May 2026
Volume 2
Issue 2

DOI: 10.48308/ijce.2026.246372.1029

Interdisciplinary Journal of Civil Engineering
Shahid Beheshti University

Prediction of the Backbone Curve for Bolted Extended Endplate Moment Connections Using Machine Learning Methods

Authors:

Matin Alizadeh[1], Mahmoud R. Shiravand[1,*], Amir Saedi Daryan[1]

Abstract

This study develops an interpretable artificial intelligence framework for predicting the nonlinear behavior and backbone curve of Bolted Extended End-Plate (BEEP) connections under seismic loading. An analytical-developmental approach based on machine learning was adopted, using an enriched database of BEEP connections to establish relationships between connection parameters and nonlinear moment–rotation responses. Several machine-learning algorithms were investigated, with particular emphasis on ensemble-learning methods, including Random Forest (RF), Extreme Gradient Boosting (XGBoost), and Light Gradient Boosting Machine (LightGBM). The developed models were evaluated using standard statistical performance metrics, including the coefficient of determination (R^2), root mean square error (RMSE), and mean absolute error (MAE). The results demonstrate that the developed ensemble-learning models achieved high predictive accuracy on unseen test data, with R^2 values of 0.95 or higher. This indicates that the proposed models can effectively capture the nonlinear characteristics of BEEP connection behavior and accurately reproduce their moment–rotation backbone curves. Furthermore, interpretability techniques were employed to quantify the contribution and relative importance of input variables to the predicted response. The analysis identified the governing connection parameters influencing nonlinear behavior, providing valuable insight into the mechanics of BEEP connections. Overall, the proposed framework provides an accurate, interpretable, and practical data-driven methodology for backbone-curve prediction and seismic performance assessment of BEEP connections.

Keywords: Bolted Endplate Moment Connection, Machine Learning, Ensemble Learning, Backbone Curve, Non-Linear Behavior

1. Faculty of Civil, Water and Environmental Engineering, Shahid Beheshti University, Tehran, Iran

*Corresponding Author: m_shiravand@sbu.ac.ir

1. Introduction

Bolted extended endplate (BEEP) moment connections are a prevalent choice in steel Moment Resisting Frames (MRFs) due to their ease of fabrication and trusted seismic performance, as indicated in many studies (Adey et al., 1997; Adey et al., 2000; Astaneh-Asl, 1995; Agerskov, 1977). Three configurations that based on several seismic design codes considered as prequalified connections are shown in figure 1. A bolted extended endplate moment connection consists of an end plate to which the beam is connected in the fabrication shop using either a complete-joint-penetration groove weld or double-sided fillet welds. The end plate is then connected to the column flange using high-strength bolts. Based on the number of bolts in the tension region, this type of connection is classified into two categories: 4-bolted and 8-bolted connections. However, their behavior under severe cyclic loading is highly nonlinear and complicated. Accurate prediction of these behaviors is essential for reliable performance-based seismic design. The response is governed by the interaction of multiple components, including endplate yielding, prying action, bolt elongation and fracture, and yielding or local buckling in the beam and column panel zone. Traditional analytical methods, including prescribed expressions in design codes like Eurocode 3 and AISC 358, often fail to accurately predict the non-linear behavior of these kinds of engineering problems. While effective for monotonic loading, they struggle to capture the complex details of cyclic degradation, strength deterioration, and component interaction that define the connection's true nonlinear response and ultimate failure modes. Advancements in computer processing power have enabled researchers to tackle problems that were once difficult to address, allowing for more realistic and precise simulations. Furthermore, advanced ML models enable quick and accurate estimation of complicated connection behavior, offering a significant reduction in computational time compared to conventional finite element simulations.

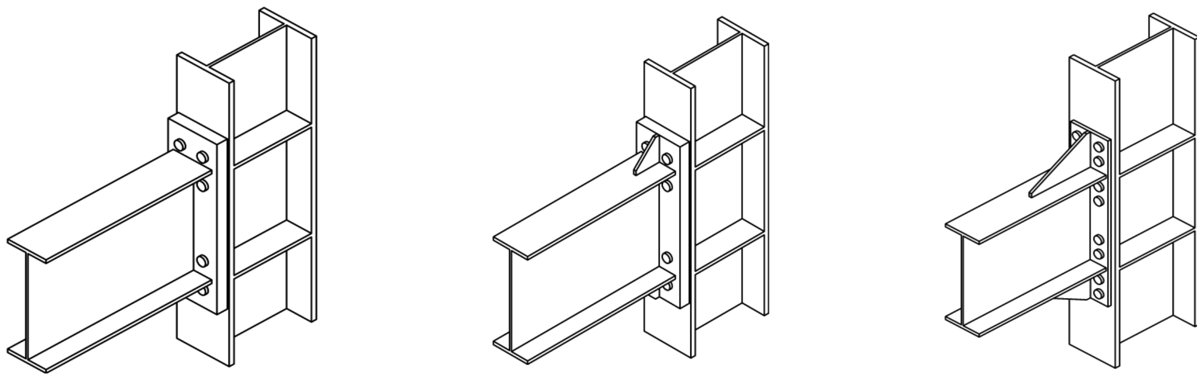


Figure 1. Three main configurations of Bolted Extended Endplate moment connection.

Traditional analytical methods for analyzing and designing steel structural connections often involve tedious, time-consuming hand calculations that inherently have a significant risk of error, potentially leading to misinformed design outputs due to the numerous factors. Overcoming this problem, numerical methods such as Finite Element Analysis (FEA) are often employed, as this approach typically yields a more realistic characterization of structural phenomena, such as the behavior of a steel connection. However, due to the high degree of nonlinearity inherent in these problems, stemming from material yielding, geometric changes, and contact mechanics, FEA also presents considerable challenges, specifically high computational cost and long simulation times. The mentioned problems often limit researchers from using this method in large-scale parametric studies and routine design procedures.

To bridge this gap, data-driven methods offer a powerful alternative. This study presents a framework that integrates a rich database with interpretable ensemble machine learning to predict the nonlinear backbone curve of BEEP connections. A backbone curve is defined as the envelope of the hysteretic action-deformation response from a cyclic test, meaning the peak points in each cycle of the hysteretic curve obtained from load-displacement response are connected to form the "backbone" curve as shown in Figure 2.

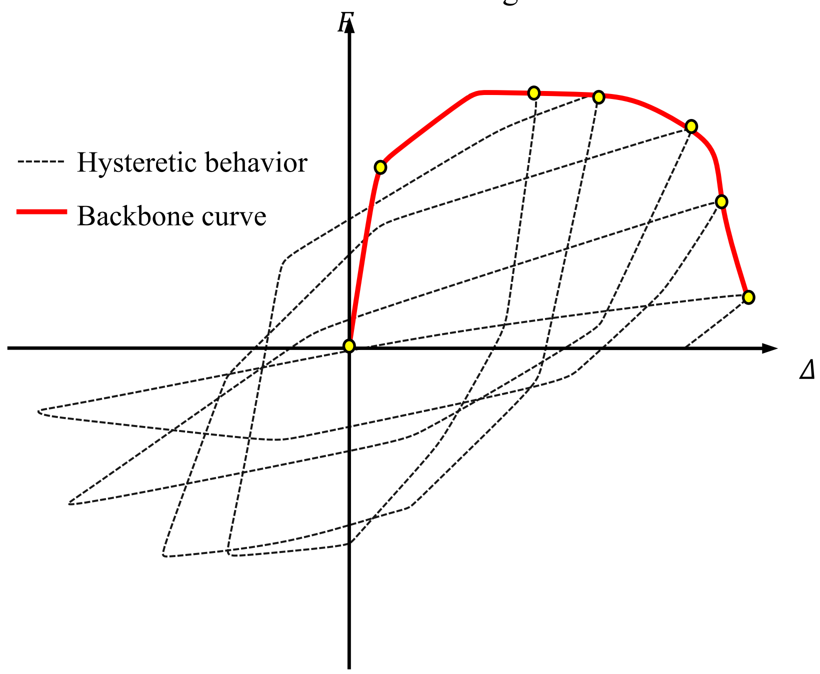


Figure 2. Backbone representation of hysteretic behavior.

The backbone curve provides critical data points such as initial stiffness, yield strength, ultimate moment capacity, and rotation capacity needed for reliable performance-based seismic design and assessment (Federal Emergency Management Agency [FEMA], 2005). This approach uses SHAP to make the model's prediction and behavior more comprehensive and easier to interpret, assuring the model's predictions are consistent with physics-based engineering principles. The study presents a complete workflow including data collection, model development, and the creation of a reliable AI-driven predictive tool for practical engineering applications. The overall workflow of the presented study is shown in Figure 3.

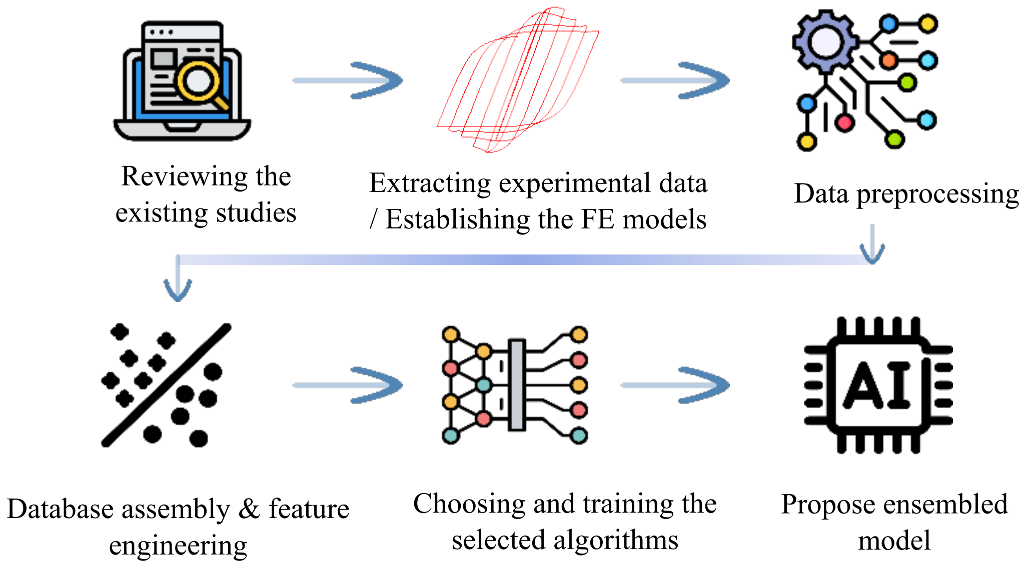


Figure 3. Overall workflow of this study.

