Proceedings of the 10th World Congress on Civil, Structural, and Environmental Engineering (CSEE 2025)
Barcelona, Spain - April, 2025
Paper No. ICGRE 158
DOI: 10.11159/icgre25.158

Machine Learning Prediction of Headed Stud Shear Resistance in Profiled Corrugated Sheets for Steel-Concrete Composite Slabs

Mohamed Elnakeb[1], Karim Abdelhady[1], May Haggag[1], Ibrahim Abotaleb[1]
[1]The American University in Cairo
New Cairo, Cairo, Egypt
mohamedelnakeb@aucegypt.edu; karimabdelhay@aucedgypt.edu; mayhaggag@aucegypt.edu;
ibrahimsalah@aucegypt.edu

Abstract - Steel-concrete composite structures depend on headed shear studs for effective load transfer and composite action. Current design codes, such as EN 1994-1-1 and AISC 360-16, provide empirical formulas to estimate stud shear resistance. However, these codes often fail to account for failure mechanisms in modern profiled steel sheeting, resulting in unreliable predictions. This study evaluates the limitations of both codes using 611 push-out tests, revealing insufficient safety factors—1.09 for EN provisions and 0.83 for AISC, both below the target of 1.25. To address these shortcomings, a machine learning-based approach was developed. Key preprocessing steps included outlier removal, feature scaling, and selection of critical features using XGBoost. Four models—XGBoost, LightGBM, Random Forest, and Decision Tree—were evaluated, with Random Forest achieving superior performance ($R^2 = 0.9149$, RMSE = 6.87, APE = 9.38%), outperforming traditional codes. A hybrid approach was devised by incorporating a safety factor of 1.25 into machine learning predictions. Adjusted predictions closely aligned with experimental results, yielding an average ratio of 1.24 and a robust R^2 of 0.93. Standard deviation comparisons highlighted a reduction of over 73% compared to EN provisions and 59% relative to AISC, ensuring improved reliability. The proposed methodology bridges the gap between empirical limitations and real-world behavior, providing a precise, data-driven tool for shear resistance estimation. By integrating machine learning, this approach enhances safety, precision, and applicability in structural design, addressing critical challenges in modern composite construction.

***Keywords*:** Machine Learning, Steel-Concrete Composite Structures, Headed Shear Studs, Structural Design Reliability

1. Introduction

Steel structures are essential in modern construction due to their strength, versatility, and adaptability. Combining steel and concrete in composite systems enhances performance by utilizing the strengths of both materials. Headed shear studs, welded to steel beams and embedded in concrete, transfer shear forces at the steel-concrete interface, enabling effective composite action. This improves stiffness, strength, and ductility while reducing deflections and material costs.

In profiled steel sheeting, studs are typically placed through the troughs, penetrating the concrete slab above. The sheeting provides additional shear resistance and supports wet concrete during casting. Factors influencing stud effectiveness include stud height, concrete and steel strengths, and sheeting geometry.

Current design codes, such as EN 1994-1-1 [1] and AISC 360-16 [2], face challenges in predicting shear resistance for headed studs in modern profiled steel sheeting. EN 1994-1-1, based on only 57 push-out tests, often overestimates resistance, especially for open-trough profiles. AISC 360-16 inadequately addresses concrete failure mechanisms and newer rib geometries. Both codes struggle with slender rib profiles, leading to inconsistent resistance predictions and failing to meet mandated safety factors.

To address these challenges, advancements such as the Luxembourg and Stuttgart models and machine learning (ML) approaches have emerged [3]. The Luxembourg model refines EN 1994-1-1 equations by incorporating additional parameters like rib width, embedment depth, and concrete tensile strength, achieving safety factors between 1.21 and 1.28. The Stuttgart model introduces entirely new equations that account for weld collar dimensions and decking geometry, delivering safety factors around 1.47 with improved accuracy for complex geometries.

Machine learning models, such as the one by Degtyarev and Hicks [4], enhance predictive accuracy using large datasets and advanced algorithms. Trained on 452 push-out tests with 24 variables, including stud diameter, concrete strength, and rib dimensions, their model achieved an R^2 of 0.95 and RMSE of 12.20 kN. However, challenges like overfitting, variable redundancy, and underrepresented configurations limit generalizability.

