Journal of Soft Computing in Civil Engineering 10-3 (2026) 1980

Contents lists available at SCCE

Journal of Soft Computing in Civil Engineering

Journal homepage: www.jsoftcivil.com

Optimizing CFRP Thickness and Fiber Orientation for Seismic Retrofitting Using LightGBM and Whale Optimization Algorithm

Ahmed Salama [1,2,*]; **Ahmed Hassan Hassan** [3]; **Njood Aljarrah** [4]

1. Assistant Professor, Department of Civil Engineering, Faculty of Engineering, Ajloun National University (ANU), Ajloun 26810, Jordan
2. Assistant Professor, Department of Civil Engineering, Al-Azhar University, Nasr City 11884, Cairo, Egypt
3. Assistant Professor, Department of Cybersecurity, Faculty of Information Technology, Middle East University (MEU), Amman, Jordan.
4. Part-time Lecturer, Department of Computer Sciences, Faculty of Information Technology and Computer Sciences, Yarmouk University (YU), Irbid, Jordan

* Corresponding author: *A.Elsayedsalama@anu.edu.jo*

https://doi.org/10.22115/scce.2025.1980

ARTICLE INFO

Article history:
Received: 11 November 2024
Revised: 04 June 2025
Accepted: 22 September 2025

Keywords:
Whale optimization algorithm (WOA);
Light GBM;
Carbon fiber reinforced polymer (CFRP);
Seismic retrofitting; optimizing;
Machine learning.

ABSTRACT

This study aims to optimize Carbon Fiber Reinforced Polymer (CFRP) configurations—specifically, thickness and fiber orientation—for seismic retrofitting of reinforced concrete (RC) frames. To achieve this, the Whale Optimization Algorithm (WOA) is integrated with a Light Gradient Boosting Machine (LightGBM) model to enhance predictive accuracy and minimize seismic damage indicators such as inter-story drift and base shear. A synthetically generated dataset comprising over 10,000 samples was used to train and evaluate the model under varying structural and seismic conditions. Before optimization, the model achieved a Mean Absolute Error (MAE) of 0.005 and an R^2 score of 0.89 on the testing set. After WOA-based hyperparameter tuning, the MAE decreased to 0.004 and the R^2 improved to 0.92. The optimization also reduced inter-story drift from 0.035 to 0.029 and base shear from 2200 kN to 1950 kN. These results demonstrate the effectiveness of combining WOA and LightGBM in simultaneously optimizing CFRP parameters. The proposed framework offers a computationally efficient, data-driven approach to retrofitting design, with improved accuracy and generalizability compared to traditional methods. The model's adaptability across various seismic intensities and structural profiles highlights its potential for broader application in performance-based seismic design.

How to cite this article: Salama A, Hassan A, Aljarrah N. Optimizing CFRP thickness and fiber orientation for seismic retrofitting using lightGBM and whale optimization algorithm. Journal of Soft Computing in Civil Engineering. 2026;10(3):1980. https://doi.org/10.22115/scce.2025.1980

2588-2872/ © 2025 The Authors. Published by Pouyan Press.
This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

Carbon Fiber Reinforced Polymer (CFRP) composites are gaining recognition for their exceptional mechanical properties, including high strength-to-weight ratios and superior fatigue resistance, making them highly suitable for reinforcing structural elements exposed to seismic forces. Aljabbri [1] examines the numerical modeling of beam-column joints strengthened with CFRP, demonstrating that specific configurations can enhance structural performance under seismic loads. The study highlights the significance of reinforcement details and the application of external CFRP layers, which can significantly affect the seismic behavior of structural joints. Similarly, Kang et al. [2] report that CFRP grid reinforcement can greatly improve energy dissipation capacity, a key factor in mitigating seismic forces during earthquakes.