2. Literature Review

2.1. Experimental and Analytical Studies on Bolted Extended End-Plate Connections

Research on bolted end-plate moment connections has evolved considerably over the past several decades, encompassing analytical formulations, experimental investigations, component-based mechanical models, and, more recently, studies addressing seismic performance. The development of these connections has primarily focused on understanding the mechanisms governing stiffness, strength, ductility, prying action, plastic hinge formation, and failure. Particular attention has been devoted to the behavior of the end plate and column flange in the tension region, as well as to the influence of geometric parameters such as end-plate thickness, bolt diameter and spacing, bolt configuration, and the presence of stiffeners. Although the precise identification of the first published investigation of bolted end-plate connections is difficult based on the available literature, (Disque, 1962) is widely cited as one of the earliest researchers to formally investigate bolted end-plate connection behavior.

A particularly important contribution to the understanding of seismic behavior was provided by (Ashtaneh-Asl, 1995), who conducted cyclic tests on unstiffened four-bolt extended end-plate moment connections. Although the specimens had been designed according to the AISC recommendations available at the time rather than specifically for seismic applications, the first specimen exhibited satisfactory ductility, with local buckling of the beam flange governing the observed failure mechanism. In a second test, an I-shaped shim was introduced between the end plate and column with the objective of improving the connection response. The modified specimen performed satisfactorily until yielding of the shim under compression. The study demonstrated the potential of modifying the force-transfer mechanism at the connection interface to improve seismic behavior, while also indicating the need for adequate strength of the shim to fully exploit this concept

(Sumner & Murray, 2002) experimentally investigated the cyclic behavior of these connections under reversed cyclic loading. Seven specimens were tested, including six four- and eight-bolt beam-to-column connections and one four-bolt end-plate connection incorporating a concrete slab on the beam. The experimental program was designed to evaluate the influence of the concrete slab on the flexural performance of the connection and to compare the load-carrying behavior of four- and eight-bolt configurations. The results indicated that the presence of the concrete slab increased the moment demand by approximately 27% compared with the corresponding connection without a concrete slab.

A major challenge in the design of bolted end-plate connections is the large number of possible connection configurations, which makes it difficult to establish a unified design procedure capable of accounting for their different geometric characteristics and behavioral mechanisms. (Sumner, 2003) addressed this problem by developing a comprehensive design methodology that could be applied to a wide range of bolted end-plate connection configurations. The proposed methodology covered eight different connection geometries and established a systematic framework for evaluating the key mechanical characteristics of the connections, including stiffness, strength, and ductility. Instead of treating each connection configuration as an independent design case, the proposed approach represented the behavior in terms of the principal connection components and their governing limit states. This systematic methodology provided a more consistent basis for the design and assessment of bolted end-plate connections and subsequently contributed to the refinement of seismic design provisions, influencing the development and application of provisions incorporated into AISC standards.

(Guo et al., 2006) investigated the cyclic response of six four-bolt beam-to-column connections with different end-plate thicknesses and configurations, including specimens with and without stiffeners. The experimental results demonstrated that the governing failure mechanism was strongly dependent on the relative strength of the connection components. Stronger configurations developed plastic hinges in the beam, whereas weaker configurations exhibited significant panel-zone rotation or failure of bolts and complete-joint-penetration welds. The introduction of end-plate stiffeners substantially increased the connection resistance and facilitated the development of beam plastic hinges without requiring excessively thick end plates. These findings further demonstrated the importance of controlling the relative strength of the connection components to achieve a desirable seismic failure mechanism.

For the purpose of controlling the location of plastic hinge formation, (Stratan et al., 2017) investigated an alternative approach by incorporating haunches near the intended plastic hinge region of end-plate beam-to-column connections. The specimens were subjected to cyclic loading in accordance with European seismic design provisions. The results demonstrated that haunch geometry could effectively influence the location and development of plastic hinging.

Plastic hinge initiation occurred at drift ratios of approximately 0.04 for haunches inclined at 35° and approximately 0.035 for 45° haunches. The observed response indicated that the investigated configurations could satisfy the seismic performance requirements associated with the EQUALJOINTS framework.

(Morrison & Qayyum, 2019) addressed the limitations associated with the use of prequalified stiffened eight-bolt extended end-plate connections for large rolled beam sections exceeding approximately 600 mm in depth. Their experimental investigation explored the possibility of achieving satisfactory seismic performance without conventional end-plate stiffeners. The authors employed a circular arrangement of bolts above and below the beam flanges to achieve a more uniform distribution of forces within the bolt group. In addition, localized heating of the beam flanges was used to promote plastic hinge formation at a predetermined location, while reinforcement of the beam web was introduced to prevent an undesirable loss of resistance. Cyclic testing confirmed the ability of the proposed configuration to develop a plastic hinge at the intended location, demonstrating a potential alternative for accommodating large beam sections within seismic moment-resisting systems.

More recently, (Ozkilic, 2023) conducted an extensive experimental investigation of beam-to-column connections incorporating thin, unstiffened end plates and large-diameter high-strength bolts. Six specimens with different end-plate thicknesses were subjected to both monotonic and cyclic loading to investigate the governing failure mechanisms and determine whether the response could be controlled by the flexural capacity of the end plate. Particular attention was given to the potential of large-diameter bolts to reduce the pinching of hysteretic responses, which is an important consideration in the seismic assessment of bolted connections. The experimental results demonstrated that connections employing relatively thin end plates could achieve both substantial strength and ductility. These findings indicate that, when appropriately proportioned, thin end-plate configurations may provide a viable solution for seismic moment-resisting frames while reducing the material requirements traditionally associated with thicker end plates.

(Wang et al., 2025) investigated the seismic response of two small-scale beam-to-column specimens incorporating stiffened end plates through reversed cyclic loading. The specimens were assembled in a cruciform configuration, and lateral loading was applied at the top of the column to evaluate hysteretic behavior, ductility, energy dissipation, and the development of yielding and failure within the connection components. The two specimens exhibited distinct governing failure mechanisms. The first specimen experienced failure associated with shear of the column web, whereas the second developed a plastic hinge at the beam end. Despite these differences, both specimens demonstrated favorable hysteretic behavior and satisfactory seismic performance. The results further emphasized the importance of the relative strength and deformation capacity of individual connection components in determining the global failure mechanism and seismic response of the connection.

2.2. Studies on the Application of Machine Learning to the Prediction of Steel Connection Behavior

One of the early applications of artificial neural networks to the prediction of steel connection behavior was reported by Abdalla and (Stavroulakis, 1995), who addressed the difficulty of accurately modeling the nonlinear behavior of semi-rigid steel connections. To overcome the limitations associated with conventional analytical representations, they employed a back-propagation neural network to model the behavior of semi-rigid connections. Their work demonstrated the potential of artificial neural networks as a data-driven approach for representing the complex mechanical response of structural connections. Building upon this concept, (Stavroulakis et al., 1997) addressed the problem of predicting the mechanical response of semi-rigid steel connections and establishing their moment–rotation relationships. They developed an artificial neural network model trained to learn the relationship between the connection parameters and its mechanical response. The results demonstrated that the neural network could successfully predict the behavior of semi-rigid steel connections and provide their corresponding moment–rotation relationships.

(Shah et al., 2018) addressed the need for reliable predictive models capable of estimating the moment–rotation behavior of steel connections without requiring extensive analytical calculations. To develop such models, they experimentally investigated 32 unbolted steel connections and established a database from the resulting experimental data. Three artificial intelligence approaches, namely genetic linear programming, artificial neural networks, and adaptive neuro-fuzzy inference systems, were then employed to predict the moment–rotation behavior. The study demonstrated the potential of these data-driven approaches for accurately representing the mechanical response of steel connections based on experimental data.

(Horton et al., 2021) focused on the more challenging problem of predicting the complete cyclic response of steel connections with reduced beam sections (RBS). Since the hysteretic behavior of such connections involves complex nonlinear mechanisms, including stiffness degradation, strength deterioration, and energy dissipation, conventional simplified models may not fully capture the actual response. To address this problem, the authors developed deep neural networks (DNNs) capable of learning the cyclic behavior of RBS connections from finite element simulations and experimental data. The trained models were able to reproduce the hysteretic response with very low prediction errors, including errors of less than 5% for maximum force and energy dissipation, while requiring only a fraction of the computational time associated with conventional numerical analyses. The study demonstrated that deep learning can provide a rapid and accurate alternative to conventional analytical and numerical approaches, while also offering the possibility of application to connections with different geometries and material properties after appropriate retraining.

(Mohammadnia & Moradi, 2022) addressed the need for an efficient method for predicting the backbone response of end-plate connections equipped with shape memory alloy (SMA) bolts, which can provide self-centering behavior to the connection. To overcome the limitations associated with repeatedly conducting finite element analyses or experimental tests, they developed an artificial neural network-based prediction tool. The network was trained using data from 72 finite element specimens and 7 experimental specimens. The developed model achieved coefficients of determination ranging from 0.92 to 0.99, indicating a high level of predictive accuracy in reproducing the backbone response of the investigated connections.

For the purpose of optimization problems of designing steel connections, (Grubits et al., 2025) developed an interpretable computational framework that combined finite element analysis with genetic algorithms to identify optimized end-plate connection configurations. The optimization procedure produced a configuration that achieved a 10.4% improvement in the overall performance objective compared with the best verified model and reduced material volume by 39.3% compared with the best validated model. Furthermore, a 53.6% reduction in equivalent plastic strain was obtained compared with the configuration exhibiting the highest level of plastic deformation. These results demonstrated the potential of advanced computational and optimization techniques for improving both the efficiency and performance of structural connection design.

