput. Sci.*, vol. 3, no. 2, Art. no. 120, Jan. 2022. https://doi.org/10.1007/s42979-022-01018-y

[39].Ç.B. Erdaş, E. Sümer, and S. Kibaroğlu, "Neurodegenerative diseases detection and grading using gait dynamics," *Multimed. Tools Appl.*, vol. 82, no. 15, pp. 22925–22942, Feb. 2023. https://doi.org/10.1007/s11042-023-14461-7