Case-Based Reasoning for Predicting Bond Strength in Fiber Reinforced Polymer (FRP) and Concrete

Nadia Nassif[1], *M. Talha* Junaid[1,2*], *Salah* Altoubat[1,2], *Mohamed* Maalej[1,2], *Samer* Barakat[1,2], *Abdulrahman* Metawa[1], and *Raghad* Awad[3]

[1]Civil and Environnemental Engineering Department, College of Engineering, University of Sharjah, Sharjah, P.O.Box 27272, United Arab Emirates
[2] Sustainable Construction Materials and Structural System Research Group, Research Institute of Sciences and Engineering, University of Sharjah, Sharjah, United Arab Emirates.
[3] Department of Civil Engineering, McMaster University, 1280 Main St W, Hamilton, ON L8S 4L8, Canada

Abstract. The increasing adoption of Fiber Reinforced Polymer (FRP) bars in concrete structures necessitates accurate prediction of bond strength to ensure structural integrity and reliability. This study introduces Case-Based Reasoning (CBR) as an interpretable and efficient approach for predicting FRP-concrete bond strength. Utilizing a dataset of 227 experimental results, the CBR model achieves high accuracy, with an R^2 of 0.98 and a low Mean Squared Error (MSE) of 0.226 MPa. Sensitivity analysis identifies critical parameters such as bar diameter (d_b), concrete compressive strength (fc'), cover-to-bar diameter ratio (c/d_b), and embedment length-to-bar diameter ratio (l_d/d_b), demonstrating their varying influence across different surface types: helical lugged, spiral-wrapped, and sand-coated. The findings emphasize the practical applicability of CBR in tailoring design strategies for FRP-reinforced structures, offering engineers an interpretable and reliable tool to optimize performance while reducing computational complexity.

1 Introduction

The growing demand for sustainable construction materials has driven interest in fiber-reinforced polymer (FRP) bars as a replacement for traditional steel reinforcement in concrete structures. FRP bars offer superior corrosion resistance, a high strength-to-weight ratio, and durability in aggressive environments, making them well-suited for infrastructure applications where steel may be prone to deterioration. However, a critical factor for the effective performance of FRP-reinforced concrete is the bond strength between the FRP bars and the concrete matrix, which governs load transfer and structural integrity. Therefore, accurately predicting this bond strength is essential, particularly in conditions where

* Corresponding author: mjunaid@sharjah.ac.ae

environmental exposures vary significantly. Machine learning (ML) methods have emerged as powerful tools for predicting FRP-concrete bond strength, providing new avenues for enhancing predictive accuracy and understanding of the parameters that influence bond performance.

Studies on FRP-concrete bond strength prediction have leveraged various ML techniques to improve prediction accuracy over traditional empirical methods. Basaran et al. [1] explored a range of ML algorithms, including Gaussian Process Regression (GPR) and Artificial Neural Networks (ANN), in comparison with traditional code-based models. Their work highlighted GPR as the most accurate approach for bond strength prediction, particularly in the context of experimental methods such as pull-out and hinged beam tests. Furthermore, this study revealed that conventional code formulations tended to yield conservative estimates under diverse testing conditions.

In another study, Baghaei and Hadigheh [2] developed an ANN model with Bayesian optimization to assess bond durability under chemically aggressive environments, establishing that ML models could successfully capture the long-term performance characteristics of FRP-concrete interfaces. Similarly, Li et al. [3] proposed a hybrid model combining Extreme Learning Machine (ELM) with Particle Swarm Optimization (PSO) to predict bond strength accurately; a sensitivity analysis within their study identified the bar diameter-to-depth ratio as the most significant parameter affecting bond strength.

More recent studies have employed advanced ML techniques to address limitations in data variability and model interpretability. Zhang et al. [4] introduced a transfer learning framework that leverages natural language processing (NLP) for feature encoding, enhancing prediction accuracy even with limited datasets by facilitating feature generalization across diverse tabular data. Meanwhile, Kumar et al. [5] utilized ensemble learning models in combination with SHapley Additive exPlanations (SHAP) for feature importance, offering an automated solution for bond strength prediction across varied conditions (Winter, 2002). Tang et al. [6] examined the bond behavior of FRP bars in coral aggregate concrete, using SHAP to identify critical parameters impacting bond stress and failure modes.

To address the gaps identified in previous studies, we propose a novel approach leveraging Case-Based Reasoning (CBR) for predicting FRP-concrete bond strength. While ML methods such as ANN, ELM, GPR have demonstrated significant predictive capabilities, they often face challenges related to model interpretability, computational complexity, and the need for extensive datasets. Additionally, these methods are less effective at providing case-specific insights, which are critical for engineering decision-making in practical applications.

