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 the bridge pier component is an essential condition that directly affects the seismic performance of the pier and thus the earthquake resistance of the bridge. At present, 61.3 per cent of the short pier specimens in the database have a shear span ratio of less than 2.0, and 68.5 per cent of the specimens with a high axial compression ratio are greater than 0.4. There is no experimental data on irregular-shaped bridge piers of long-span bridges in areas with high-intensity earthquakes, and no experimental data on shear failure in the plastic hinge region under cyclic loading from strong earthquakes has been obtained. Generally speaking, the accuracy of the old formulas is lower than that of machine learning models.*

Keywords: *Bridge Piers, Seismic Performance, Shear Strength, Machine Learning, Bridge Engineering*

Introduction

Bridge piers are the basic support structures of a bridge, and therefore, the seismic response of the piers will affect the overall earthquake resistance of the bridge. Shear bearing capacity is a measure of the seismic performance of bridge piers and an indicator of shear failure[1]. UHPC has a very high strength, good toughness and durability, so it is now being used in the construction of bridges. It can improve the strength and durability of structures in a good way, and this is one of the future directions for the development of bridge structures[2]. As the principal load-bearing parts of bridges, the shear capacity of bridge piers is directly related to the seismic resistance of the entire bridge. If the shear strength is too small, the bridge pier will be damaged by a strong earthquake and the whole bridge will collapse[3].

Many experiments and finite element analyses of the shear behaviour of bridge piers have been conducted by Chinese and foreign scholars. Li Jianzhong and others[4] have carried out quasi-static tests on bridge piers with different axial compression ratios and shear-span ratios and found that a certain range of axial compression ratios can significantly increase the shear bearing capacity, and that the smaller the shear-span ratio, the higher the probability of shear failure for the bridge pier. Liu Jun and others[5] have shown that an increase in the stirrup ratio and concrete strength can reasonably improve the shear capacity of the plastic hinge zone in bridge piers and reduce the risk of brittle failure. Zhang Lei and others[6] have developed a reasonable finite element model of bridge piers, simulated the development process of shear damage in bridge piers under the influence of a strong earthquake, and proposed a calculation formula for the shear bearing capacity of bridge piers that considers the cumulative effect of damage.

A seismic database of 400 groups of bridge pier tests and numerical simulations has been constructed, shear bearing capacity has been taken as the main index, correlations among various parameters and their impact on the shear and seismic performance of bridge piers have been investigated, six machine learning models have been trained and validated through eight rounds of random sampling, their accuracy has been evaluated with R^2 and RMSE, and finally, they have been compared with the three traditional code methods. Based on the above results, the parameters in the current database are unevenly distributed, and there is a lack of data under some extreme working conditions; thus, a Random Forest model was used to obtain the highest accuracy ($R^2 = 0.99$, RMSE = 6.4). It has been improved significantly from the traditional model; that is to say, R^2 has increased by 216% and RMSE has decreased by 90%, so a new method for seismic performance assessment and design optimisation of bridge piers has been established.

1. Establishment of the Experimental Database

In order to obtain accurate predictions of the shear bearing capacity of bridge piers by means of machine learning, an experimental database of the seismic performance of bridge piers with various cross-sectional shapes, different material types, multiple design parameters, various strengthening methods and different working conditions needs to be built in this study. The four rules of database construction are all-encompassingness, representativeness, accuracy and normality. All kinds of information should be comprehensive, such as the geometry of piers, materials, reinforcement, strengthening, loading conditions, seismic performance indicators, etc.; representativeness is achieved by collecting data on commonly used structural forms and design parameter ranges for bridge piers in engineering; accuracy means that all data are based on the results of indoor model tests and full-scale tests published in academic papers and master's theses; normativity is expressed by using a single unit and uniform names and codes for all parameters.

2. Seismic Performance of Bridge Piers Based on Machine Learning

2.1 Machine Learning Models

The main direction of research on the seismic performance of bridges based on machine learning is to determine how much structural characteristic parameters affect the seismic performance of bridges through the key indicator of shear bearing capacity, and then to build a prediction model for the shear bearing capacity and seismic performance of bridges by means of regression training. Six typical supervised learning algorithms that will be used for the prediction of shear bearing capacity of bridges are introduced in this paper. The Calculation Principles of each algorithm are as follows:

Random Forest Regression (RFR) constructs several decision trees and combines the prediction results from them to build a model of non-linear relationships; it is relatively immune to overfitting and can be used with all kinds of engineering features.

b: Gradient Boosting Decision Tree (GBDT) is a kind of forward distribution that iteratively adds weak learners to correct the errors of the previous models based on gradient descent.

c. Adaptive Boosting (AdaBoost) is a kind of boosting method that repeatedly adds weak classifiers, adjusts the weights of the training samples to give more weight to difficult-to-classify ones, and thus improves the accuracy of the whole model.

d: Ridge regression is a type of regularised linear regression that does not have the unbiasedness property of Ordinary Least Squares; it is less accurate but more stable, and it can help to prevent overfitting.

e: Linear Regression (LR) is to minimize the sum of the squares of the residuals and find the best linear combination coefficients; it is a very simple type of linear model.

f: Lasso Regression - An L1 penalty term is added to the cost function to reduce model complexity and automatically select some features; it is often used when there are many features.