This study addresses the gap by applying machine learning to a larger dataset with fewer, carefully selected variables. By reducing complexity and focusing on critical predictors, it aims to develop a more practical and accurate tool for shear resistance estimation in steel-concrete composite structures.

2. Methodology

2.1. Dataset Description

The dataset analyzed in this study includes 611 push-out test results on headed stud shear connectors in steel-concrete composite systems, compiled by Vigneri et al [3]. from various experimental investigations. It captures key geometric and material parameters such as stud diameter (19–22 mm), welded stud length (70–200 mm), sheeting thickness (0.6–1.2 mm), and rib depth (40–136 mm). Material properties include concrete compressive strength (24–58.1 MPa) and stud tensile strength (417–570 MPa). The dataset also considers the number of studs per rib (up to two) and their positions within the ribs (centered, staggered, or offset), which influence load distribution and resistance. Lightweight concrete was excluded, and a minimum concrete strength threshold ensured comparability across tests. Table 1 summarizes the variables, their units, and ranges.

Table 1: Dataset Description

Symbol	Name of Variable	Unit	Min Value	Max Value
nr	Number of studs per rib	count	1	3
Pos.	Position of the stud in the trough	N/A	N/A	N/A
Welding	Type of welding procedure	N/A	N/A	N/A
et	Transversal spacing of the studs in the trough	mm	0	191
eL	Longitudinal eccentricity of the stud in the trough	mm	-52.75	52.75
dnom	Nominal diameter of the shank of the stud	mm	10	22
hscm	Mean (measured) height of the stud after welding	mm	70	200
t	Nominal thickness of the sheeting	mm	0.6	1.29
Sheeting	Steel deck product name	N/A	N/A	N/A
hp	Nominal (net) height of the trough	mm	38.1	152.4
btop	Nominal top width of the trough	mm	63.5	240
bbot	Nominal bottom width of the trough	mm	40	200
fcm	Mean value of concrete cylinder compressive strength	MPa	17.82	52
fum	Mean ultimate tensile strength of the stud material	MPa	417	570
TL	Transversal load	N	N/A	N/A
Pe	Experimental resistance per connector	kN	18.02	144.2

2.2. Current Code Provisions

2.2.1. EN 1994-1-1

The EN 1994-1-1 code determines the shear resistance of headed stud connectors in steel-concrete composite beams by by evaluating two failure modes—steel failure and concrete-related failure—and using the lesser value for design (CEN, 2004). Steel failure resistance depends on the stud material's ultimate tensile strength (fu), stud shank diameter (d), and a a reduction factor (kt), which accounts for the geometry of profiled sheeting and stud arrangement. The reduction factor kt considers parameters like the number of studs per rib (nr), rib base width (b0), sheeting rib height (hp), and welded stud height (hsc). Concrete-related failure resistance depends on the concrete's compressive strength (fc), modulus of elasticity (Ec), stud diameter (d), and a factor (α) reflecting the embedment depth-to-diameter ratio (hsc/d). This factor adjusts resistance based on embedment depth, as greater depth improves concrete shear resistance. The code sets upper limits for kt to prevent overestimations [1]. The equations below outline these methodologies.

$$r_t = \min\{r_{t,s}\,, r_{t,c}\} \tag{1}$$

$$r_{t,s} = 0.80 \cdot k_t \cdot f_u \pi \frac{d^2}{4}\,,\ r_{t,c} = 0.29 \cdot k_t \cdot d^2 \alpha \sqrt{f_c E_C} \tag{2}$$

$$k_t = \min\{\frac{0.7}{\sqrt{n_r}} \frac{b_0}{h_p}\left(\frac{h_{sc}}{h_p} - 1\right), k_{t,max}\}\,,\ \alpha \begin{cases} 0.2\left(\frac{h_{sc}}{d} + 1\right) & for\ 3 \leq \frac{h_{sc}}{d} \leq 4 \\ 1 & for \frac{h_{sc}}{d} > 4 \end{cases} \tag{3}$$