Overall, the development of machine-learning applications for steel connections has progressed from the early use of artificial neural networks for representing moment–rotation behavior to more advanced approaches capable of predicting complete cyclic responses and optimizing connection configurations. These studies have demonstrated that machine-learning models can capture highly nonlinear relationships between connection parameters and structural response without explicitly defining all of the underlying mechanical relationships. In particular, the increasing availability of experimental and numerical databases has created opportunities for developing predictive models that can complement conventional analytical and numerical methods.

3. Methodology

3.1. Data Extracted from Previous Studies

The proposed framework is built upon a robust, physically grounded database, which is then used to train and validate the interpretable ML models. The methodology is divided into database generation, feature engineering, and the ML training pipeline as shown in Figure 3. A comprehensive database of bolted extended end-plate moment connections was developed by combining data extracted from previously reported experimental tests with additional numerical data generated through finite element (FE) analyses. The experimental database was compiled from a wide range of previously conducted studies, covering different connection configurations, geometric properties, material characteristics, and bolt arrangements. For each experimental specimen, the corresponding moment–rotation response was collected and the backbone curve was systematically extracted from the reported test results. Data extraction procedure have been shown in Figure 4.

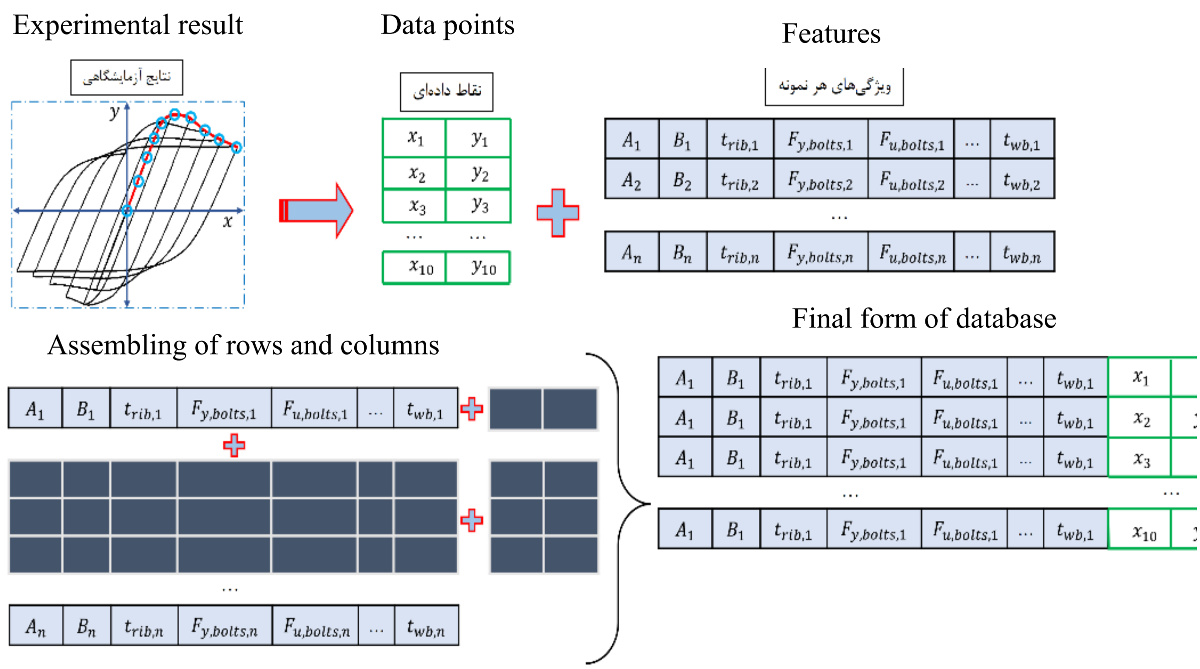


Figure 4. Data extraction procedure

The required data were extracted from the response curves reported in the selected experimental studies using a high-precision, pixel-based data extraction software. To minimize data loss and ensure accuracy, the extraction was performed manually. For each specimen, ten representative points were carefully selected from the cyclic response curve, with emphasis on points defining the backbone curve. More points were selected in critical regions associated with yielding, softening, and other significant changes in response, while fewer points were used in approximately linear regions. The main limitation during this process was the relatively low raster quality of some graphs reported in the literature. The extracted data were then incorporated into the main database, with ten rows generated for each specimen, corresponding to the ten selected points and retaining the associated geometric and material properties. Table 1 represents the studies that have been used to constructed the experimental database of this article.

3.2. Development of Numerical Specimens

To extend the range of connection configurations and provide a broader dataset for machine learning, additional numerical specimens were developed using the finite element software ANSYS. The numerical database was generated by systematically varying the principal geometric and material parameters governing the behavior of bolted extended end-plate connections. The resulting FE analyses provided additional moment–rotation responses and backbone curves, thereby complementing the experimentally extracted data and increasing the size and diversity of the database. Before incorporating the numerical results into the final database, the reliability of the developed FE modeling approach was evaluated through validation against available experimental results. The numerical predictions were compared with corresponding experimental responses in terms of the moment–rotation behavior, strength, stiffness, and deformation capacity. The details of the FE modeling procedure and the validation of the numerical results are presented in the subsequent sections.

Table 1. Summary of Experimental Studies and Variables Included in the Database

No.	Author(s)	Number of Specimens	Variables
1.	(Ghobarah et al., 1992)	5	Column web thickness, end-plate stiffener, bolt diameter
2.	(Sumner & murray, 2002)	6	Plate thickness, number of bolt rows
3.	(Shi et al., 2010)	6	End-plate thickness, bolt diameter
4.	(Morrison et al., 2017)	8	End-plate thickness
5.	(Ozkilic, 2023)	6	End-plate thickness
6.	(Wang et al., 2025)	2	Column web thickness

3.2.1. Numerical Modeling Process

The numerical models required for generating the training data were developed using the commercial finite element software ANSYS. The software was selected because of its advanced element library and its capability to account for the three major sources of nonlinearity in bolted end-plate connections: material, geometric, and contact nonlinearities. These capabilities are particularly important for simulating phenomena such as plate yielding, bolt failure, local buckling, separation, and sliding between connection components; so three sources of nonlinearity have been taken to account for this purpose.

To simulate the cyclic nonlinear behavior of steel, the combined hardening model was employed, incorporating both kinematic and isotropic hardening. This model is capable of reproducing the evolution of the yield surface and the plastic response of steel under cyclic loading. The model was originally formulated based on the Armstrong–Frederick kinematic hardening rule and subsequently developed by Chaboche. In the adopted model, kinematic hardening is represented by the sum of two back-stress components, which evolve according to the Armstrong–Frederick rule, as expressed in Eq. (1). The first term represents the development of plastic hardening with increasing plastic deformation, while the second term controls the saturation of the back stress and prevents unlimited hardening. Thus, this formulation enables the translation of the yield surface during cyclic loading.

$$d\sigma_{\mathrm{ij}} = \frac{\mathrm{C}}{\sigma_0}\left(\sigma_{\mathrm{ij}} - \alpha_{\mathrm{ij}}\right)d\varepsilon^{\mathrm{p}} - \gamma\alpha_{\mathrm{ij}}\varepsilon^{\mathrm{p}} \qquad (1)$$

The isotropic hardening component is described using an exponential saturation law, as given in Eq. (2). This relation governs the change in the size of the yield surface as plastic deformation accumulates. Therefore, while kinematic hardening controls the translation of the yield surface, isotropic hardening governs its expansion or contraction.

$$\sigma_{0(\varepsilon^{\mathrm{p}})} = \sigma_{\mathrm{init}} + \mathrm{Q}_{\infty}\left[1 - \mathrm{e}^{-\mathrm{b}\times\varepsilon^{\mathrm{p}}}\right] \qquad (2)$$

Materials used for every steel part in the models, excluding the bolts, are variants of St-37 and St-52 grade steels. For the bolts, a bilinear kinematic hardening model was adopted, with a post-yield stiffness equal to 3% of the elastic modulus of steel (E = 210,000 MPa). Bolt grades 8.8 and 10.9 were considered. The corresponding stress–strain curve used for the bolt material is presented in Figure 5.

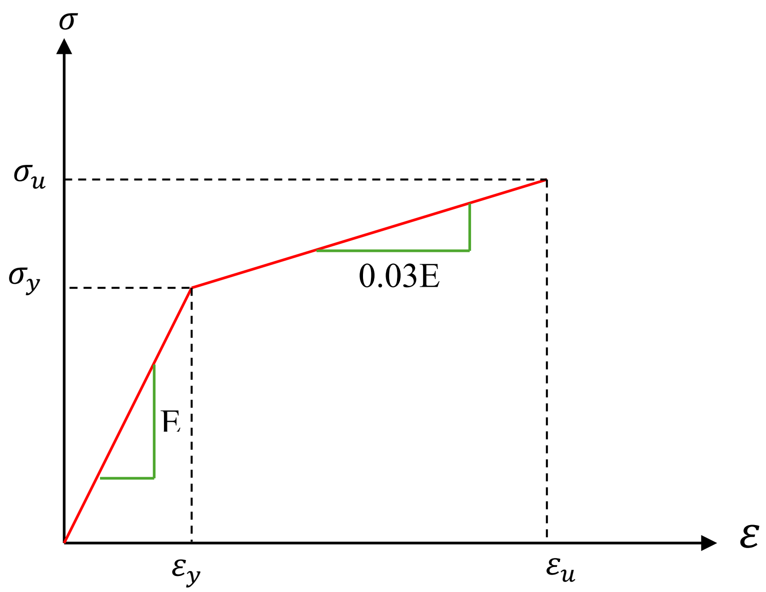


Figure 5. Bilinear stress-strain curve used for modeling bolts.

