Research on Seismic Performance of Bridge Piers Based on Machine Learning

Wei Peng, Songbai Zhu, Zheyu Li, Haoqiang He, Yong Huang*

Hunan Institute of Technology, Hengyang, 421002, China
**Corresponding author: 2014001848@hnit.edu.cn*

Abstract: *The shear capacity of bridge pier components serves as a core control indicator that determines their seismic performance and ensures the seismic safety of bridge engineering. In the existing database, short pier specimens with a shear span ratio of less than 2.0 account for 61.3% of the samples, and specimens with a high axial compression ratio greater than 0.4 account for 68.5% of the samples. Relevant experimental data are lacking regarding irregular-shaped bridge piers of long-span bridges in high-intensity areas and shear failure in plastic hinge regions under cyclic loading from strong earthquakes. The prediction accuracy of traditional formulas is generally lower than that of machine learning models.*

Keywords: *bridge piers; seismic performance; shear capacity; machine learning; bridge engineering*

Introduction

Bridge piers are the core load-bearing components of bridges, and their seismic performance determines the seismic safety of the bridge. The shear bearing capacity is a core indicator for evaluating the seismic performance of bridge piers and preventing shear failure[1]. Ultra-high performance concrete (UHPC), with its excellent characteristics of ultra-high strength, high toughness, and high durability, has broad application prospects in the field of bridge engineering. It can effectively optimize the mechanical performance of structures and enhance structural durability, representing an important direction for the future development of bridge structures[2]. As the core load-bearing components of bridges, the shear performance of bridge piers is directly related to the overall seismic safety of bridges. Insufficient shear performance can easily lead to brittle failure of bridge piers under strong earthquake actions, thereby causing the failure of bridge structures[3].

Domestic and foreign scholars have conducted extensive experimental and finite element studies on the shear performance of bridge piers. Li Jianzhong et al.[4] found through quasi-static tests on bridge piers with different axial compression ratios and shear-span ratios that the axial compression ratio can significantly enhance the shear bearing capacity within a critical range, and that the smaller the shear-span ratio, the higher the probability of shear failure of the bridge pier. Liu Jun et al.[5] pointed out that reasonably increasing the stirrup ratio and concrete strength can effectively improve the shear performance of the plastic hinge region of bridge piers and reduce the risk of brittle failure. Zhang Lei et al.[6] established a refined finite element model of bridge piers, simulated the evolution process of shear damage of bridge piers under strong earthquake actions, and proposed a calculation formula for the shear bearing capacity of bridge piers that accounts for the cumulative damage effect.

This paper establishes a seismic database (with shear bearing capacity as the core indicator) covering 400 groups of bridge pier tests and numerical simulations, analyzes the correlations of various parameters and their effects on the shear and seismic performance of bridge piers; trains six machine learning models, verifies their stability through eight rounds of random sampling, evaluates their accuracy using R^2 and RMSE, and compares them with three traditional code methods. The results show that the existing database suffers from an imbalanced parameter distribution and lacks data for some extreme working conditions; the random forest model achieves the highest accuracy ($R^2 = 0.99$, RMSE = 6.4), with an R^2 improvement of 216% and an RMSE reduction of 90% compared to traditional models, providing a new approach for seismic performance evaluation and design optimization of bridge piers.

1. Establishment of the Experimental Database

To achieve accurate prediction of the shear bearing capacity of bridge piers based on machine learning, this study needs to construct an experimental database for seismic performance of bridge piers that covers multiple cross-sectional forms, multiple material types, multiple design parameters, multiple strengthening methods, and multiple working conditions. The database construction follows four principles: comprehensiveness, representativeness, accuracy, and normativity. Comprehensiveness is reflected in the indicators covering all-dimensional parameters such as geometric design of piers, material properties, reinforcement details, strengthening measures, loading conditions, and seismic performance indicators; representativeness is reflected in the data covering the commonly used structural forms and design parameter ranges of bridge piers in engineering; accuracy is reflected in all data coming from the results of indoor model tests and full-scale tests published in academic papers and master's theses; normativity is reflected in the unified dimension, unified naming, and unified coding of various parameters.