The orientation of CFRP fibers plays a crucial role in determining their mechanical performance. Medeghini et al. [3] emphasize that fiber orientation relative to the loading direction profoundly impacts the residual tensile strength of fiber-reinforced composites. This finding is especially pertinent for seismic applications, where the direction of applied forces can vary significantly. The ability to control fiber orientation through advanced manufacturing techniques allows for customized mechanical properties that can be optimized using machine learning algorithms to predict performance under various loading conditions.

Machine learning techniques provide a powerful tool for analyzing extensive datasets generated from experimental and simulation studies, enabling the identification of optimal CFRP configurations. Metaheuristic algorithms, for instance, can systematically explore design parameters such as CFRP thickness and fiber orientation. These methods facilitate the development of predictive models that inform design decisions, ultimately enhancing structures' seismic performance. Integrating these computational techniques with experimental validation ensures optimized designs are practical and effective in real-world applications.

Furthermore, structures' seismic resilience can be significantly enhanced by adopting hybrid systems that combine CFRP with other materials. Li [4] investigates the performance of steel-PVA hybrid fiber cementitious composites within concrete-filled steel tubes (CFST), demonstrating that increasing core concrete strength and adjusting the axial compression ratio improve the seismic performance of composite columns. This finding suggests that hybridization and optimized CFRP configurations can significantly enhance structural resilience.

Advanced numerical methods, such as finite element analysis (FEA), play a critical role in optimizing CFRP applications. Wang et al. [5] demonstrate that FEA, incorporating fiber elements, can accurately simulate the complex interactions in reinforced concrete structures under seismic loading. By integrating CFRP into these models, researchers can simulate various scenarios to assess the influence of CFRP thickness and fiber orientation on structural performance. This approach provides a deeper understanding of how CFRP configurations can be optimized for specific seismic conditions.

Experimental studies are pivotal in validating theoretical predictions derived from numerical models. Golias et al. [6] provide empirical evidence supporting the effectiveness of X-shaped CFRP ropes in improving the hysteretic behavior and energy dissipation of beam-column joints under cyclic loading. Such validations are crucial for refining optimization algorithms and ensuring that the proposed designs meet performance criteria.

The growing body of research on fiber-reinforced concrete (FRC) and its seismic performance further supports the application of machine learning and metaheuristic algorithms in optimizing CFRP thickness and fiber orientation. Studies have demonstrated that fibers can significantly enhance the ductility and energy absorption capacity of concrete structures [7,8]. Leveraging these insights, researchers can develop machine learning models to predict optimal fiber configurations, facilitating the design of more resilient structures.

Additionally, the integration of distributed sensing technologies, such as fiber optic sensors, offers real-time monitoring of structural health during seismic events. These data are invaluable for refining optimization algorithms and improving the predictive capabilities of machine learning models. Lindsey et al. [9] explore distributed acoustic sensing (DAS) for seismic monitoring, which enhances the understanding of structural responses to seismic forces and informs future design improvements.

Studies have shown that CFRP retrofitting can substantially increase peak load capacity, initial stiffness, and ductility in reinforced concrete (RC) structures. For example, Zhou demonstrated that CFRP retrofitting led to a 43.89% increase in peak load, a 39.27% increase in initial stiffness, and a 30.1% improvement in ductility in RC frames subjected to seismic loading [10]. This improvement is primarily due to the high strength-to-weight ratio of CFRP, which effectively confines concrete elements, enhancing their load-bearing capacity and energy dissipation during seismic events [11].

Optimizing CFRP thickness and fiber orientation is essential for maximizing the benefits of retrofitting. Jafari and Mahini explored various CFRP configurations at structural joints and revealed that specific arrangements enhance the seismic fragility capacity of RC frames [12]. Strategic placement of CFRP can improve energy dissipation and lateral capacity, both critical for maintaining structural integrity during earthquakes.

Incorporating machine learning techniques, such as LightGBM, into the optimization process provides a data-driven approach to determining the most effective CFRP configurations. LightGBM, well-suited for large datasets, efficiently models complex relationships between CFRP parameters (e.g., thickness and fiber orientation) and structural performance metrics (e.g., load capacity and ductility). By training models on experimental data, researchers can predict the performance of various CFRP configurations and identify optimal solutions for maximizing seismic resilience [13].