Table 2: Cases of EN 1994-1-1 Code

$k_{t,max}$	Thickness of the sheeting t	Studs not exceeding 20 mm in diameter and welded through profiled steel sheeting	Case	Profiled sheeting with holes and studs 19mm or 22mm in diameter	Case
$n_r = 1$	t <= 1mm	0.85	A	0.75	C
	t > 1mm	1.0	B	0.75	D
$n_r = 2$	t <= 1mm	0.70	E	0.60	G
	t > 1mm	0.80	F	0.60	H

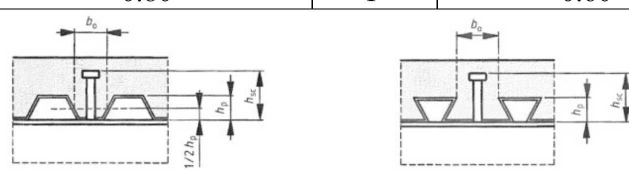

Figure 1. Cross Section of a Headed Stud Connection in Steel-Concrete Composite in EN 1994-1-1

2.2.1. AISC 360-16

The AISC 360-16 provisions for headed stud connectors in steel-concrete composite structures determine shear resistance as the lesser capacity between two failure modes: steel failure and concrete-related failure (AISC, 2016). Steel failure resistance is based on the stud's ultimate tensile strength (fu), diameter (d), and reduction factors Rg and Rp. Rg accounts for the number of studs per rib (1.0 for one stud, 0.85 for two), while Rp adjusts for stud position within the rib (0.75 for embedment ≥50 mm, 0.6 for less favourable positions). Concrete-related failure resistance depends on the concrete compressive strength (fc), elastic modulus (Ec), and stud diameter. Unlike EN 1994-1-1, AISC 360-16 uses a simplified approach without additional embedment geometry factors [2]. The equations below outline the methodology.

$$r_t = \min\{r_{t,s}\ , r_{t,c}\} \tag{4}$$

$$r_{t,s} = 1.00 \cdot Rg \cdot\ Rp \cdot\ f_u \pi \frac{d^2}{4}\ , r_{t,c} = 0.5 \cdot \pi \frac{d^2}{4} \sqrt{f_c E_C} \tag{5}$$

$$Rg = \begin{cases} 1.0 & for\ n_r = 1 \\ 0.85 & for\ n_r = 2 \end{cases},\ Rp = \begin{cases} 0.75 & for\ e_{mid-ht} \geq 50mm \\ 0.6 & otherwise \end{cases} \tag{6}$$

Table 3: Cases of AISC 360-16 Code

	Rg = 1	**Rg = 0.85**
Rp = 0.75	W	Y
Rp = 0.6	X	Z

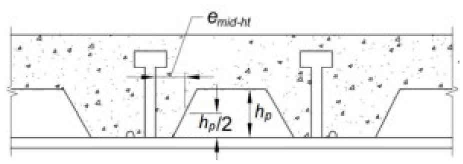

Figure 2. Cross Section of a Headed Stud Connection in Steel-Concrete Composite in AISC 360-16

2.3. Machine Learning
2.2.2. Machine Learning

The methodology begins with implementing basic machine learning models, including Decision Tree, Random Forest, and Linear Regression, in default settings to explore the dataset, identify patterns, and diagnose issues such as overfitting, underfitting, or feature significance. Insights from this step guide data refinement and model development.

Data preparation ensures dataset quality and reliability. Invalid or non-numeric values are removed, irrelevant features are discarded, and categorical variables are encoded using binary, ordinal, or one-hot methods. Numerical features are normalized with standard or Min-Max scaling to ensure consistency, while logarithmic and square root transformations address skewness and outliers. Statistical methods are applied to detect and remove outliers in the target variable [5]. Gradient boosting algorithms rank feature importance, identifying key predictors for advanced modeling.