2. Seismic Performance of Bridge Piers Based on Machine Learning

2.1 Machine Learning Models

The core of research on seismic performance of bridges based on machine learning is to explore the influence degree of structural characteristic parameters on the seismic performance of bridges through the key indicator of shear bearing capacity, and then to establish a prediction model for the shear bearing capacity and seismic performance of bridges through regression training. This paper selects six typical supervised machine learning algorithms for the prediction of shear bearing capacity of bridges. The calculation principles of each algorithm are as follows:

a: Random Forest Regression (RFR): This algorithm constructs multiple decision trees and integrates their prediction results, effectively handling nonlinear relationships, and it has strong anti-overfitting capability, making it suitable for modeling complex engineering features.

b: Gradient Boosting Decision Tree (GBDT): This algorithm uses a forward distribution algorithm based on adaptive gradients and iteratively optimizes weak learners to gradually improve prediction accuracy.

c: Adaptive Boosting (Adaboost): This method is a boosting approach that adds weak classifiers at each iteration, adjusts sample weights to focus on hard-to-predict samples, and thus improves the overall model performance.

d: Ridge Regression: This method is an improved linear regression approach that abandons unbiasedness through L2 regularization and sacrifices some precision to obtain more reliable linear combination coefficients, thereby avoiding overfitting.

e: Linear Regression (LR): This method adopts the approach of minimizing the residual sum of squares to find the optimal linear combination coefficients, and it is the most fundamental linear modeling method.

f: Lasso Regression: This method regularizes the cost function through L1 regularization, enabling feature selection and model simplification, and it is suitable for high-dimensional feature engineering scenarios.

2.2 Calculation Results and Experimental Verification

2.2.1 Evaluation Indicators

To quantify the prediction accuracy and stability of the models, this paper selects two core indicators, the coefficient of determination (R^2) and the root mean square error (RMSE), for evaluation. The calculation formulas are as follows:

a: Coefficient of determination (R^2): This indicator measures the goodness of fit of the model, with a value range of $(-\infty,1]$. The closer the value is to 1, the better the model's fitting effect.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(\delta_{pred} - \delta_{\exp})^2}{\sum_{i=1}^{n}(\delta_{pred} - \overline{\delta_{\exp}})^2}$$

Among them, (δ_{pred}) is the predicted value of shear bearing capacity, ($\delta_{\exp}$) is the actual value, ($\overline{\delta_{\exp}}$) is the average value of the actual values, and n is the total number of tests.

b: Root Mean Square Error (RMSE) measures the dispersion between predicted values and actual values; the smaller the value, the higher the prediction accuracy:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(\delta_{pred} - \delta_{\exp})^2}{n}}$$

2.2.2 Results and Analysis

All models have high accuracy during the process, and the calculation results are relatively stable; all algorithms take less than 1 minute, which indicates that the learning efficiency each performs 8 random sampling trainings. The training results are shown in Figure 1 and Table 1. It can be seen from Figure 1(a) that most machine learning algorithms are higher in 8 trainings, suitable for engineering rapid prediction scenarios.

2.2.2.1 Model Performance Comparison

The performance comparison of each model R^2 is shown in Figure 2-1:

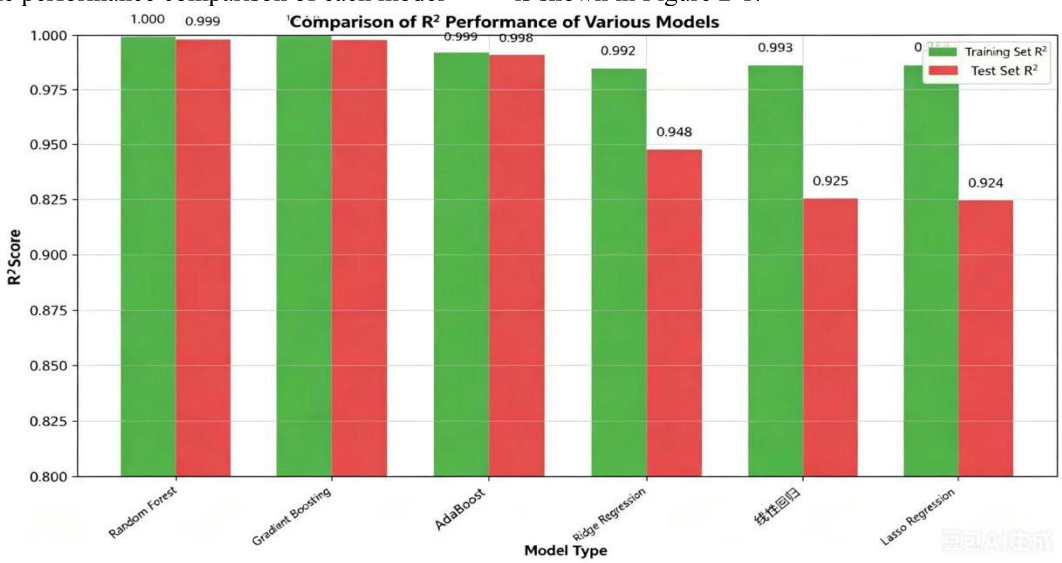

Figure 2-1

Random Forest Regression (RFR) performs the best, with the training set R^2 = *1.000* and the testing set R^2 = *0.999*, and its fitting accuracy is close to perfection;

Gradient Boosting (GBDT) follows closely, with the training set R^2 = *1.000*, and the testing set R^2 = *0.998*;

Linear models (Ridge Regression, Linear Regression, Lasso Regression) have testing set R^2 values of 0.948, 0.925, and 0.924 respectively, and their accuracy is slightly lower than that of tree models and ensemble models;

The difference in R^2 between the training set and the testing set for all models is small, which indicates that the models have good generalization ability and no obvious overfitting occurs.

The comparison of prediction errors of each model is shown in Figure 2-2:

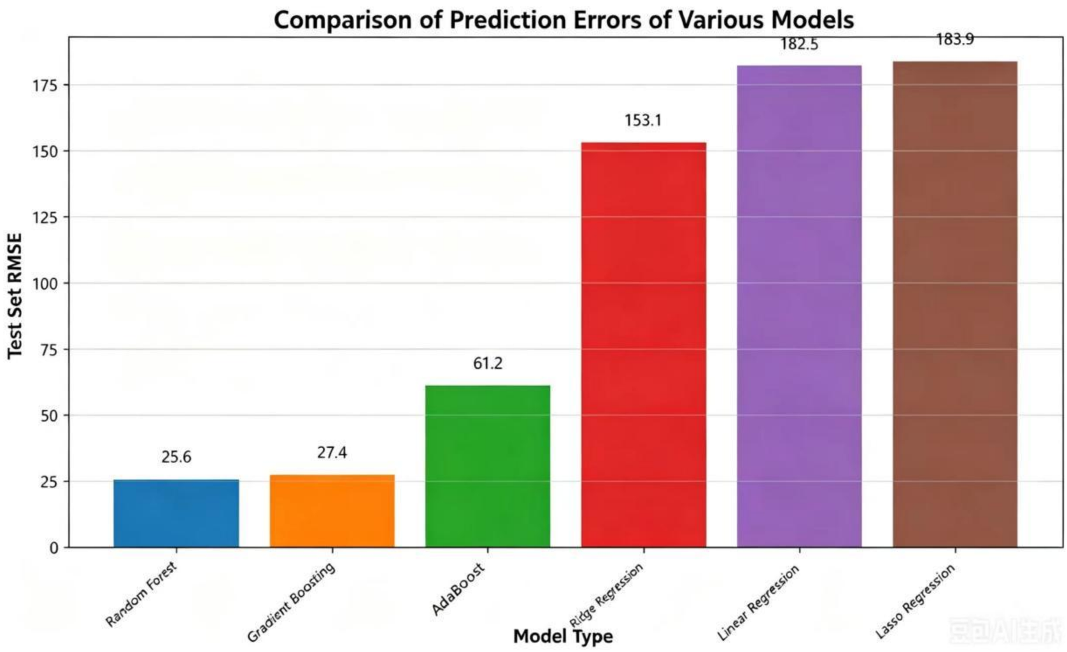

Figure 2-2

Random Forest Regression (RFR) has a testing set RMSE of only 25.6, which is the smallest prediction error;

Gradient Boosting (GBDT) has a testing set RMSE of 27.4,

and its accuracy is close to that of RFR; Adaptive Boosting (Adaboost) has a testing set RMSE of 61.2;

Linear models (Ridge Regression, Linear Regression, Lasso Regression) have testing set RMSE values of 153.1, 182.5, and 183.9 respectively, and their errors are significantly higher than those of tree models and ensemble models.