The Whale Optimization Algorithm (WOA) complements the machine learning approach by offering a robust optimization framework. WOA, inspired by whale hunting behavior, effectively explores the solution space to identify optimal CFRP retrofitting parameters. When applied to optimizing CFRP configurations, WOA helps identify the best combinations of thickness and fiber orientation, yielding significant performance improvements in seismic scenarios. The synergy between WOA and LightGBM leads to more efficient and effective retrofitting strategies, ultimately enhancing the safety and durability of RC structures. Experimental validations further support the theoretical and computational findings. For instance, Laseima et al. demonstrated significant improvements in the shear efficiency and overall seismic performance of RC beam-column joints retrofitted with CFRP sheets [14]. These findings underscore the importance of CFRP configuration, as variations in orientation and thickness significantly affect structural performance. Beyond performance enhancements, CFRP retrofitting addresses existing structural deficiencies. Hejazi's research on flange-bonded CFRP sheets highlights their role in relocating plastic hinges in beam-column joints, improving stability and load capacity in retrofitted frames [15]. This is particularly valuable for older structures that may not meet contemporary seismic design standards.

As the field of structural engineering evolves, the combination of advanced materials, computational techniques, and real-time monitoring will play a critical role in optimizing CFRP applications for seismic resilience. Ongoing research underscores the importance of interdisciplinary collaboration, uniting materials science, structural engineering, and data science to address seismic challenges. Despite the substantial advancements in CFRP applications for seismic retrofitting, previous studies have either focused on simulation-based optimization or experimental validations of individual parameters such as thickness or orientation. The novelty of this study lies in the dual-level integration of a machine learning model (LightGBM) with the Whale Optimization Algorithm (WOA), allowing for a simultaneous and data-driven optimization of CFRP thickness and fiber orientation. This combination not only improves predictive accuracy but also significantly reduces computational time, providing a more efficient, scalable, and intelligent framework for structural retrofitting design. Additionally, the use of a large, synthetically generated dataset that covers diverse seismic and structural conditions enhances the generalizability and applicability of the proposed approach to real-world scenarios.

While numerous studies have explored the use of CFRP in enhancing seismic resilience—either through finite element simulations [1,5] or experimental validation of specific configurations [6,10,14]—most of these works address parameter effects in isolation and lack integrated optimization strategies. Moreover, existing machine learning applications in this context are often limited to predictive modeling [20–23] without exploring metaheuristic-enhanced design frameworks. This reveals a critical knowledge gap in the simultaneous, data-driven optimization of CFRP retrofitting parameters under varying seismic conditions. Motivated by this gap, the present study proposes a combined LightGBM–WOA framework that not only predicts structural performance with high accuracy but also determines optimal CFRP configurations that minimize seismic damage indicators such as inter-story drift and base shear. Integrating experimental validation, numerical modeling, and real-time monitoring is crucial for developing effective optimization strategies. Future research should focus on refining these approaches and exploring new materials and configurations to further improve the seismic performance of reinforced structures.

The novelty of this research lies in its dual-focus methodology: it not only leverages the predictive strength of LightGBM but also incorporates the optimization power of WOA to provide a holistic design solution. Unlike earlier studies, this integrated framework enables simultaneous prediction and optimization of CFRP parameters, resulting in a retrofitting approach that is both adaptive to diverse seismic demands and efficient in resource utilization. This study aims to optimize CFRP thickness and fiber orientation through machine learning and metaheuristic algorithms and offers a promising avenue for enhancing structural resilience under seismic loads.

To guide the reader through the study, the remainder of this paper is structured as follows: Section 2 presents the dataset description, including the synthetic generation methodology and parameter ranges used in the simulations. Section 3 outlines the methodology, detailing the integration of the LightGBM model with the Whale Optimization Algorithm (WOA) and the evaluation metrics used. Section 4 discusses the results, including predictive performance, optimization outcomes, and comparative analysis. Section 5 provides an in-depth discussion highlighting the advantages, limitations, and practical implications of the proposed framework. Finally, Section 6 concludes the study with a summary of key findings, contributions, and suggestions for future research directions.