Due to the occurrence of large plastic deformations in the investigated connection models, a nonlinear analysis capable of updating the stiffness matrix throughout the deformation process was required. Therefore, a Large Deformation analysis was adopted to account for geometric nonlinearity and the effects of significant plastic deformations. To represent the initial geometric imperfections of the structural members, an Eigenvalue Buckling analysis was first performed to identify the critical buckling modes of the connection components. The resulting buckling mode shape was then scaled and introduced into the original geometry as an Initial Geometric Imperfection. This approach provides a more realistic representation of the physical specimens and facilitates the development of actual instability mechanisms, such as local buckling of the beam flange, beam web, and end plate, during the subsequent nonlinear analysis. Figure 6 shows the first mode shape of one of the models for implementing it in the main cyclic loading analysis as a geometric imperfection.

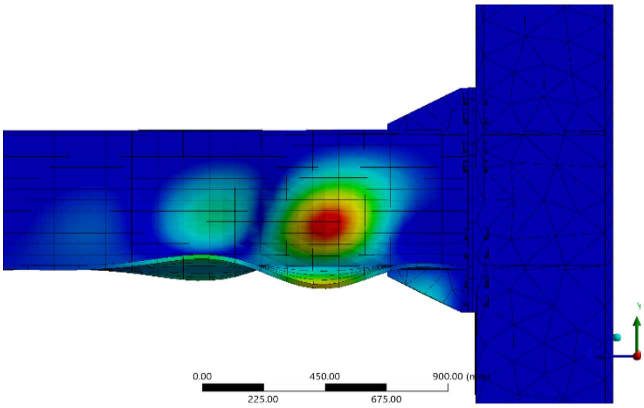


Figure 6. First mode shape for one of the analyzed models

To simulate the interaction between the connection components, surface-to-surface contact elements CONTA174 and TARGE170 were employed to account for separation, sliding, and the pinching effect under cyclic loading. All components, including the beam, column, end plate, and bolts, were modeled using a full three-dimensional solid representation with SOLID186, a 20-node higher-order solid element with three degrees of freedom per node. The bolts were explicitly modeled to accurately capture force transfer and their interaction with the plate holes, while simplified cylindrical bolt heads and nuts were adopted to improve numerical stability and reduce computational cost. This modeling approach enables realistic simulation of stress distribution, local yielding, prying action, and local buckling in the connection.

Figure 7 shows the complete meshed and assembled finite element model of one of the investigated bolted extended end-plate connection, including the beam, column, end plate, and bolts.

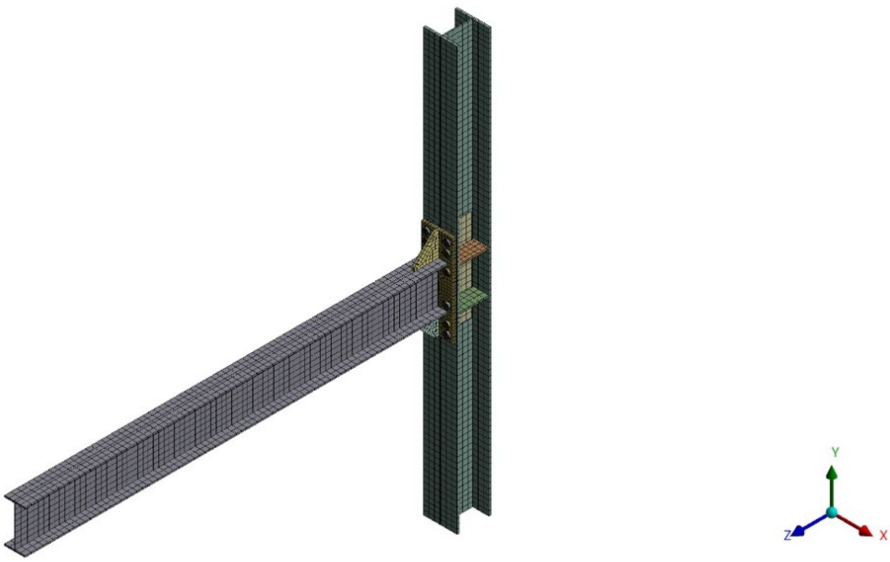


Figure 7. Mesh and assembled geometry of a specimen.

Since the investigated connections are assumed to represent part of a moment-resisting frame, the column length was defined based on the locations of the inflection points, corresponding to half the height of the stories above and below the connection. The upper and lower column ends were modeled as pinned supports, with all translational degrees of freedom restrained and rotation about the principal connection axis released. At the beam ends, the boundary conditions were defined to prevent axial shortening or elongation and maintain symmetry, ensuring that the numerical sub-structure realistically represents the experimental setup and the behavior of a moment-resisting frame. The adopted connection configuration and boundary conditions are illustrated in Figure 8.

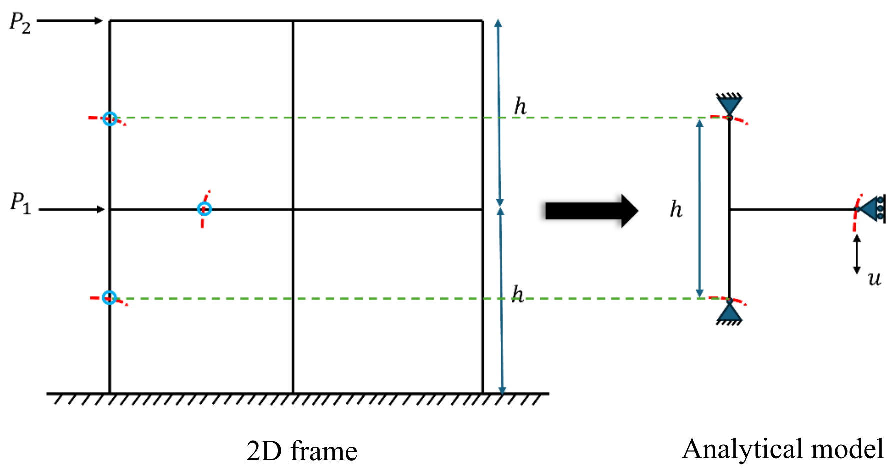


Figure 8. Idealization of the two-story moment-resisting frame and the corresponding analytical sub-structure with boundary conditions

Also, to prevent the lateral-torsional buckling of the models, the beams were provided with lateral restraints at specified locations along the beam length. A displacement-controlled loading protocol was adopted for the numerical models. This loading approach provides improved numerical stability in the nonlinear and post-yield response regions. The load was applied at the free end of the beam, starting from zero and continuing until a maximum interstory drift of 7% was reached.

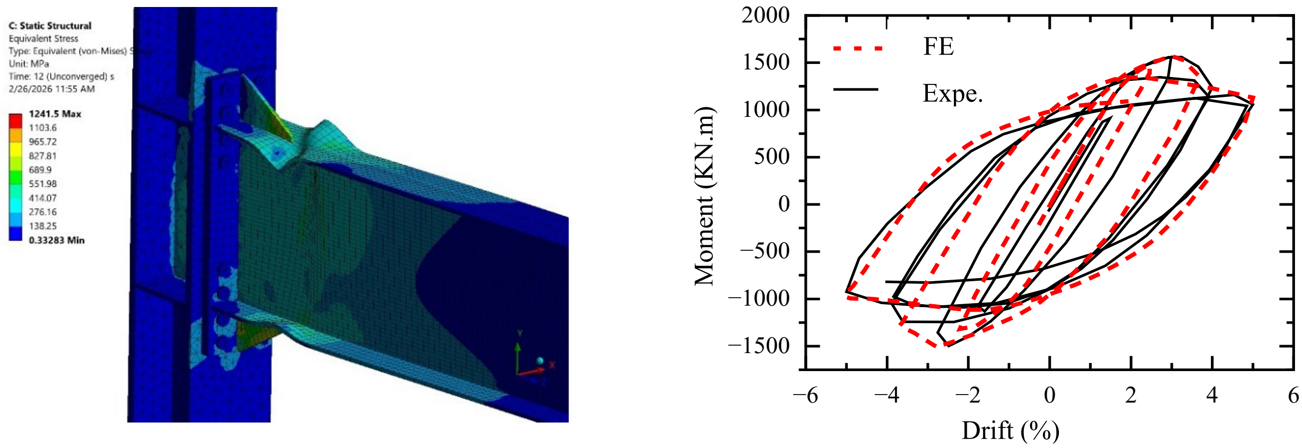
3.2.2. FE Models Verifications

To evaluate the accuracy and reliability of the developed finite element (FE) modeling approach, eight selected experimental specimens described in Table 2 were reproduced numerically and their responses were compared with the corresponding experimental results. The selected specimens represent different connection configurations and behavioral characteristics, providing a suitable basis for assessing the capability of the developed numerical model to capture the nonlinear response of the investigated connections.