2.2 Calculation Results and Experimental Verification

2.2.1 Evaluation Indices

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

a. Coefficient of Determination (R^2): The degree of fit of the model to the data, which is between (-∞, 1). The closer the value is to 1, the better the model fits the data.

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

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 the models have good accuracy in the process, and the calculation results are relatively stable; all the algorithms take less than 1 minute, and the learning efficiency of each is 8 random-sampling trainings. The training results are as follows: Figure 1 and Table 1. As shown in Figure 1(a), most of the machine learning algorithms are better after 8 rounds of training and are suitable for fast prediction in engineering.

2.2.2.1 Model Performance Comparison

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

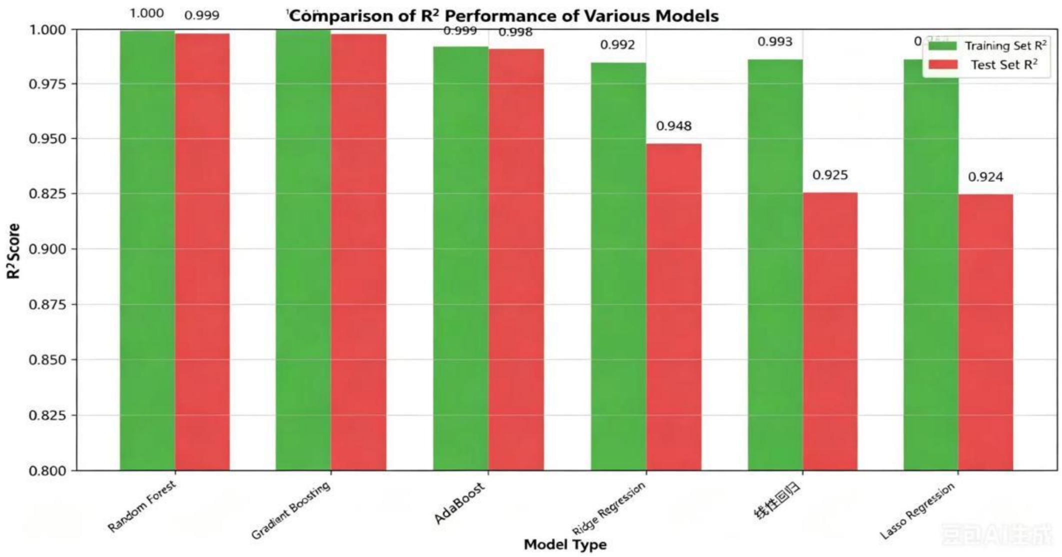

Figure 2-1

Random Forest Regression (RFR) is the best one; its training set is 1.000 and its test set is 0.999, and its fitting accuracy is almost perfect.

Gradient Boosting (GBDT) is next, and the training set =1.000 and the test set =0.998.

Linear models (Ridge Regression, Linear Regression, Lasso Regression) have test set 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 R^2 values of the training set and the test set for all models are close; therefore, it can be concluded that the models have good generalisation capabilities and have not been overfit.

The comparison of prediction errors for all the models is shown in Figure 2-2.