2. Research methodology

2.1. Dataset description

The dataset used for this regression forms the backbone of the entire development process, as it guides in the development of the machine learning model through optimization. The dataset has been carefully designed to encompass a wide range of parameters essential for understanding the structural behavior of RC frames during seismic events. The dataset consists of over 10,000 synthetic data points considering several structural configurations, material properties, seismic demand parameters, and retrofitting strategies [16]. The synthetic dataset was generated using a parametric simulation framework that integrates probabilistic modeling of structural dimensions, material properties, and seismic demand variables. Latin Hypercube Sampling (LHS) was employed to ensure uniform coverage across the multidimensional parameter space, maintaining realistic correlations between features. The data reflect plausible physical configurations based on code-compliant ranges derived from existing literature and engineering standards [17,18]. Several key feature groups form the structure, each giving insight into multiple features of the RC frame seismic response.

It comprises two categories: the first relates to Structural Features that describe the geometry and the material properties of the RC frame. These vary between the frame height, which falls between 10 and 30 meters, and the number of stories, which range between 3 to 10. Further refinement gives a story height range of 2.5 to 3.5 meters and a span length ranging between 5 to 10 meters, giving a fine view of the general structure. Furthermore, beam widths and depths, as well as column widths and depths, represent an important dataset regarding the load-carrying elements of the structure, while properties characterizing the resistance of concrete in the range of 25 to 50 MPa—and rebars in the range of 400 to 600 MPa—represent critical material properties, along with the reinforcement ratio ranging between 1% and 4%, influencing the structural stiffness and capacity to resist seismic forces.

It represents the second group, consisting of the CFRP features that are the central optimization parameters in the study. The most common retrofitting materials are CFRP, and under seismic loads, their performance is highly dependent on two major variables: the CFRP material's thickness and the fiber's orientation. The dataset contains variations in the thickness of the CFRP materials from 1 to 5 mm, along with variations in fiber orientation angles from 0° to 90°. These parameters directly influenced how the CFRP strengthened the RC frame, and indeed, different configurations offered differing degrees of enhancement to the seismic resilience of the structure.

The third set would be the Seismic Demand Parameters, representing a quantity characterizing the external forces imposed on the structure due to the earthquake. It includes PGA, representing the maximum ground motion of the earthquake in terms of gravitational acceleration in the range between 0.1g and 1.0g, and spectral acceleration in the range between 0.5g and 3.0g, indicating the structure's response at specific frequencies. Magnitude of the earthquake (Richter scale) ranges from 5.0 to 8.0, indicating a broad measure of the overall severity of the earthquake; ground motion duration ranges from 10 to 60 seconds, and frequency content is categorized as either low or high.

Other seismic features of importance include the damping ratio, between 2% and 7%, a measure of the dissipative ability of the structure, as well as distance to fault, from 5 up to 50 km, related to the intensity of the seismic load experienced by the structure. The condition of the soils can be either soft, medium, or rock to take into account the amplification of the ground motion.

The Damage Indicators group focuses on two important metrics related to the structural response due to seismic events: inter-story drift ratio and base shear force. Inter-story drift ratio defines the relative displacement of adjacent floors and is a more important indicator; higher values will imply greater potential damage. Other significant measures of the structure's capacity to resist seismic loads include base shear force, which is defined as the total lateral force at the base. Both of these parameters will be considered important for assessing different CFRP configurations in the effectiveness of seismic damage mitigation. The key outcome variable that the machine learning model trains to predict corresponds to the Optimization

Target, representing the corresponding reduction in inter-story drift due to specified configurations of CFRP thickness and fiber orientation. By optimizing this target, the model should be able to find the properties of CFRPs which minimize seismic damage, hence a powerful tool for guiding retrofitting strategy in RC structures. The dataset is, in fact, designed to be a rich starting point for the training of machine learning and the realization of the optimization process. The features, related to the detailed structural, CFRP properties, seismic demand parameters, and damage indicators, ensure the proper modeling of the complex interactions among these variables to offer appropriate and efficient seismic retrofitting solutions.