The refined dataset is used to train and optimize advanced models, including XGBoost, LightGBM, and Random Forest, as well as simpler models like Decision Tree [6]. Hyperparameter tuning systematically enhances model performance. Effectiveness is assessed using Root Mean Square Error (RMSE), which quantifies prediction error (lower values indicate better performance), and R^2, which measures the proportion of variance explained by the model (values closer to 1 indicate a stronger fit). Together, these metrics provide a comprehensive evaluation of accuracy and robustness.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(yi - yactual[i])^2}{n}} \tag{7}$$

$$R2 = 1 - \frac{\sum_{i=1}^{n}(yi - yactual[i])^2}{\sum_{i=1}^{n}(yi - ymean)^2} \tag{8}$$

3. Results and Discussion

3.1. Applying Current Code Provisions

This study evaluates the accuracy and reliability of EN 1994-1-1 (Eurocode) and AISC 360-16 in predicting stud shear resistance by comparing theoretical predictions with experimental results from a dataset of 611 push-out tests. Both codes aim to achieve safety factors between 1.25 and 1.50 to balance structural safety with efficiency, accounting for uncertainties in material properties, construction tolerances, and load assumptions.

The Eurocode frequently overestimates shear resistance, especially for open-trough profiles, creating a risk of non-conservative designs where actual stud capacity is lower than predicted. A scatter plot of experimental versus theoretical values highlights significant discrepancies and variability in predictions. Specific cases, such as Case 0 and Cases C and D, exhibit substantial inaccuracies, as indicated in Figure 3. These deviations suggest limitations in the code's representation of critical parameters. The average safety factor for the Eurocode across all cases is 1.09, significantly below the target value of 1.25. Boxplot analysis Figure 3 further reveals high variability in predictions, particularly in configurations where the code fails to provide reliable resistance estimates.

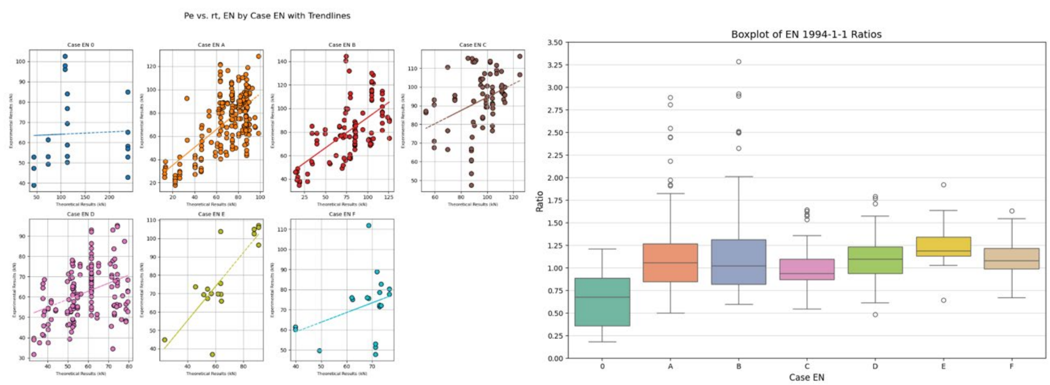

Figure 3. Scatter Plot of the Dataset Factor if Safety for both EN 1994-1-1 and AISC 360-16 Codes

Similar trends are observed with the AISC 360-16 code, which also overestimates stud resistance compared to experimental results. The scatter plot in Fig. 8 illustrates consistent overestimation, raising concerns about the reliability of its predictions under specific conditions. Analysis of all cases Figure 4 confirms significant inaccuracies, with none of the cases achieving the required safety factor. The average safety factor for the AISC code is 0.83, far below the target of 1.25, and inadequate even to match experimental results. Boxplot analysis Figure 4 highlights the high variability in predictions, emphasizing the code's inability to ensure consistent safety margins.