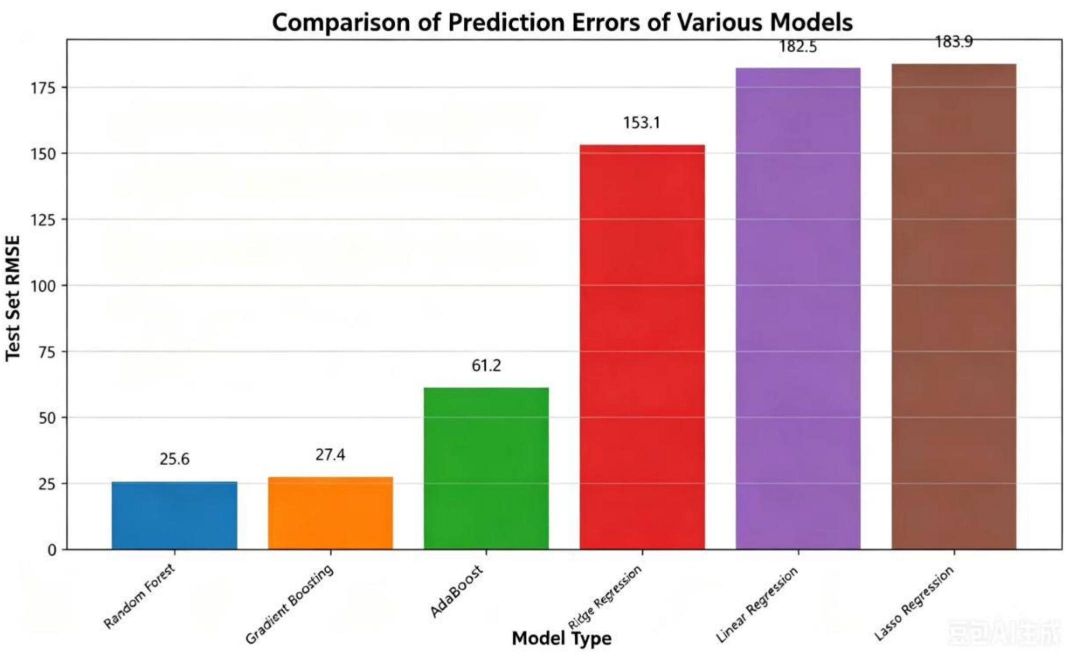

Figure 2-2

Random Forest Regression (RFR) has a test set RMSE of only 25.6, and this is the smallest prediction error.

The test set RMSE of Gradient Boosting (GBDT) is 27.4.

And it is close to that of RFR; the testing set RMSE of Adaptive Boosting (Adaboost) is 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 much higher than those of tree models and ensemble models.

2.2.2.2 Best Model Validation and Feature Analysis

Prediction effect of the best model (Random Forest): Based on the test set of the Random Forest model, the predicted values are in good agreement with the actual values and close to the ideal prediction line; therefore, it can be concluded that the model has captured the nonlinear relationship between bridge shear capacity and the input features accurately and is suitable for seismic performance evaluation.

Random Forest Top Feature Importance (as shown in Figure 2-3):

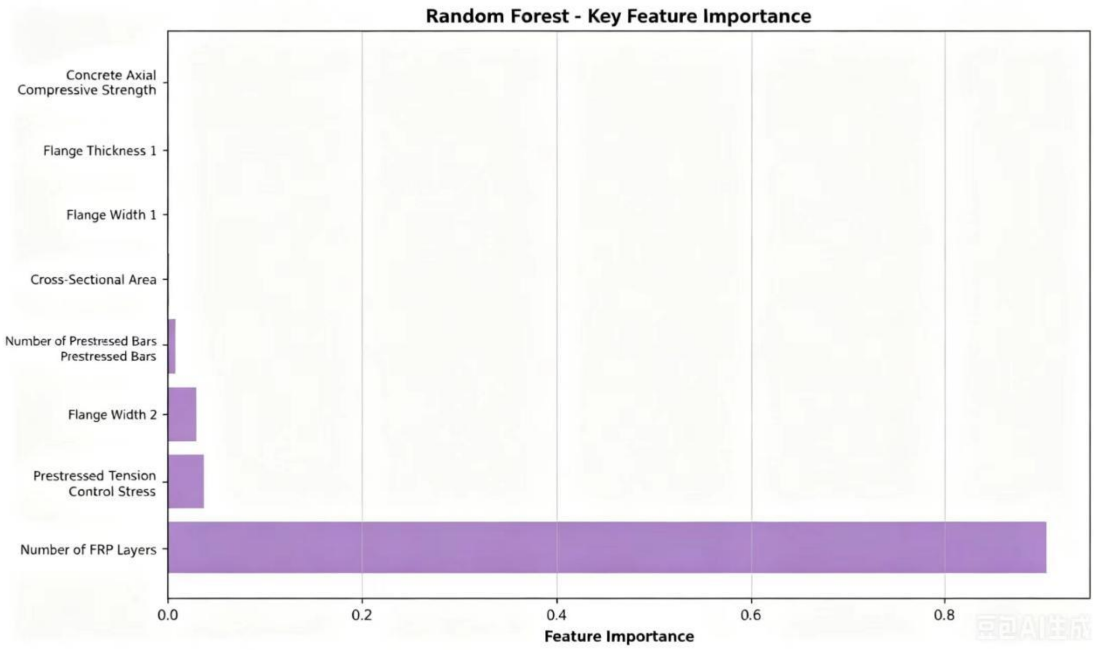

Figure 2-3

The number of FRP layers is the main reason for the change in the shear capacity of the bridge, and its relative importance is about 0.9; therefore, it can be concluded that the number of FRP reinforcement layers affects both the shear performance and the seismic resistance of the bridge to a certain extent.