2.2. Feature engineering

Feature engineering is one of the most crucial steps for any model development in machine learning, as it forms the direct basis of a model's capability to understand and make predictions. In this analysis, the prepared raw dataset had to be surveyed judiciously so that any model may learn something from the input data efficiently. It consisted of a series of steps in order: normalization, handling missing value, feature selection; all these together gave the final dataset used for training the model, LightGBM [19,20].

Normalization has been done on the dataset to ensure features from different units and scales do not disproportionately affect model learning. For example, the structural dimensions were measured in mm and seismic demand parameters included Peak Ground Acceleration, measured in terms of g-gravitational acceleration. As most of these features will be continuous in nature, they needed normalization into a standard range-usually 0 to 1 using min-max scaling [18,21].

This step did not allow any one feature to be out-weighing the others because of its numerical scale and allowed a model to learn the relative importance of each variable without bias. Handling missing values in this dataset proved to be another important approach to feature engineering. Since this was a synthetically generated dataset that was very complete, whatever missing or wrong values were found needed detection and treatment to avoid bias or error in the model. Whatever needed it, simple imputations were used: missing value imputation with the median or mean of the respective feature. This helped retain integrity in the dataset while keeping enough variability for model training. Meanwhile, feature selection played the most important role in refining the dataset for optimal model performance. Several structural, seismic, and CFRP-related features are not contributing equally to the prediction of the reduction in inter-story drift. Therefore, a few strategies were enforced based on the correlation analysis and feature importance ranking from preliminary machine learning models. These methods therefore captured those features most strongly related to the optimization target of reduced inter-story drift, while minimizing noise from less-informative features [22].

Based on the selection, the most critical features identified for the study were CFRP thickness and fiber orientation. Thickness in CFRP, which ranges from 1 to 5 millimeters, directly influences the stiffness and strength in the RC frame. Greater reinforcement with thicker layers of CFRP usually reduces structural deformation during seismic events. However, for greater thickness, the gain in performance is usually at a diminishing rate and also more costly; optimization will therefore be imperative to derive the right balance between effectiveness and practicality. Orientation of the fibers, from 0° to 90°, affects how CFRP fibers interact with seismic forces.

Different orientations alter the mechanical and behavioral performance of the reinforced concrete. Where other angles may provide better resistance to lateral forces, the angle of fiber alignment can enhance structural ductility and energy dissipation, reducing the possibility of brittle failure under a strong seismic event. Besides the CFRP-related features, a number of seismic demand parameters have also been selected due to their considerable influence on the structural response. In this respect, the study directly included as factors the direct indicators of seismic event intensity PGA and spectral acceleration. These parameters are critical because they define the magnitude of the forces acting on the structure; hence, they determine the effectiveness of the applied reinforcement with CFRP. Other parameters to retain for investigation included earthquake magnitude, ground-motion duration, and soil conditions, including these factors in the dynamic response modification of the RC frame.

The structural features also included beam and column dimensions, concrete strength, and rebar strength on feature selection because they are related to the inherent seismic resilience of the RC frame. Such a combination of structural properties with the CFRP parameters enables modeling to acquire how different configurations of material properties and geometries interact with seismic load, hence normally developing a more holistic understanding of the retrofitting process. In this respect, preparation of the data for training the LightGBM model was done by selecting and engineering features that capture the core interactions between the properties of CFRP, structural characteristics, and seismic forces. Also, such a process of normalization, missing data handling, and feature selection has been done with cares, ensuring the proper prediction of the model on inter-story drift reduction to guide the optimization of CFRP thickness and fiber orientation in seismic retrofitting.