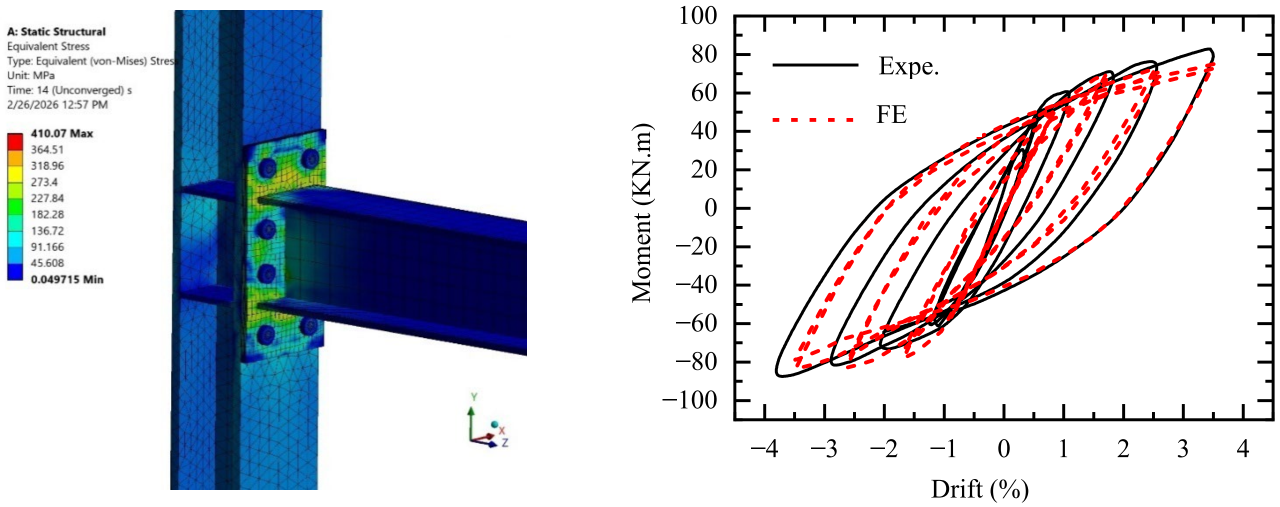
Table 2. Selected experimental tests for verification process

No.	Reference	Selected specimen(s)	Failure mode
1	(Sumner & Murray, 2002)	8ES-1.25-1.75-30	Plastic hinge formation in the beam
2	(Gou et al., 2006)	S1, S2, S5, S6	Plastic hinge formation in the beam and panel-zone rotation
3	(Shi et al., 2010)	JD2, JD3, JD5	Plastic hinge formation in the beam and panel-zone rotation

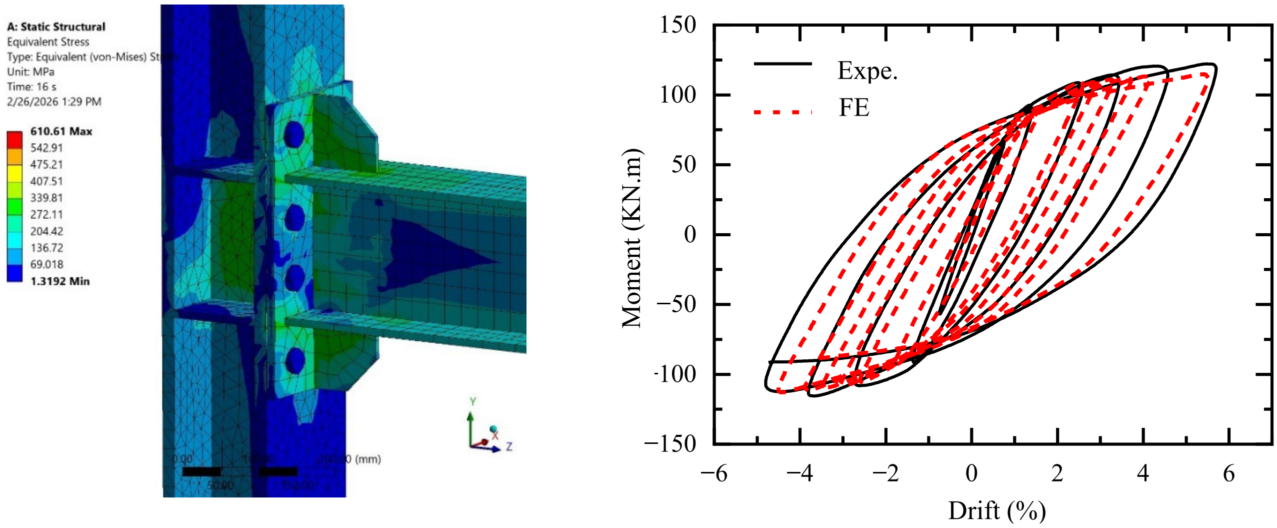
Based on the verification results presented herein, the reliability of the developed numerical models was confirmed, providing a reliable basis for generating the numerical database. Accordingly, the validated FE models were employed to conduct the required numerical analyses under cyclic loading and to generate the additional data required for the subsequent development of the artificial ML models.



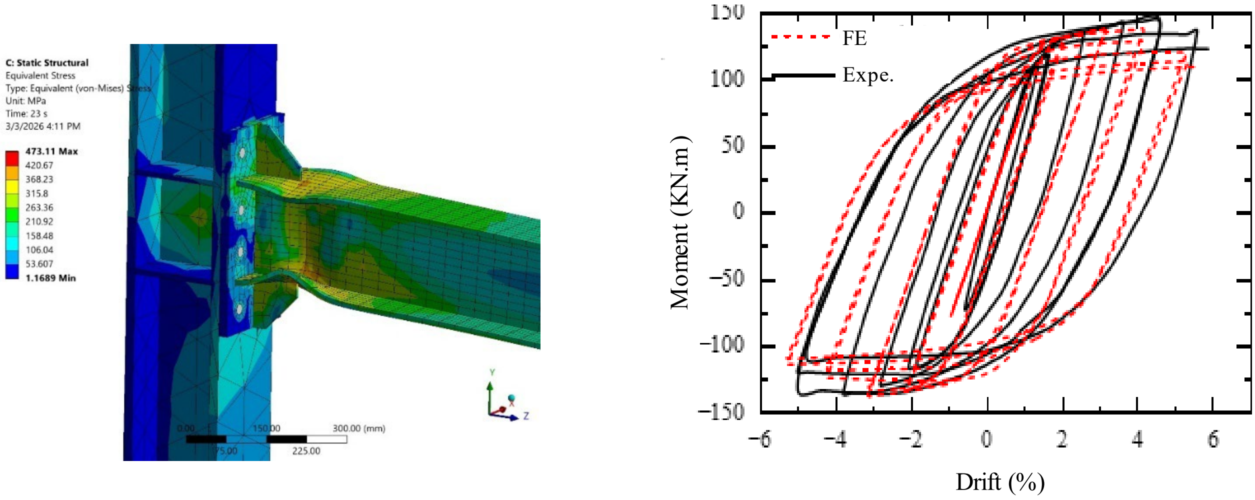
(a)



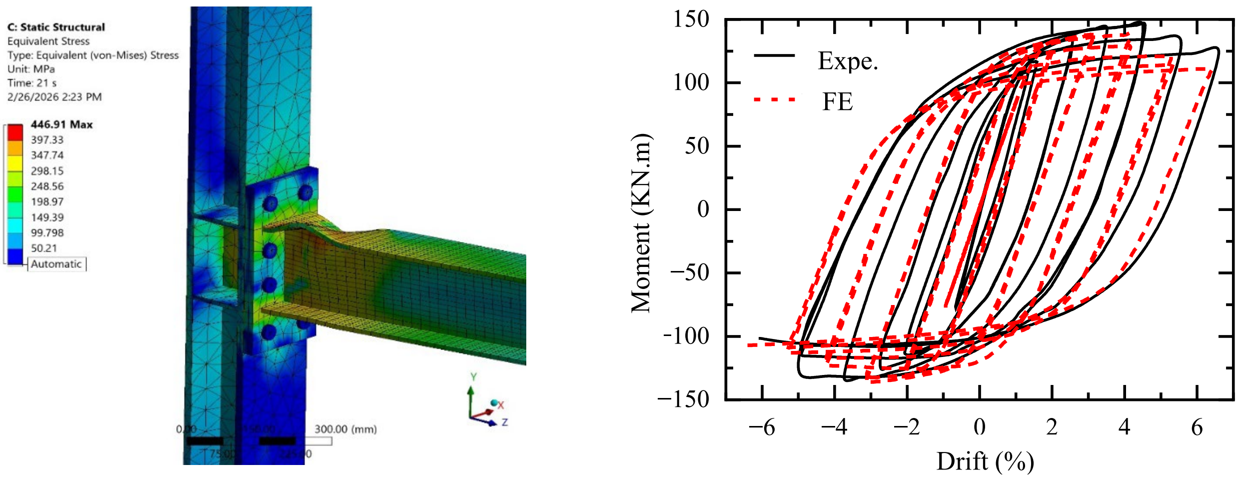
(b)


A: Static Structural
Equivalent Stress
Type: Equivalent (von-Mises) Stress
Unit: MPa
Time: 16 s
2/26/2026 1:29 PM
610.61 Max
542.91
475.21
407.51
339.81
272.11
204.42
136.72
69.018
1.3192 Min
Expe.
FE
Moment (KN.m)
Drift (%)

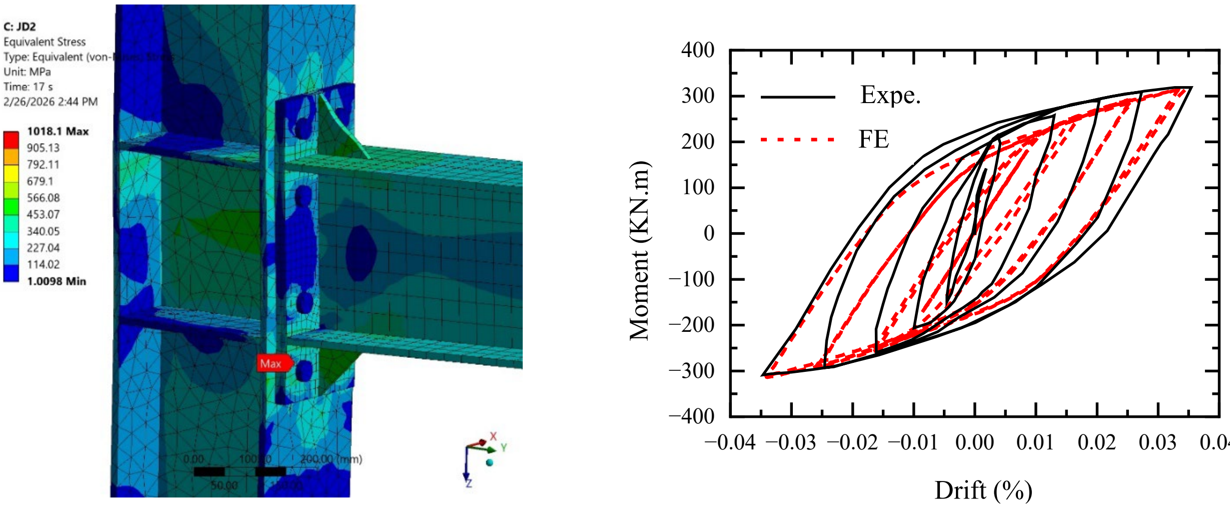
(c)