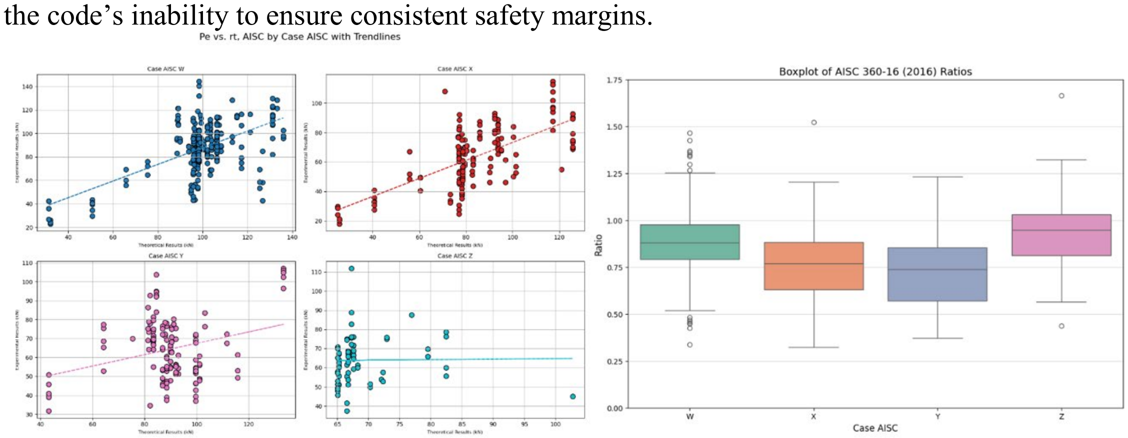

Figure 4. Scatter Plot of the Dataset Factor if Safety for both EN 1994-1-1 and AISC 360-16 Codes

A combined analysis of safety factors for both codes further underscores their inadequacy in reliably estimating stud shear resistance. The majority of data points fall short of the targeted safety factor of 1.25, as shown in Figure 5**,** this highlights the urgent need for substantial revisions to both codes to address these discrepancies and improve their accuracy and applicability in modern construction

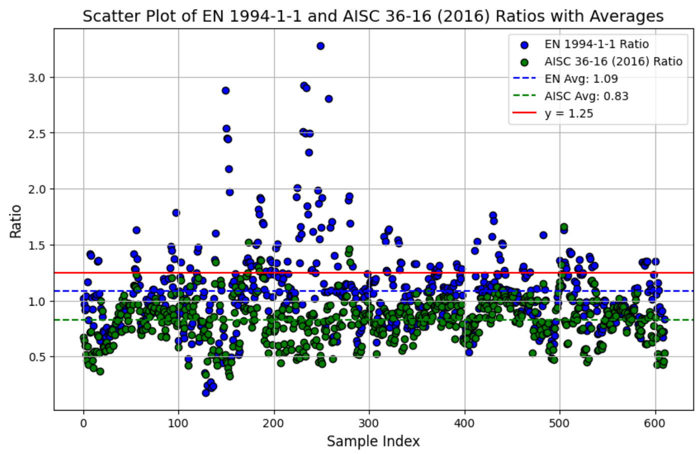

Figure 5. Scatter Plot of the Dataset Factor if Safety for both EN 1994-1-1 and AISC 360-16 Codes

3.2. Machine Learning Model

3.2.1 Unmodified Approach

The study evaluated preliminary machine learning models, including Linear Regression, Decision Tree, Random Forest, and XGBoost, using R^2 and RMSE. Random Forest outperformed the others, achieving an R^2 of 0.72, the lowest RMSE (11.21), indicating superior predictive accuracy. However, these preliminary models lacked sufficient reliability for structural applications, necessitating the development of more advanced techniques.

Final model performance was assessed with XGBoost, LightGBM, Random Forest, and Decision Tree. Random Forest achieved the highest R^2 (0.9149), adjusted R^2 (0.9063), and the lowest RMSE (6.87) and RMSE percentage (9.38%), making it the most accurate and reliable model. XGBoost and LightGBM also performed well, with R^2 values of 0.9116 and 0.9044, respectively, while Decision Tree showed lower stability in cross-validation (R^2 CV = 0.7083). Random Forest's minimal error and consistency confirmed it as the optimal model for further analysis.