2.3. Machine learning model: lightGBM

This study selected the LightGBM (Light Gradient Boosting Machine) algorithm as the core machine learning model due to its ability to handle large datasets with multiple features and its efficiency in performing complex optimization tasks. LightGBM is a gradient-boosting framework that utilizes decision trees as base learners, allowing it to effectively model non-linear relationships between features. It excels in scenarios with high-dimensional data and large datasets, making it ideal for predicting complex outcomes such as the seismic performance of reinforced concrete (RC) frames, where interactions between CFRP properties and seismic demand parameters are intricate and multi-faceted [17].

LightGBM offers several key advantages that make it suitable for this study. First, it is highly efficient in terms of both memory usage and computational speed. It implements a leaf-wise tree growth strategy, which allows for deeper trees and better accuracy compared to level-wise growth strategies used by other gradient-boosting frameworks. This approach helps capture intricate patterns and relationships between features like CFRP thickness, fiber orientation, and seismic demand parameters. Additionally, LightGBM can handle categorical data, missing values, and highly skewed feature distributions, making it robust for real-world applications like seismic retrofitting optimization [23].

Furthermore, LightGBM uses gradient-based one-sided sampling (GOSS) and exclusive feature bundling (EFB) to handle large datasets more efficiently. GOSS helps in sampling instances that contribute more to the gradient, while EFB bundles features with low mutual information to reduce the dimensionality of the feature space. Both improve computational efficiency without sacrificing predictive accuracy [24].

The primary task of the LightGBM model in this study is to predict the optimization target, which is the reduction in inter-story drift ratio due to different CFRP configurations. This reduction is a function of CFRP thickness, fiber orientation, and other structural and seismic parameters. LightGBM is trained to model the non-linear relationship between these inputs and the optimization target. The training process minimizes a loss function that measures the error between the predicted and actual values of the reduced drift [25].

The objective function for LightGBM in regression tasks (such as predicting reduced inter-story drift) is typically based on minimizing the Mean Squared Error (MSE) [24,26], defined as Eq. (1).

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \quad (1)$$

Where:

- y_i is the actual reduced drift for the i^{th} observation,
- $\hat{y}_i$ is the predicted reduced drift for the i^{th} observation,
- n is the number of observations in the dataset.

In each iteration of the training process, LightGBM updates its decision trees to minimize the MSE by following the gradient of the loss function. The gradient boosting approach ensures that the model improves iteratively, with each new tree correcting the residuals (errors) from the previous iteration. The model also learns the importance of features by calculating how much each feature contributes to reducing prediction errors. Features such as CFRP thickness and fiber orientation are expected to have high importance scores because they are the primary parameters affecting the structural response during seismic events [27].

2.4. Dataset splitting and cross-validation

The dataset is divided into training and testing sets to ensure the model can generalize well to unseen data. The training set is used to fit the LightGBM model, allowing it to learn the relationships between the features and the target variable. The testing set held during training is then used to evaluate the model's predictive performance on new, unseen data [28]. This study split the dataset into training and testing sets using an 80/20 split. This means that 80% of the data was used for model training, while the remaining 20% was reserved for testing. This ensures that the model's performance can be evaluated regarding its ability to predict reduced drift for data it has not seen before, providing an unbiased assessment of its generalization capacity.

To further improve the model's robustness and avoid overfitting, k-fold cross-validation was applied during the training process [29]. Cross-validation involves dividing the training data into k subsets (folds), training the model on k-1 folds, and validating it on the remaining fold. This process is repeated k times, with each fold used exactly once for validation. The final model performance is averaged over all folds, ensuring that the model is not overly sensitive to any particular subset of the data.

Mathematically, cross-validation for the MSE can be represented as Eq. (2).

$$\text{CV-MSE} = \frac{1}{k}\sum_{j=1}^{k} \frac{1}{n_j}\sum_{i=1}^{n_j} \left(y_i^{(j)} - \hat{y}_i^{(j)}\right)^2 \quad (2)$$

where:

- k is the number of folds,
- n_j is the number of observations in the j^{th} fold,
- $y_i^{(j)}$ and $\hat{y}_i^{(j)}$ are the actual and predicted reduced drift for the i^{th} observation in the j^{th} fold.