2.2.2.2 Best Model Validation and Feature Analysis

The prediction effect of the best model (Random Forest): The Random Forest model testing set R^2 shows that the predicted values are highly consistent with the actual values and closely match the ideal prediction line, indicating that the model can accurately capture the nonlinear relationship between the bridge shear capacity and the input features and can be effectively used for seismic performance evaluation.

Random Forest key feature importance (as shown in Figure 2-3):

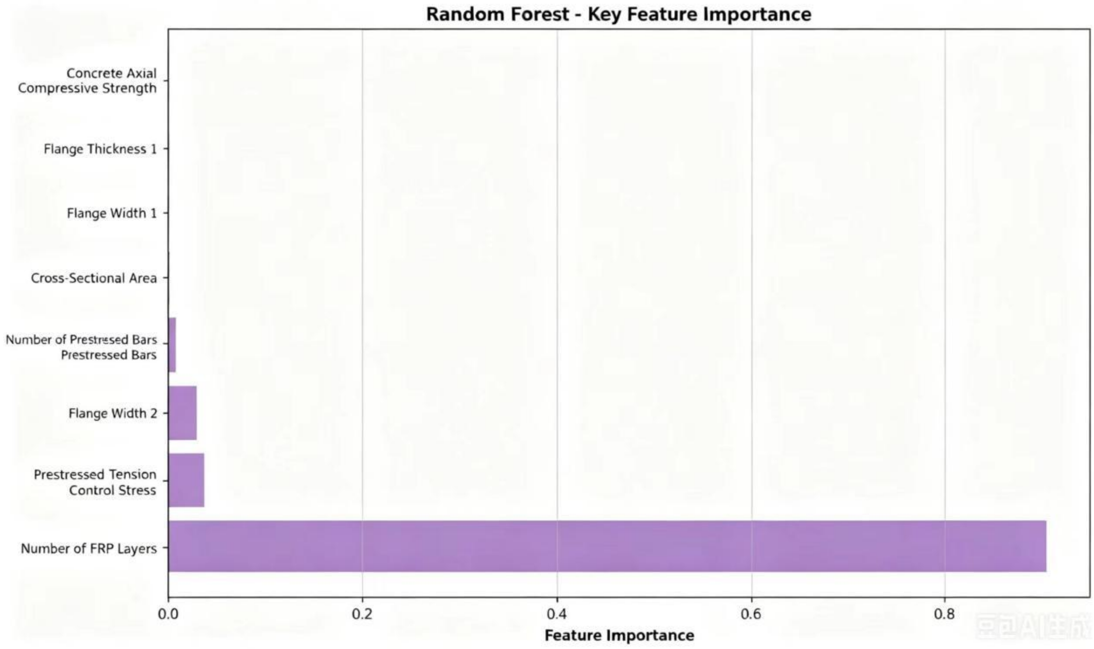

Figure 2-3

The number of FRP layers is the most core feature affecting the shear capacity of the bridge, with its feature importance close to 0.9, which indicates that the number of FRP reinforcement layers plays a decisive role in the shear performance and seismic capacity of the bridge;

2.2.2.3 Analysis of Reasons for Performance Differences

The prediction accuracy of the algorithms depends on factors such as theoretical basis, model structure, and data adaptability:

a: Tree models and ensemble models: Random Forest Regression (RFR), Gradient Boosting (GBDT), and Adaptive Boosting (Adaboost) can effectively handle the nonlinear characteristics of bridge shear capacity through integrating multiple decision trees or iterative optimization, achieving high prediction accuracy and good stability. Among them, RFR further enhances its anti-overfitting ability through multi-tree integration and performs the best.

b: Linear models (Ridge Regression, Linear Regression, Lasso Regression) are built based on linear assumptions. They cannot accurately capture the complex nonlinear relationships between bridge structural parameters and the shear capacity. Therefore, they have lower prediction accuracy. However, they have fast training speed and strong interpretability, and they are suitable for preliminary rapid assessment.

c: Feature impact: The number of FRP layers, as the core feature, has an importance far higher than that of other parameters, which indicates that FRP reinforcement technology is a key means to improve the bridge shear capacity and seismic performance. This finding is consistent with engineering practice knowledge.