Unlike ensemble models or neural networks, where feature importance is often derived through post-hoc techniques (e.g., SHapley Additive exPlanations (SHAP)), CBR inherently incorporates feature significance into its predictions. By evaluating the similarity contributions of individual features, the approach offers an intuitive understanding of how each variable influences bond strength predictions.

2 Methodology

2.1 Data Collection and Preparation

This study utilized a dataset containing 227 experimental results from FRP bond strength tests. The database was gathered by [7] from 15 peer-reviewed studies. This dataset was successfully utilized by several researchers such as [8]. Variables included in the data set are:

- Surface type (*Surf*): Describes the condition or treatment of the surface. The values 1, 2, and 3 for the bar surface variable correspond to helical lugged, spiral-wrapped, and sand-coated surfaces of FRP bars.
- Bar diameter (d_b): Diameter of the reinforcing bar, measured in millimeters.
- Concrete compressive strength (fc'): Compressive strength of the concrete in megapascals (MPa).
- Concrete cover to bar diameter ratio (c/d_b): Ratio of the concrete cover thickness to the bar diameter.
- Development length to bar diameter ratio (l_b/d_b): Ratio of the development length of the reinforcement to its diameter.
- Bond strength (τ_b): The target variable, representing the bond strength in MPa.

The dataset was scrutinized for missing or inconsistent values, ensuring it was reliable for analysis. Variables were divided into input features and the target variable (τ_b), with no transformations applied to categorical variables (i.e., surface type). The statistical decstipion of the collected dataset is summarized in Table 1.

Table 1. Statistical description for the collected datasets.

	Surface	d_b (mm)	$f'c$ (MPa)	c/d_b (-)	l_d/d_b (-)	τ_b (MPa)
Min.	1.00	6.35	23.43	1.68	3.56	0.76
Max.	3.00	28.58	55.06	9.34	115.79	21.00
Average	1.72	14.65	40.04	3.58	30.14	6.79
STDEV	0.76	4.89	6.72	1.82	23.01	4.15

2.2 Data Normalization and Data Splitting.

To ensure all features contributed equally to the similarity computations, numerical attributes were normalized using the Min-Max scaling technique as shown in Eq. (1):

$$x_{norm} = \frac{x - x_{min}}{x_{max} - x_{min}} \qquad (1)$$

Where:

- x is the original feature value,
- x_{min} and x_{max} are the minimum and maximum values of the feature.

This step scaled all numerical features to a range of 0 to 1, preventing attributes with larger magnitudes from dominating the similarity calculations.

The dataset was randomly split into: The dataset was partitioned into a reference dataset (70%) for building the case base and a query dataset (30%) for evaluating the CBR predictive performance. This division ensured unbiased testing and generalizable results.

2.3 Model Selection

CBR mimics human problem-solving by drawing on past cases to address new ones. Its implementation involved the following key steps: similarity computation, case retrieval and prediction generation. In similarity computation, The similarity between a query case (q) and existing cases (c_i) was calculated using the Euclidean distance (d) (Eq. 2):

$$d(q, c_i) = \sqrt{\sum_{j=1}^{n} \left(q_j - c_{ij}\right)^2} \qquad (2)$$

Where:

- q_j and c_{ij} are the values of the j-th feature for the query and i-th case respectively,
- n is the number of features.

For each query case, the 5 closest cases (k) were identified from the training set based on their computed distances. Thereafter, the predicted bond strength ($\hat{\tau}_b$) for a query was calculated as a weighted average of the bond strengths ($\tau_{b,i}$) of the retrieved cases as shown in Eq. 3:

$$\hat{\tau}_b = \frac{\sum_{i=1}^{k} \frac{1}{d(q,c_i)+\epsilon} \cdot \tau_{b,i}}{\sum_{i=1}^{k} \frac{1}{d(q,c_i)+\epsilon}} \quad (3)$$

Where:

- ϵ is a small value to avoid division by zero,
- $d(q, c_i)$ is the distance between the query and the i-th case.