C: Static Structural
Equivalent Stress
Type: Equivalent (von-Mises) Stress
Unit: MPa
Time: 23 s
3/3/2026 4:11 PM
473.11 Max
420.67
368.23
315.8
263.36
210.92
158.48
106.04
53.607
1.1689 Min
0.00
150.00
300.00 (mm)
75.00
225.00
FE
Expe.
Moment (KN.m)
Drift (%)

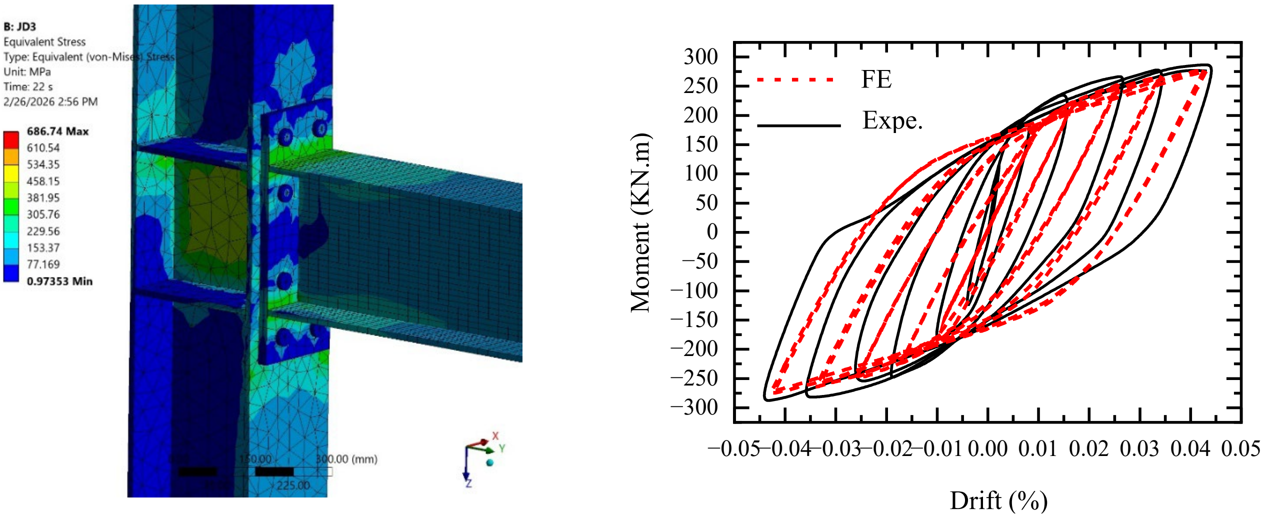
(d)


C: Static Structural
Equivalent Stress
Type: Equivalent (von-Mises) Stress
Unit: MPa
Time: 21 s
2/26/2026 2:23 PM
446.91 Max
397.33
347.74
298.15
248.56
198.97
149.39
99.798
50.21
Automatic
300.00 (mm)
Expe.
FE
Moment (KN.m)
Drift (%)

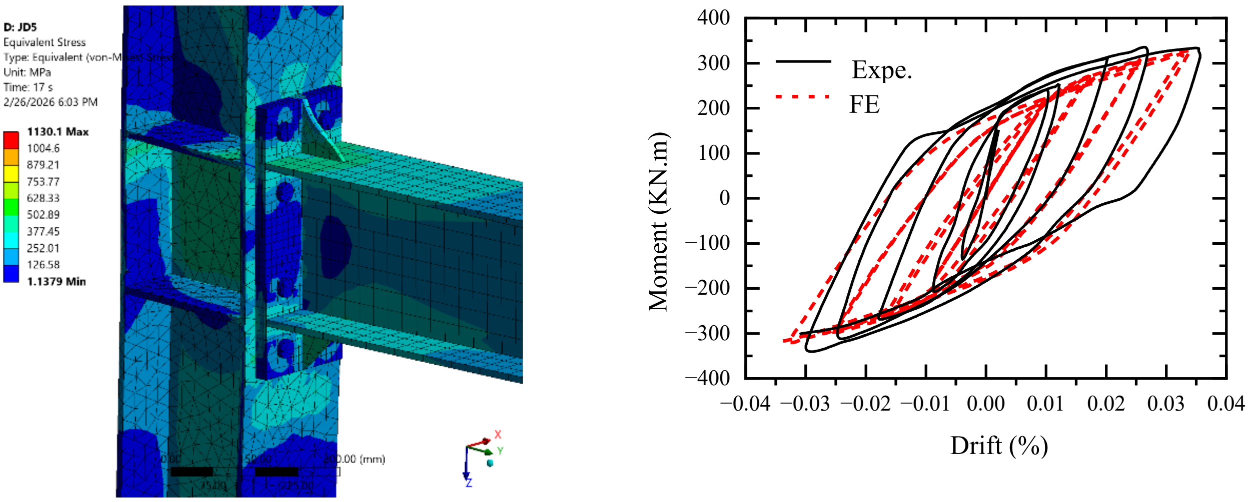
(e)



(f)



(g)



(h)

Figure 9. Experimental and numerical verification of the developed FE models for the selected specimens: (a) Specimen 8ES-1.25-1.75-30, (b) Specimen S1, (c) Specimen S2, (d) Specimen S5, (e) Specimen S6, (f) Specimen JD2, (g) Specimen JD3, and (h) Specimen JD5.

3.3. Feature Selection and Preprocessing

A total of 17 geometric and material parameters were identified as input features for each specimen. These features are listed in Table 3, along with their mean, coefficient of variation, and

standard deviation. In this table, $t_{Rib,}$ $T_{u,bolts,}$ D_{bolts} are thickness of rib stiffeners, bolts pre-tensioning force, bolts diameter, yield and ultimate strength of steel material, respectively The selected features are also illustrated in Figure 10. These features describe the complete geometry of the beam, column, and endplate, as well as the material properties (e.g., $F_{y,plates}$, $F_{u,plates}$) of each component and the bolt properties (e.g., diameter, pretension, F_{ybolts}, $F_{u,bolts}$). The target variable for prediction was the moment at each point along the backbone curve extracted from the hysteretic behavior of the experimentally tested specimens. Although Figure 11 illustrates the relationships among the input features using the Pearson correlation matrix, the results provide an initial assessment of the degree and direction of linear correlation between the features.

Table 3. Features in the training database
(*Dimensions in millimeters; Stress in megapascals***)**

Features	**Mean**	**Standard deviation**	**CV**
A	74	75	1.00
B	123	131	1.06
C	144	150	1.04
t_{Rib}	6	7	1.02
$F_{y,bolts}$	926	105	0.11
$F_{u,bolts}$	1086	114	0.10
$T_{u,bolts}$	258827	164708	0.63
D_{bolts}	25	5	0.21
t_{EP}	33	16	0.48
H_{EP}	703	242	0.34
$F_{y,plates}$	334	40	0.12
$F_{u,plates}$	490	60	0.12
L_b	3042	1113	0.36
H_b	467	175	0.37
W_b	196	43	0.21
t_{fb}	15	4	0.26
t_{wb}	10	3	0.30

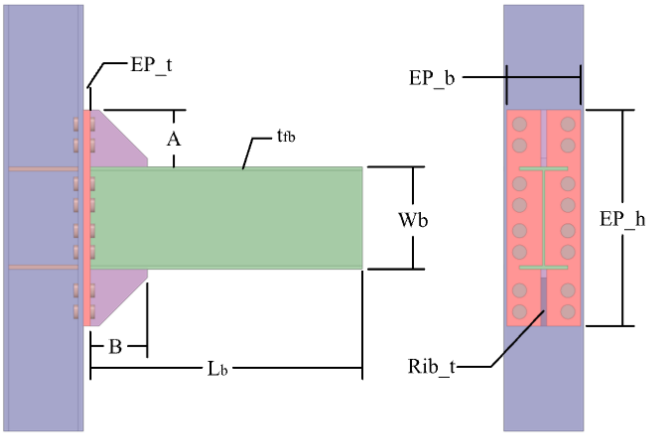


Figure 10. geometric features of the specimens.

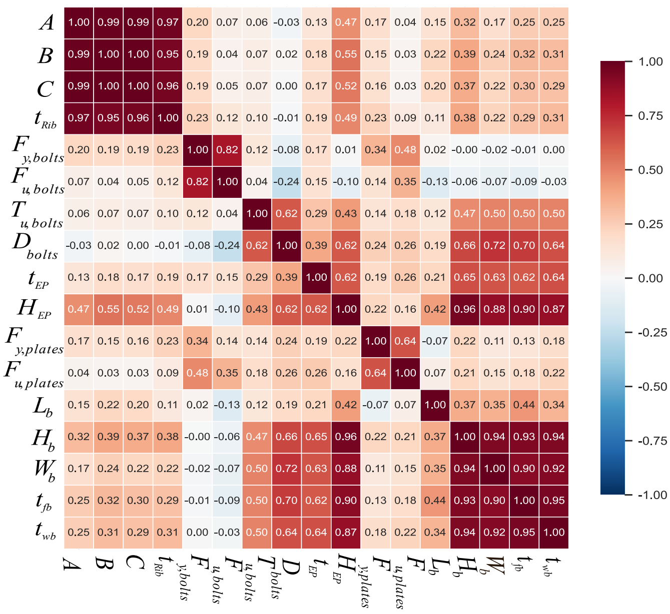


Figure 11. Features Pearson correlation matrix.