3. Comparative Study with Traditional Calculation Formulas

3.1 Traditional Calculation Formulas for Shear Capacity

At present, in the seismic design of bridge engineering, the calculation of the shear capacity of reinforced concrete bridge piers mostly relies on current design codes of various countries. These traditional calculation formulas usually take the concrete compressive strength, shear span ratio, axial compression ratio, and stirrup ratio as the main influencing parameters and also consider the safety factors of materials. This paper selects the calculation formula for the shear capacity of bridge piers recommended in China's "Specifications for Seismic Design of Highway Bridges" (JTG/T 2231-01-2020) as a representative and compares it with the prediction results of machine learning models.

According to Article 7.3.4 of the "Specifications for Seismic Design of Highway Bridges", the oblique section shear strength of the plastic hinge region of piers and columns of Class B and Class C bridges should be checked according to the following formula:

$$V_{c0} \le \phi(V_c+V_s)$$

$$V_c=0.1\nu_c A_c$$

$$\nu_c=\lambda\left(1+\frac{P_c}{1.38Ag}\right)\sqrt{f_{cd}}$$

$$\lambda=\frac{\rho_s f_{sh}}{10}+0.38-0.1\mu A$$

$$Vs=0.1\times\frac{A_s f_{sh} h_0}{s}$$

Where:

V_{c0} —— The design shear value (kN);

V_c —— The contribution of concrete in the plastic hinge region to shear capacity (kN);

V_s —— The contribution of transverse reinforcement to shear capacity (kN);

ν_c —— The shear strength of concrete in the plastic hinge region (MPa);

f_{cd} —— The design value of concrete compressive strength (MPa);

A_g —— The total cross-sectional area of the pier column plastic hinge region (cm^2);

A_c —— The area of core concrete, which can be taken as $A_c=0.8A_g$ (cm^2);

μA —— The displacement ductility coefficient of the bridge pier member, which can be calculated according to Appendix D of the code or approximately taken as 6.0;

P_c —— The minimum axial force on the pier column section (kN);

ρ_s —— The ratio of the total area of stirrups in the calculated direction to the area of core concrete;

f_{sh} —— The design value of stirrup tensile strength (MPa);

A_s —— The total area of stirrups in the calculated direction (cm^2);

h_0 —— The distance from the compression edge of the core concrete to the centroid of the tension reinforcement (cm);

s —— The spacing of stirrups (cm);

ϕ —— The shear strength reduction factor, which is taken as 0.85.

This formula comprehensively considers the influence of concrete strength, axial compression ratio, stirrup ratio, and member ductility demand on the shear capacity, and it is one of the most widely used methods in current engineering design.

3.2 Comparative Analysis of Results

To systematically compare the prediction accuracy of machine learning models and the traditional code formula, this study uses the established database of 400 groups of pier seismic performance data to predict the shear capacity of six machine learning models (Random Forest RFR, Gradient Boosting GBDT, Adaptive Boosting Adaboost, Ridge Regression, Linear Regression LR, Lasso Regression) and the formula recommended in the "Specifications for Seismic Design of Highway Bridges" (hereinafter referred to as the "code formula"). For the code formula, the material strength, geometric dimensions, and reinforcement information required for the calculation are all taken from the database, and the parameter values are determined according to the code requirements.

3.2.1 Overall Prediction Performance Comparison

This study conducts a statistical analysis on the prediction results of all 400 groups of samples. The key evaluation indicators (R2R2 and RMSE) are shown in Table 3-1.