2.2.2.3 Analysis of Reasons for Performance Differences

The reason for the good prediction accuracy of the algorithm may be its theoretical foundation and model design, etc.

a. Tree Models and Ensemble Models: Due to the non-linearity of bridge shear capacity, some decision trees or iterative optimisation methods must be applied; thus, Random Forest Regression (RFR), Gradient Boosting (GBDT) and Adaptive Boosting (Adaboost) are used for this purpose, as they have good prediction accuracy and stability. RFR is better than the others at preventing overfitting and performs well, so we have chosen it.

b: Ridge Regression, Linear Regression and Lasso Regression are all based on the assumption of linearity. They do not know the irregular nonlinear relationship of the parameters of the bridge structure and shear capacity exactly. Therefore, they are not very predictive. However, they have a fast training speed and are easy to learn, so they are suitable for the first round of rapid analysis.

c: Feature Impact: The number of FRP layers is the main reason; among all the factors, this one is relatively large, so FRP reinforcement technology can be employed to enhance both the shear strength and earthquake resistance of the bridge effectively. It agrees with the engineering knowledge we have.

3. Comparison 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 is mainly based on the current design codes in various countries. The general form of the above calculation formulas is to consider the concrete's compressive strength, shear span ratio, axial compression ratio and stirrup ratio as the main factors, along with material safety factors. The formula for calculating the shear capacity of bridge piers in China's "Specifications for Seismic Design of Highway Bridges" (JTG/T 2231-01-2020) is taken as the representative one in this paper and compared 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 in piers and columns of Class B and Class C bridges should be checked according to the following formula:

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

$$V_c=0.1\nu_c A_c$$

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

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

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

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—— Shear Strength of Concrete in the Plastic Hinge Zone (MPa);

f_{cd}—— The Design Value of Concrete Compressive Strength (MPa);

A_g—— Total Cross-sectional Area of the Plastic Hinge Region in the Pier Column (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 whole area of stirrups in the calculated direction. (cm^2);

h_0—— The distance between the compression edge of the core concrete and the centroid of the tension reinforcement. (cm);

s —— The spacing of stirrups (cm);

ϕ —— The factor of reduction for shear strength is taken as 0.85..

The formula takes into account the influence of concrete strength, axial compression ratio, stirrup ratio and member ductility demand on shear capacity in a comprehensive manner, and it is one of the most commonly used methods in current engineering design.

3.2 Comparative Analysis of Results

To systematically compare the prediction accuracy of all the machine learning models with that of the traditional code formula, this paper takes the existing database of 400 groups of pier seismic performance data and predicts the shear capacity of six machine learning models (Random Forest RFR, Gradient Boosting GBDT, Adaptive Boosting Adaboost, Ridge Regression, Linear Regression LR, and Lasso Regression) and the formula specified in the "Specifications for Seismic Design of Highway Bridges" (hereinafter referred to as the "code formula"). The code formula, material strength, geometric dimensions and reinforcement information required for the calculation are all obtained from the database, and the parameter values are set according to the demands of the code.

3.2.1 General Prediction Performance Comparison

Statistics of the prediction results for all 400 groups of sample are shown in this paper. The first two evaluation indices (R2R2 and RMSE) are presented 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 higher than that of the traditional code formula. Among them, the Random Forest (RFR) model has the best prediction accuracy and an R^2 of 0.99; its minimum RMSE is 6.4. The RFR model is about 216% more accurate than the best traditional formula, and the RMSE has been reduced by about 90%. Therefore, it can be seen that the machine learning model has learned a more precise, nonlinear mapping function from the data to predict the target variable and is thus much more accurate.

3.2.2 Example Sample Prediction Verification

To show the prediction results of the machine learning model on the actual data more clearly, this paper takes 10 groups of typical samples for comparison. Table 3-2 shows the comparison results of the calculated values for the shear capacity of these 10 groups of samples and the predicted values of the RFR model. The predicted values of RFR are the output of the trained model, and the calculated values can be obtained by using 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 prediction results of the RFR model are in line with the experiment. For samples 1 and 2-8, the predicted values of the RFR model are very close to the calculated values, and the errors are in a reasonable range. For samples 9 and 10, although the predicted values are slightly higher than the experimental values, the direction of change is in line with that of the experimental data, and they still have the characteristics of high-capacity samples.

4. Conclusion

In short, compared with the traditional calculation method in the "Specifications for Seismic Design of Highway Bridges", the prediction model for the shear capacity of bridge piers based on machine learning has a higher prediction accuracy and a smaller error over the entire dataset, and it also has good fitting performance for predicting some typical samples. It will provide a more accurate and effective new way to assess the seismic performance of bridge piers based on shear capacity, and it is hoped that this will be used in the seismic design of bridges in the future.

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.