3.4. Machine Learning Algorithms and Training

Different ML algorithms were tested, including K-Nearest Neighbors (KNN) and Support Vector Regression (SVR). The primary focus, however, was on tree-based ensemble methods, namely Random Forest (RF), Gradient Boosting (GB), XGBoost, and LightGBM (LGBM), which are well-recognized for their high accuracy and efficiency. In supervised learning, machine learning algorithms are trained using a database consisting of labeled data. Labeled data contain a feature vector (inputs) and a target variable, together with their corresponding labels. The objective of supervised learning algorithms is to identify and fit a relationship between the inputs and the target variable and subsequently predict the target for new input data. Although the accuracy of the final predictions strongly depends on algorithm-related factors, such as proper hyperparameter tuning, the quality of the division of the data into training and testing sets, and the inherent ability of the algorithm to capture the patterns present in the data, the overall performance of the trained model is also significantly influenced by its ability to learn from the available database. Therefore, the availability of a comprehensive, accurate, and well-structured database is essential for achieving reliable predictions. Algorithms have been trained with 2020 data points obtained from tested specimens and numerical modeling; meaning a total of 202 specimens with 10 data points extracted from the backbone curve of the hysteretic response, along with their features, act as a training set to predict the unseen, future result. The dataset was divided into 80% training and 20% testing data to evaluate the predictive performance and generalization capability of the developed models.

4. Results and Discussion

The performance of the trained models was evaluated based on their ability to predict moment values for previously unseen specimens. The ensemble model demonstrated acceptable accuracy in predicting the moment response. Model performance was quantitatively assessed using the coefficient of determination (R^2), Root Mean Squared Error (RMSE), and Mean Absolute Error (MAE). In addition to these numerical metrics, the agreement between the predicted and actual moment values was evaluated by comparing the model predictions for the training and unseen test data. This comparison provides a clear assessment of the model's predictive capability and its ability to generalize to unseen data.

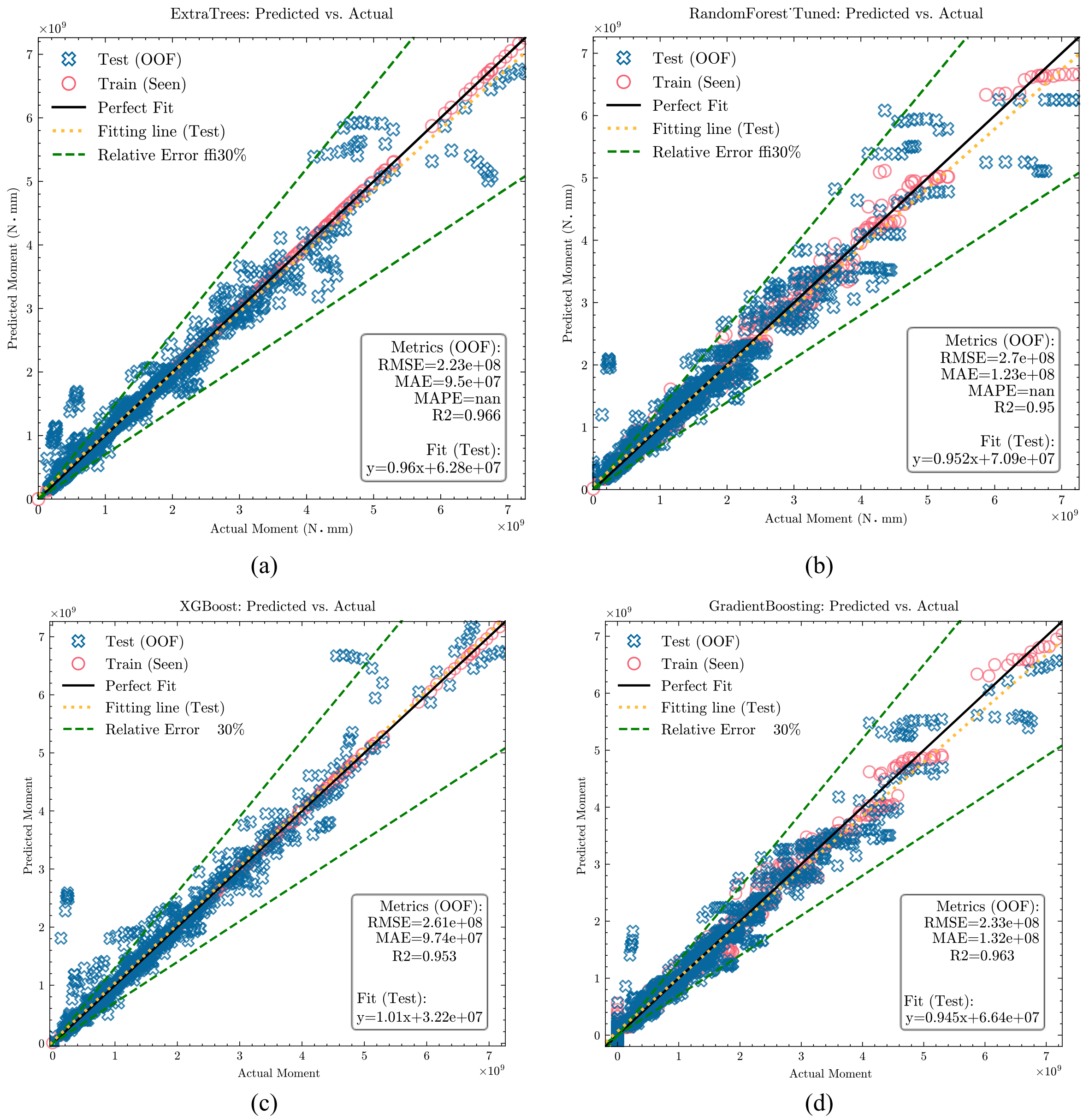


Figure 12. Predicted moment vs. actucal moment for both train and test set for four ML algorithem: (a) ExtraTrees, (b) RandomForest, (c) XGBoost, (4) GradiantBoosting

The four scatter plots in Figure 12 give a full check of the GradientBoosting, RandomForest, XGBoost, and ExtraTrees models. The solid black line on each plot shows a perfect fit, which means that predictions would be exactly equal to actual values. The predicted moment is plotted against the actual moment. The data points, which are divided into training and out-of-fold test sets, are all very close to this perfect line for all four models. These graphs show that the models are very good at making predictions. The metrics show that all of the models are very accurate, and the score R^2 for the test data is above 0.95. Also 0.963, 0.951, 0.953, 0.966 scores for GradientBoosting, RandomForest, XGboost and ExtraTrees, respectively. Additionally, the low RMSE and MAE values (for example, RMSE in the 2.6e+08 - 2.7e+08 range) show that the average error is very small. The fact that most test predictions are well within the 30% relative error limits shows that these models are strong and trustworthy.

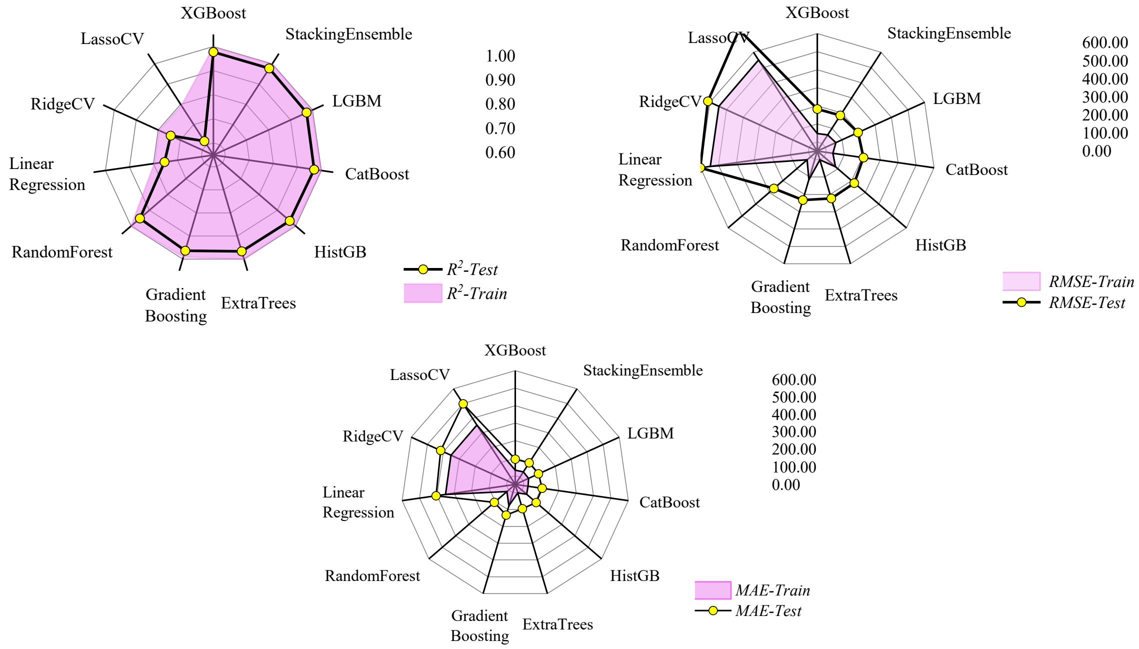


Figure 13. Comparision of models metric

The regression models were evaluated using R², RMSE, and MAE on both training and unseen test data. The advanced ensemble and boosting models, including StackingEnsemble, LGBM, CatBoost, HistGB, ExtraTrees, Gradient Boosting, RandomForest, and XGBoost, achieved substantially better performance than the linear models, with R² values close to 1.0 and lower prediction errors. The relatively small differences between training and test results indicate good generalization and limited overfitting. In contrast, Linear Regression, RidgeCV, and LassoCV showed considerably poorer performance, reflecting their limited ability to capture the nonlinear relationships in the dataset. Figure 13. SHAP analysis was used to interpret the contribution of the input features across different models. The results showed that feature effects varied considerably among the algorithms. XGBoost exhibited a wide distribution of SHAP values, with high feature values generally producing stronger positive effects, while RandomForest showed both positive and negative feature contributions. KNN displayed a narrower SHAP distribution, indicating lower sensitivity to changes in feature values. Overall, the SHAP analysis provided valuable insight into the differences in feature sensitivity and model behavior. Figure 14.