This approach enhances the model's ability to generalize by ensuring that it performs well across different subsets of the data, making it more robust and reliable for real-world applications.

2.5. Whale optimization algorithm (WOA)

The Whale Optimization Algorithm (WOA) is a nature-inspired metaheuristic optimization technique that simulates the social hunting behavior of humpback whales. Humpback whales use a unique hunting method known as "bubble-net feeding," where they create spiral-shaped bubbles around their prey. WOA mimics this spiral movement and is designed to solve optimization problems by iteratively improving a candidate solution [30–32]. In this study, WOA plays a crucial role in hyperparameter tuning for the LightGBM model and in optimizing the key variables—CFRP thickness and fiber orientation—to minimize inter-story drift and base shear forces in RC frames subjected to seismic loads.

2.5.1. Overview of WOA

The WOA operates by modeling the hunting strategy of humpback whales, which can be mathematically simplified into three main phases:

1. Encircling prey: In this phase, the algorithm assumes that the optimal solution is the target (prey) and the whale's position is updated iteratively to move closer to the optimal solution.
2. Bubble-net attacking: This represents the exploration phase where the whales search for better solutions, inspired by the spiral-shaped bubbles they create.
3. Search for prey: This is the exploitation phase, where the whales attempt to locate the prey by updating their positions based on the global best solution found so far.

The position of a whale in the solution space is represented by candidate values for the variables being optimized. In this study, those variables are the CFRP thickness and fiber orientation.

2.5.2. Mathematical model of WOA

WOA uses the following mathematical equations to simulate the encircling and spiral updating behaviors:

1. Encircling prey: Whales move toward the best solution found so far. The position of each whale is updated according to Eq. (3).

$$\begin{gathered}\vec{D} = \left|\vec{C} \cdot \vec{X}^{*}(t) - \vec{X}(t)\right| \\ \vec{X}(t+1) = \vec{X}^{*}(t) - \vec{A} \cdot \vec{D}\end{gathered} \tag{3}$$

where:

- t is the current iteration,
- $\vec{X}(t)$ is the position of the whale at iteration t_t
- $\vec{X}^{*}(t)$ is the position of the current best solution,
- $\vec{A}$ and $\vec{C}$ are coefficient vectors,
- $\vec{D}$ is the distance between the whale and the best solution found so far.

The coefficient vectors $\vec{A}$ and $\vec{C}$ are computed as Eq. (4).

$$\vec{A} = 2 \cdot \vec{a} \cdot \vec{r} - \vec{a},\ \vec{C} = 2 \cdot \vec{r} \tag{4}$$

where $\vec{a}$ is a linearly decreasing vector from 2 to 0 over the course of iterations, and $\vec{r}$ is a random vector in the range $[0,1]$.

2. Spiral updating: The spiral behavior is simulated by creating a helix-shaped movement toward the prey, given by Eq. (5).

$$\vec{X}(t+1) = \vec{D}' \cdot e^{bl} \cdot \cos(2\pi l) + \vec{X}^{*}(t) \tag{5}$$

Where:

- $\vec{D}' = \left|\vec{X}^{*}(t) - \vec{X}(t)\right|$.
- b is a constant defining the shape of the spiral,

- l is a random number in the range $[-1,1]$.

Exploration and exploitation: WOA switches between exploitation (searching around the best solution) and exploration (searching new areas of the solution space) based on a probability p. This dynamic balance between exploration and exploitation allows WOA to effectively avoid local minima and search the entire space.

2.5.3. WOA for hyperparameter tuning in lightGBM

WOA is applied in this study to perform hyperparameter tuning for the LightGBM model. LightGBM has several hyperparameters that influence its performance [33–35], such as:

- Learning rate: Determines the step size during the gradient descent update.
- Max depth: Controls the maximum depth of each decision tree.
- Number of leaves: Limits the number of leaves in each tree.
- Subsample: The fraction of data used to build each tree.
- Feature fraction: The fraction of features used in each iteration.