2.4 Variable Significance Analysis

The variable significance analysis was conducted to identify which input variables have the greatest influence on the $\hat{\tau}_b$. This was achieved through a sensitivity analysis, where each variable was varied across a defined range while holding the other variables constant at their mean values. The steps involved in this analysis are as follows: For each variable, a range of values was defined based on its observed range in the dataset. A series of test cases was created by varying the value of the target variable across its range while keeping other variables at their average values. The $\hat{\tau}_b$ was calculated for each test case using the CBR approach.
Thereafter, the variable influence was measured using Eq. (4) for the difference between maximum and minimum $\hat{\tau}_b$:

$$\Delta\tau_b = max(\tau_{b,\text{ predicted}}) - min(\tau_{b,\text{ predicted}}) \quad (4)$$

This $\Delta\tau_b$ was normalized by the range of the variable being analyzed (ΔVariable) to account for differences in variable scales as shown in Eq. (5):

$$\text{Significance Score } = \frac{\Delta\tau_b}{\Delta \text{ Variable}} \quad (5)$$

2.5 Model Evaluation

The model's performance was assessed on both training and testing datasets using the following metrics:

- Mean Squared Error (MSE):

$$MSE = \frac{1}{N}\sum_{i=1}^{N} \left(\hat{\tau}_{b,i} - \tau_{b,i}\right)^2 \quad (6)$$

- Mean Absolute Error (MAE):

$$MAE = \frac{1}{N}\sum_{i=1}^{N} \left|\hat{\tau}_{b,i} - \tau_{b,i}\right| \quad (7)$$

- Coefficient of Determination (R^2):

$$R^2 = 1 - \frac{\sum_{i=1}^{N} \left(\hat{\tau}_{b,i} - \tau_{b,i}\right)^2}{\sum_{i=1}^{N} \left(\tau_{b,i} - \bar{\tau}_b\right)^2} \quad (8)$$

Where:

- N is the number of cases,
- $\hat{\tau}_{b,i}$ and $\tau_{b,i}$ are the predicted and actual bond strengths for the i-th case,
- $\bar{\tau}_b$ is the mean of the actual bond strengths.

3 Results and Discussion

3.1 Accuracy

The CBR demonstrated high accuracy in predicting τ_b for FRP-concrete. The evaluation metrics for the query dataset (30%) are 0.226 MPa MSE, 0.248 MPa MAE and 0.98 R^2. These results indicate that the model provides consistent and accurate predictions across both datasets, with an R^2 value close to . This high predictive performance suggests that the CBR approach effectively captures the relationships between the input features and the bond strength. Moreover, the low error values highlight the robustness of the method in handling unseen cases.

3.2 Variable Significance Analysis

To identify the most influential factors affecting bond strength, a variable significance analysis was conducted. The results are summarized in Fig. 1. For helical lugged surfaces, all variables show relatively low importance scores, suggesting that bond strength is primarily influenced by the inherent mechanical interlocking of the surface. Among the factors, d_b and fc' appear to have a slightly higher influence compared to others, reflecting contributions of material and geometric properties. In contrast, spiral-wrapped surfaces display a broader distribution of variable importance, with l_d/d_b standing out as the most significant factor. Other variables, such as c/d_b and fc', also exhibit moderate importance, indicating that bond strength in spiral-wrapped surfaces is driven by a combination of geometric, material, and mechanical factors.

For sand-coated surfaces, the highest importance score is attributed to c/d_b, suggesting that bond strength is primarily governed by the quality of concrete cover surrounding the reinforcement. l_d/d_b also shows moderate importance, which is attributed to its role in ensuring sufficient anchorage, while d_b and concrete strength fc' have comparatively less influence. These findings emphasize the need for surface-specific optimization strategies. For sand-coated surfaces, optimizing c/d_b and l_d/d_b would be key to enhancing bond performance, while for spiral-wrapped surfaces, addressing multiple factors, especially l_d/d_b, would yield the best outcomes. Such insights offer valuable guidance for improving bond strength in FRP systems through tailored design approaches [9].

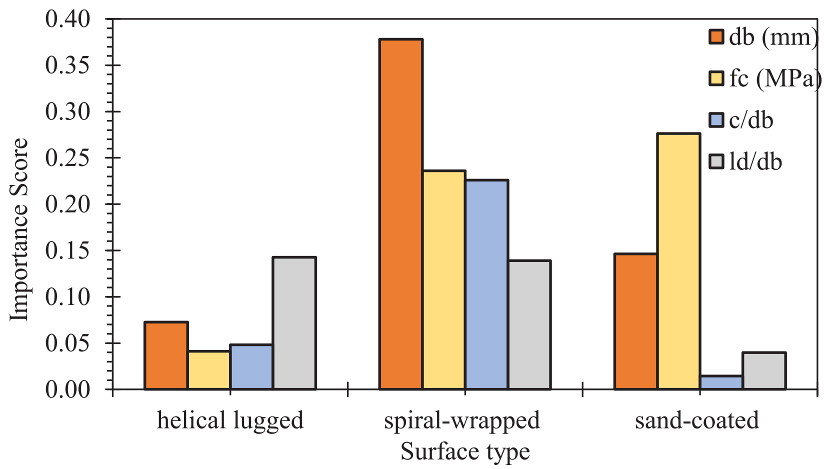


Fig. 1. Importance Scores for different.