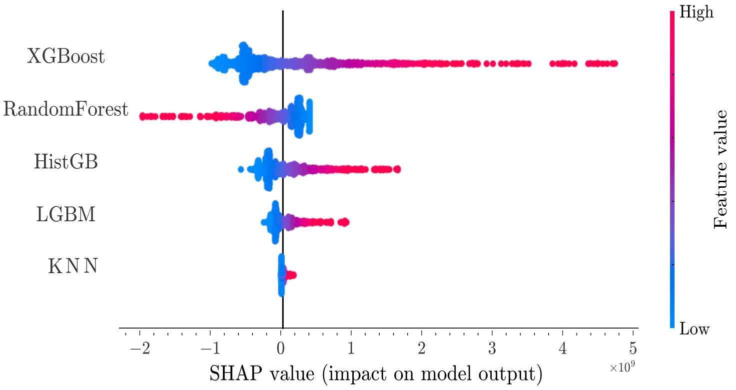


Figure 14. SHAP model importance plot.

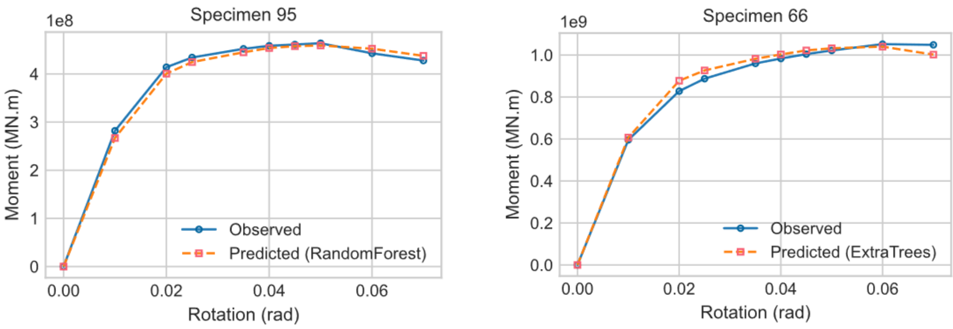


Figure 15. Predicted vs. true values for two different models.

As shown in the figure 14, stacking ensemble learning approach uses the best results of each trained algorithms to predict the final result. Figure 16 shows the prediction of the backbone curve for 10 unseen specimens selected from the test data set. As it can be seen there are good conformity with the test data.

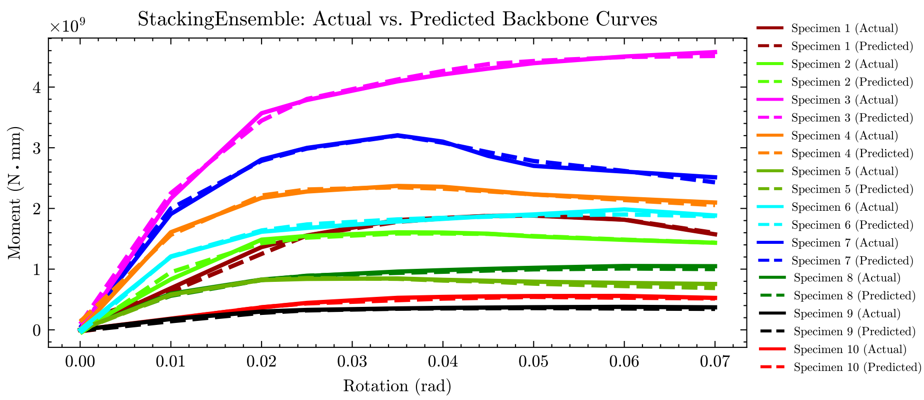


Figure 16. Predicted backbone curves for 10 unseen specimens using Stacking Ensemble method.

Also, figure 17 shows the scatter plot for the stacking ensemble method results.

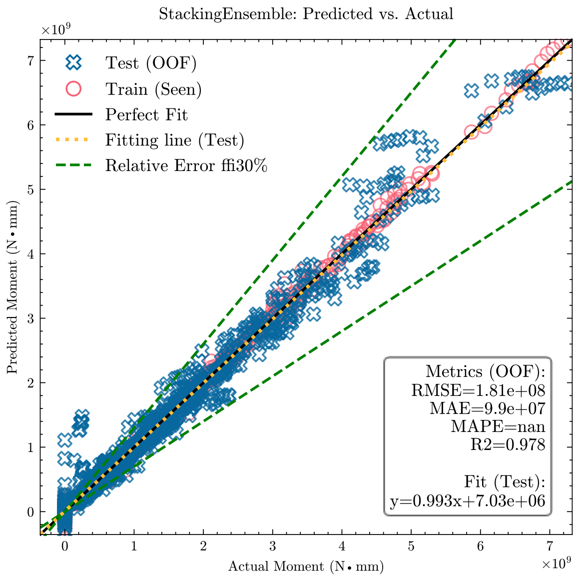


Figure 17. Predicted moment vs. actucal moment for stacking ensemble

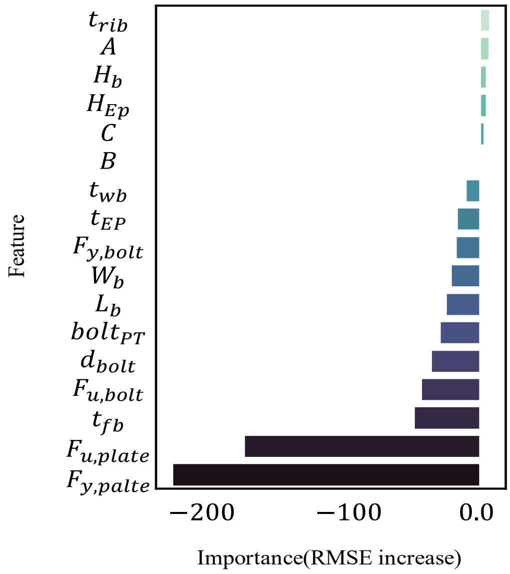


Figure 18. Feature importance metrics

5. Conclusion

An interpretable ensemble learning framework for predicting the nonlinear seismic performance of BEEP connections was successfully developed and validated in this study. The key conclusions are:

- The proposed ensemble learning framework, trained on a database collated from various experimental studies, accurately predicts the peak and ultimate moment capacities of BEEP connections.

- The framework's strong predictive output for unseen data validates its excellent generalization capabilities, providing a reliable and honest assessment of model performance.

- SHAP-based interpretability analysis confirms the physical consistency of the model's predictions, with feature importance (e.g., endplate thickness, bolt properties) aligning with established engineering mechanics.

- This framework provides a transparent, computationally efficient, and accurate tool to support performance-based seismic design, accelerate design optimization, and promote the adoption of trustworthy AI in structural engineering.

6. References

Adey, B. T., Grondin, G. Y., & Cheng, J. J. R. (1997). Extended end plate moment connections under cyclic loading (Structural Engineering Report No. 216). Department of Civil and Environmental Engineering, University of Alberta.

Adey, B. T., Grondin, G. Y., & Cheng, J. J. R. (2000). Cyclic loading of end plate moment connections. *Canadian Journal of Civil Engineering, 27*(4), 683–701.

Astaneh-Asl, A. (1995). *Seismic design of bolted steel moment-resisting frames* (Steel Tips). Structural Steel Education Council.

Agerskov, H. (1977). Analysis of bolted connections subject to prying. *Journal of the Structural Division, 103*(11), 2145–2163.

American Institute of Steel Construction. (2022). *Specification for structural steel buildings (ANSI/AISC 360-16)*. American Institute of Steel Construction.

European Committee for Standardization. (2005). *Eurocode 3: Design of steel structures—Part 1-8: Design of joints (EN 1993-1-8)*. European Committee for Standardization.

Sumner, E. A., & Murray, T. M. (2002). Behavior of extended end-plate moment connections subject to cyclic loading. *Journal of Structural Engineering, 128*(4), 501–508.

Sumner, E. A., III. (2003). *Unified design of extended end-plate moment connections subject to cyclic loading* [Doctoral dissertation, Virginia Polytechnic Institute and State University].

Guo, B., Gu, Q., & Liu, F. (2006). Experimental behavior of stiffened and unstiffened end-plate connections under cyclic loading. *Journal of Structural Engineering, 132*(9), 1352–1357.

Ghobarah, A., Korol, R. M., & Osman, A. (1992). Cyclic behavior of extended end-plate joints. *Journal of Structural Engineering, 118*(5), 1333–1353.

Shi, G., Shi, Y., Wang, Y., & Bijlaard, F. S. K. (2010). Monotonic loading tests on semi-rigid end-plate connections with welded I-shaped columns and beams. *Advances in Structural Engineering, 13*(2), 215–229.

American Society of Civil Engineers. (2017). *Seismic evaluation and retrofit of existing buildings*. American Society of Civil Engineers.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research, 12*, 2825–2830.

Mohammadnia, M., & Moradi, S. (2022). Artificial neural network–based predictive tool for modeling of self-centering endplate connections with SMA bolts. *Journal of Structural Engineering, 148*(12), 04022198.

Horton, T. A., Hajirasouliha, I., Davison, B., & Ozdemir, Z. (2021). Accurate prediction of cyclic hysteresis behaviour of RBS connections using deep learning neural networks. *Engineering Structures, 247*, 113156.

Shah, S. N. R., Sulong, N. H. R., & El-Shafie, A. (2018). New approach for developing soft computational prediction models for moment and rotation of boltless steel connections. *Thin-Walled Structures, 133*, 206–215.

Wang, K., et al. (2025). Experimental and numerical research on hysteretic behavior of bolted end-plate connections. *Structures*, 108397.

Stavroulakis, G. E., Avdelas, A. V., Abdalla, K. M., & Panagiotopoulos, P. D. (1997). A neural network approach to the modelling, calculation and identification of semi-rigid connections in steel structures. *Journal of Constructional Steel Research, 44*(1–2), 91–105.