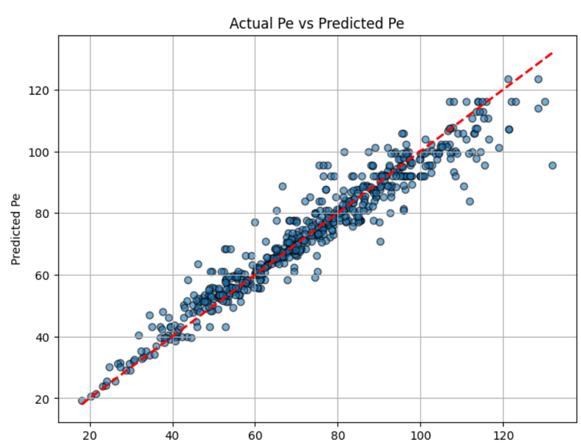

Figure 6. Scatter Plot of Actual V Predicted for the Training Set Using Random Forest

Applying Random Forest to the training set showed strong alignment between predicted and actual values, with points clustering closely around the y = x line, particularly in the mid-range. For EN cases, the model performed exceptionally for Cases EN 0, A, and E, achieving R^2 values of 0.98, 0.96, and 0.97, respectively, with standard deviations reduced to 0.07, 0.07, and 0.08. Cases EN B, C, and F showed moderate accuracy, with R^2 values of 0.86, 0.77, and 0.73 and standard deviations of 0.10, 0.09, and 0.10, respectively. Compared to the EN code provisions, which exhibited an average standard deviation of 0.317, the machine learning model reduced variability by over 73%, achieving an average standard deviation of 0.085. For AISC cases, Random Forest achieved high R^2 values for Cases Y (0.94), X (0.92), and W (0.89), with slightly lower accuracy for Case Z (0.84). Standard deviations were significantly reduced to 0.07 for Cases Y and X, 0.08 for Case W, and 0.08 for Case Z. In contrast, the AISC provisions had standard deviations of 0.196, 0.177, 0.185, and 0.180 for these cases, with an average of 0.185. The machine learning model reduced variability by over 59%, highlighting its superior accuracy and consistency compared to traditional code provisions.

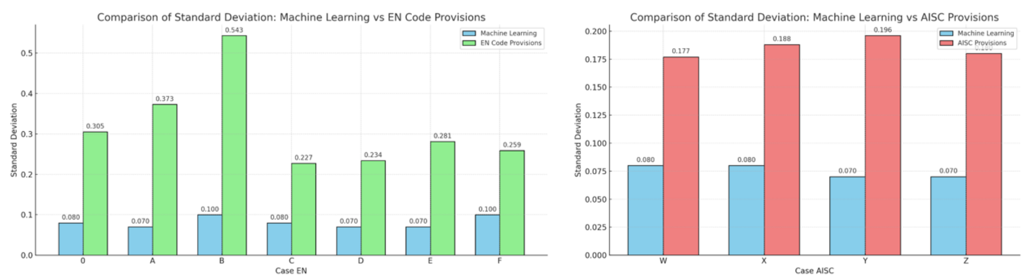

Figure 7. Code Provisions Standard Deviations V Machine Learning Prediction Standard Deviations

3.2.2 Modified Approach

A hybrid approach was developed by dividing machine learning-predicted Pe values by a safety factor of 1.25 to ensure structural reliability. The adjusted predictions were compared to experimental results using a scatter plot, with the trend line showing a strong linear correlation ($R^2 = 0.93$).

The adjusted predictions closely aligned with experimental values, with an average ratio of 1.24, near the target safety factor of 1.25. The low spread around the average ratio highlights the hybrid approach's reliability and precision, ensuring both accuracy and safety in shear resistance predictions.