3.3 Sensitivity Analysis

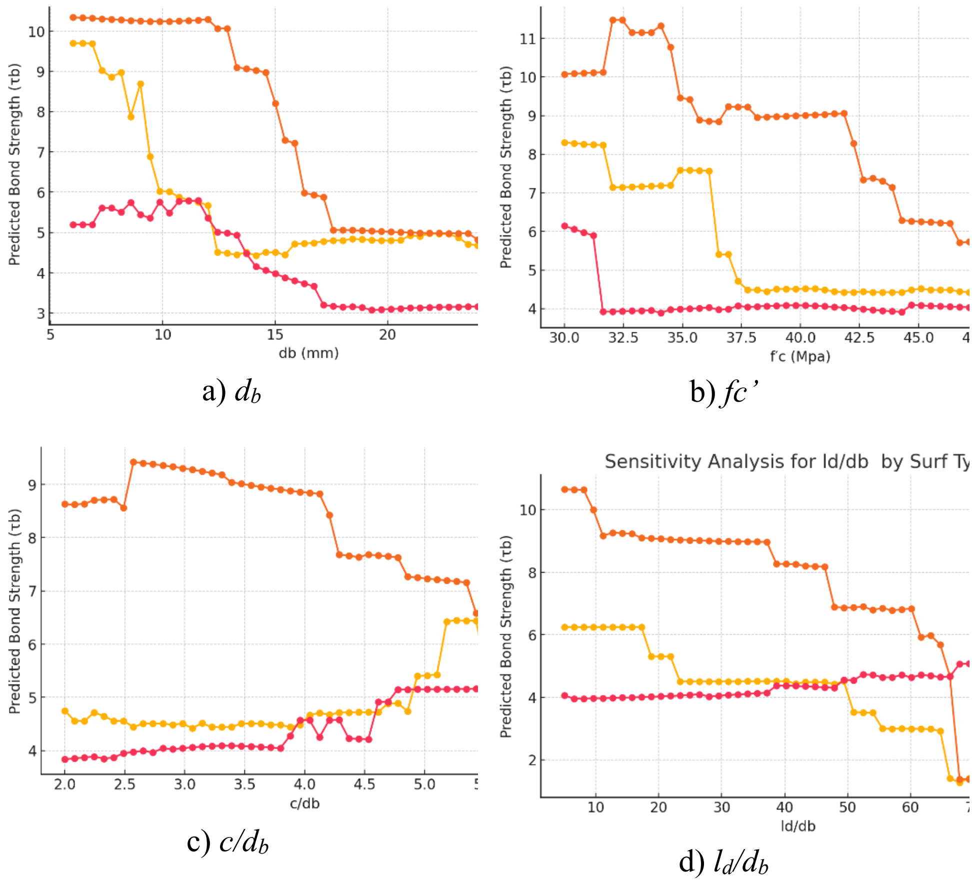
The sensitivity analysis results in Fig. 2 provide insights into how predicted τ_b responds to variations in individual variables across the three surface types: helical lugged, spiral-wrapped, and sand-coated. For d_b, bond strength decreases significantly as d_b increases, with the helical lugged surface showing the steepest decline, suggesting that smaller bar diameters are most effective in enhancing bond performance due to improved mechanical interlocking. Spiral-wrapped surfaces exhibit a gradual decline, while sand-coated surfaces show consistently lower bond strength across all diameters, likely due to weaker surface adhesion. For concrete compressive strength fc', bond strength decreases as fc' increases, particularly for helical lugged and spiral-wrapped surfaces, which may be attributed to reduced interfacial slip resistance associated with higher concrete stiffness. Sand-coated surfaces, on the other hand, exhibit stable but low bond strength, suggesting that surface coatings dominate their performance regardless of concrete strength.

For c/d_b, bond strength increases substantially with larger c/d_b ratios for helical lugged and spiral-wrapped surfaces, with the former demonstrating the highest sensitivity. This highlights the critical role of concrete cover in enhancing confinement and anchorage. In contrast, sand-coated surfaces show minimal improvement, indicating that their bond strength relies less on confinement and more on surface properties such as adhesion. For l_d/d_b, bond strength increases with greater embedment for helical lugged and spiral-wrapped surfaces, plateauing at higher values, which reflects sufficient anchorage beyond a certain embedment length. Sand-coated surfaces maintain consistently lower bond strength across the entire range of l_d/d_b, likely due to their dependence on surface-level interactions rather than geometric or mechanical factors. These findings emphasize the distinct bond mechanisms of the three surface types, underscoring the importance of tailoring design strategies to their specific behaviors.