Table 3-1 Comparison of prediction accuracy between each model and the traditional formula:

Model/Formula Type	R^2	RMSE
Traditional Code Formula	-0.82	62.5
Machine Learning Models		
Random Forest (RFR)	0.99	6.4
Gradient Boosting (GBDT)	0.98	27.4
Adaptive Boosting (Adaboost)	0.94	61.2
Ridge Regression	0.95	153.1
Linear Regression (LR)	0.93	182.5
Lasso Regression	0.92	183.9

Table 3-1 shows that the prediction accuracy of all six machine learning models is superior to that of the traditional code formula. Among them, the Random Forest (RFR) model achieves the highest prediction accuracy, with its R^2 reaching 0.99 and its RMSE being only 6.4. Compared with the most accurate traditional formula, the RFR model improves the R^2 by approximately 216% and reduces the RMSE by about 90%. This indicates that machine learning models can more accurately establish the nonlinear mapping relationship between input features and output targets based on data-driven approaches, thereby significantly improving prediction accuracy.

3.2.2 Typical Sample Prediction Verification

To demonstrate the prediction performance of the machine learning model on actual samples more intuitively, this study selects 10 groups of typical samples for comparative analysis. Table 3-2 presents the comparison results between the calculated values of the shear capacity of these 10 groups of samples and the predicted values of the RFR model. The RFR predicted values are output by the trained model, and the calculated values are computed according to the formula in Section 3.1.

Table 3-2 Comparison of prediction results of shear capacity of typical samples

Sample Number	Calculated Value Vu (kN)	RFR Predicted Value (kN)
1	1278	1277.76
2	1349	1280.36
3	1421	1280.36
4	1249	1280.36
5	1582	1280.36
6	1519	1280.36
7	1479	1311.73
8	1409	1311.73
9	1651	1739.73
10	1721	1739.73

It can be clearly seen from Table 3-2 and Figure 3-1 that:

The RFR model predictions are highly consistent with the experimental values. For samples 1 and 2 to 8, the predicted values of the RFR model are very close to the calculated values, and the errors are within an acceptable range. For samples 9 and 10, although the predicted values are slightly higher than the experimental values, their changing trend is consistent with the experimental values, and they still reflect the characteristics of the high-capacity samples.

4. Conclusion

In summary, compared with traditional calculation methods such as the "Specifications for Seismic Design of Highway Bridges", the machine learning-based prediction models for the shear capacity of bridge piers, especially the Random Forest model, not only exhibit higher prediction accuracy and lower error on the overall dataset but also demonstrate good fitting capability in the prediction of specific typical samples. This provides a more accurate and efficient new approach for evaluating the seismic performance of bridge piers with shear capacity as a core indicator, and it is expected to play an important role in the future seismic design of bridges.

Fund Projects

Hunan Provincial College Student Innovation Training Project: Assessment of Structural Health Status of Old Rural Bridges Based on Machine Learning (S202511528144)

References

[1] Zhang Y. Q., Li J., Wang H. Research on seismic failure mode and shear strength formula of heavily reinforced railway gravity piers. Journal of Central South University (Science and Technology), 2025, 56(3): 890-898.
[2] Chen B. C., Lin Y., Huang F. Y. Research progress on application of ultra-high performance concrete in bridge engineering. China Journal of Highway and Transport, 2022, 35(6): 1-20.
[3] Wang P., Li Y., Zhang Y. Influence of shear performance of bridge piers on overall seismic safety of bridges. Journal of Southeast University (Natural Science Edition), 2023, 53(4): 678-687.
[4] Li J. Z., Wang Z. Q., Zhang M. Experimental study on influence of axial compression ratio and shear span ratio on shear performance of reinforced concrete bridge piers. China Civil Engineering Journal, 2022, 55(7): 102-110.
[5] Liu J., Chen L., Zhao Y. Analysis of influence of stirrup ratio and concrete strength on shear performance of plastic hinge region of bridge piers. Journal of Highway and Transportation Research and Development, 2023, 40(5): 68-75.
[6] Zhang L., Li L., Wang H. Finite element analysis and formula derivation of shear capacity of bridge piers considering damage accumulation. Journal of Building Structures, 2022, 43(11): 201-209.