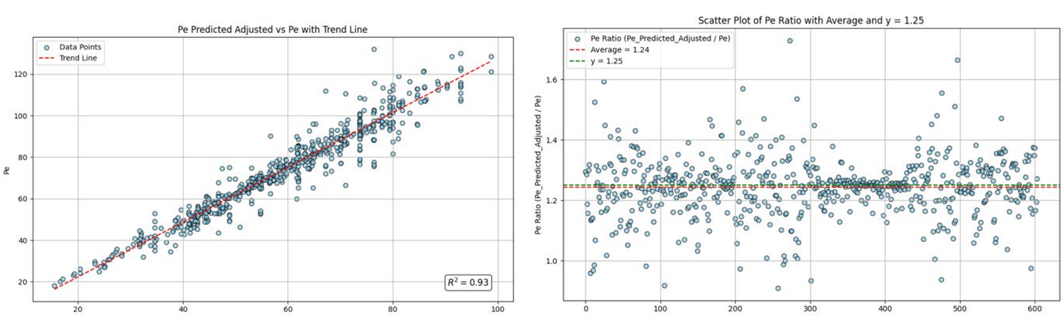

Figure 8. Modified Machine Learning Model V Actual Experimental Results

3.2.3 Comparison with Current Code Provision Results

The scatter plot shown in Figure 9 compares the ratios of machine learning predictions, EN 1994-1-1, and AISC 360-16 provisions against the safety factor of 1.25. Machine learning predictions, with an average ratio of 1.24, closely align with the target, demonstrating superior accuracy and consistency. In contrast, EN 1994-1-1 averages 1.09, falling short, and AISC 360-16 underpredicts with 0.83. This underscores the machine learning model's reliability and its outperformance of both codes.

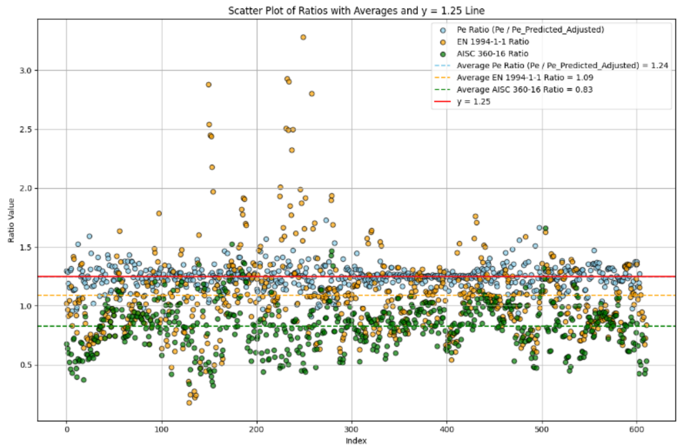

Figure 9. Comparison of the Modified Machine Learning Model with the Other Code Provisions

3. Conclusions

- A machine learning-based approach was developed to improve shear resistance predictions for headed studs, addressing limitations in EN 1994-1-1 (average ratio 1.09, below the target safety factor of 1.25) and AISC 360-16 (average ratio 0.83, significantly underpredicted).
- The Random Forest model achieved the best results with $R^2 = 0.9149$ and RMSE = 6.87. Adjusted predictions, divided by 1.25 for safety, achieved an average ratio of 1.24 and $R^2 = 0.93$, aligning closely with the target safety factor while maintaining a strong linear correlation with experimental results.
- Standard deviations were significantly reduced with the machine learning model (0.07 to 0.10), compared to EN provisions (0.227 to 0.543) and AISC provisions (0.177 to 0.196), achieving a 73% reduction in variability compared to EN and 59% compared to AISC.
- Machine learning predictions consistently reduced variability and maintained accuracy across all cases, outperforming traditional design provisions while ensuring structural safety margins.
- This study demonstrates the potential of machine learning to enhance structural design practices, offering greater precision and reliability when combined with safety adjustments to meet accuracy and safety standards.

References

[1] European Committee for Standardization (CEN), “Eurocode 4 (EN 1994-1-1): Design of composite steel and concrete structures – Part 1-1: General rules and rules for buildings,” Brussels, Belgium, 1994.
[2] “Specification for structural steel buildings.,” Chicago, IL, USA, 2016.
[3] V. Vingneri, S. J. Hicks, A. Taras, and C. Odenbreit, “Design models for predicting the resistance of headed studs in profiled sheeting,” *Steel and Composite Structures*, vol. 42, no. 5, pp. 633–647, 2022.
[4] M. Razavi Setvati and S. J. Hicks, “Machine learning models for predicting resistance of headed studs embedded in concrete,” *Eng Struct*, vol. 254, p. 113803, Mar. 2022, doi: 10.1016/j.engstruct.2021.113803.
[5] J. Freidman, T. Hastie, and R. Tibshirani, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. New York, NY, USA: Springer, 2001.
[6] L. Breiman, “Machine Learning,” *Mach Learn*, vol. 45, no. 1, pp. 5–32, 2001, doi: 10.1023/A:1010933404324.