3.4 Practical implications

The application of CBR for predicting FRP-concrete bond strength provides engineers and practitioners with a user-friendly and effective tool for decision-making in real-world projects. Practitioners can utilize CBR models to predict bond strength with minimal computational effort, as the method leverages existing data and past cases, eliminating the need for extensive computational resources or complex model training. This allows engineers to make quick, informed predictions for bond performance, particularly when designing reinforcement strategies for challenging environments such as coastal infrastructure or chemically aggressive settings.

By identifying the most influential variables, such as c/d_b for sand-coated surfaces or l_d/d_b for spiral-wrapped surfaces, practitioners can optimize design parameters to enhance structural performance. For example, engineers can prioritize increasing c/d_b when using sand-coated bars to improve bond performance or ensure adequate embedment length l_d/d_b for spiral-wrapped surfaces to achieve reliable anchorage. The interpretability of CBR also enables practitioners to trace predictions back to specific cases, fostering confidence in design decisions and facilitating validation against experimental or field data. This practical integration of CBR supports tailored, efficient designs that balance safety, performance, and cost-effectiveness.



a) d_b b) fc'

c) c/d_b d) l_d/d_b

Fig. 2. Sensitivity analysis for different variables at varying surface types.

4 Results and Discussion

This study demonstrates the effectiveness of Case-Based Reasoning (CBR) in predicting FRP-concrete bond strength, providing a transparent and interpretable alternative to traditional machine learning models. By leveraging case-specific insights, the CBR model achieves high predictive accuracy, making it a reliable tool for practical engineering applications. Sensitivity analysis highlights the distinct roles of key parameters—such as bar diameter, concrete compressive strength, concrete cover, and embedment length—across different FRP surface types. These insights enable practitioners to tailor design strategies to optimize bond performance, such as enhancing concrete cover for sand-coated bars or minimizing bar diameter for helical lugged surfaces. The simplicity and interpretability of CBR make it an accessible and efficient solution for real-world applications, particularly in scenarios requiring quick, data-driven decisions for sustainable and durable infrastructure. Future work can expand the dataset and explore hybrid approaches to further enhance predictive capabilities while retaining CBR's practical benefits.

References

1. Basaran, B., Kalkan, I., Bergil, E., and Erdal, E. (2021). "Estimation of the FRP-concrete bond strength with code formulations and machine learning algorithms." *Compos. Struct.*, **268**, 113972.
2. Baghaei, K. A., and Hadigheh, S. A. (2021). "Durability assessment of FRP-to-concrete bonded connections under moisture condition using data-driven machine learning-based approaches." *Compos. Struct.*, 114576.
3. Li, R., Liu, L., and Cheng, M. (2022). "Estimating the bond strength of FRP bars using a hybrid machine learning model." *Buildings*, **12**(10), 1654.
4. Zhang, P. F., Zhang, D., Zhao, X. L., Zhao, X., Iqbal, M., Tuerxunmaimaiti, Y., and Zhao, Q. (2024). "Natural language processing-based deep transfer learning model across diverse tabular datasets for bond strength prediction of composite bars in concrete." *Comput. Civ. Infrastruct. Eng.*
5. Kumar, A., Arora, H. C., Kumar, P., and Nehdi, M. L. (2024). "Automated explainable ensemble learning prediction of FRP-bar-to-concrete bond strength and failure pattern under diverse exposure scenarios." *Constr. Build. Mater.*, **445**, 137840.
6. Tang, Y., Zhou, C., Wang, W., Zhang, W., and Cheng, Y. (2024). "Machine learning driven bond performance prediction between FRP bars and coral aggregate concrete." *Constr. Build. Mater.*, **442**, 137684.
7. Bolandi, H., Banzhaf, W., Lajnef, N., Barri, K., & Alavi, A. H. (2019). An intelligent model for the prediction of bond strength of FRP bars in concrete: a soft computing approach. *Technologies*, *7*(2), 42.
8. Al-Hamd, R. K. S., Albostami, A. S., Alzabeebee, S., & Al-Bander, B. (2024). An optimized prediction of FRP bars in concrete bond strength employing soft computing techniques. *Journal of Building Engineering*, *86*, 108883.
9. Nassif, N., Talha Junaid, M., Maalej, M., Altoubat, S., Barakat, S., & Metawa, A (2025). Fiber-Reinforced Polymer (FRP) Bar Bond Strength Prediction Using SHapley Additive exPlanations (SHAP)-Optimized Artificial Neural Network (ANN). In *CISSC 2025* (pp. 102-110).