These hyperparameters significantly affect the model's ability to capture the relationship between CFRP parameters and seismic performance. WOA optimizes these hyperparameters by minimizing a predefined loss function (such as Mean Squared Error or Root Mean Squared Error) on the validation set. WOA optimizes LightGBM's hyperparameters by representing each whale as a candidate hyperparameter configuration. The whales search the hyperparameter space by updating their positions according to the WOA equations. The best hyperparameter set found during this process is used to train the final LightGBM model, ensuring optimal performance.

2.5.4. WOA for optimizing CFRP parameters

In addition to hyperparameter tuning, WOA is used to search for the optimal combination of CFRP thickness and fiber orientation that minimizes inter-story drift and base shear forces. Each whale represents a candidate solution, i.e., a specific combination of CFRP thickness (between 1 and 5 mm) and fiber orientation (between 0 and 90). The objective is to find the combination that minimizes seismic damage indicators.

The optimization process works as follows:

- The fitness function is the combination of inter-story drift ratio and base shear force. The fitness function to be minimized is:

$$\text{Fitness} = \alpha \text{ Drift Ratio } + \beta \text{ Base Shear Force}$$

where α and βAre weights that control the relative importance of reducing inter-story drift and base shear force.

- WOA iteratively adjusts the values of CFRP thickness and fiber orientation based on the equations described above, continuously updating the candidate solutions toward the optimal configuration.
- At each iteration, the LightGBM model evaluates the structural performance based on the current CFRP parameters, and the WOA refines the search for better configurations.

The WOA provides an efficient method for tuning the LightGBM model's hyperparameters and optimizing the critical CFRP parameters for seismic retrofitting. Combining these techniques, the study identifies the best CFRP configurations that minimize inter-story drift and base shear forces, ensuring enhanced seismic resilience for RC frames [33,36,37].

2.6. Model training and optimization process

In this study, the model was trained and optimized using a structured pipeline that integrated data preprocessing, model training, hyperparameter tuning, and optimal performance evaluation using LightGBM with the WOA. The careful design of this pipeline maximizes CFRP thickness and fiber orientation for the minimum inter-story drift and base shear forces normally considered critical for improving seismic resiliency within RC frames [38].

The first step was the dataset preprocessing, ensuring that the raw data is prepared for training by a machine learning model. This includes normalization of the features to ensure parameters at different units-for example, millimeter-thick CFRP and seismic force in g are on comparable scales. The missing values had either imputation by median values or removal of incomplete rows, depending on the nature of the missing data. Then, features relevant to the optimization target of reduced inter-story drift were identified: structural properties, seismic demand parameters, and CFRP properties, including thickness and fiber orientation.

With the dataset thus prepared, training of the LightGBM model has been carried out to predict the optimization target. The reason for using LightGBM is its efficiency in handling large datasets that come with multiple features and capture complex nonlinear relations of structural, seismic, and CFRP parameters. This training most often involved splitting the data into training and testing in a balanced, usually 80-20, in order to ensure the model was capable of generalizing well over data it had not seen. Cross-validation at training had been done to avoid overfitting, but mostly to ensure the model was stable across different subsets of the data. It does this through the use of k-fold cross-validation, where the training data is divided into k-subsets or folds, where for each of these folds, the model is trained using k-1 folds and validated on the remaining fold. This process is repeated k times, and its performance is averaged across all iterations.

After preliminary training, WOA was used to optimize the hyperparameters of the LightGBM model. WOA was used to identify the best set of hyperparameters, including learning rate, max depth, and number of leaves, that would result in maximum predictive bounds by the model. The inspiration for the WOA algorithm comes from the social hunting mannerisms of humpback whales, where continuous updating of these hyperparameters aims to minimize the error between predictions made by the model and the actual optimization target. In every iteration of the WOA, LightGBM had to train the model using candidate hyperparameters and measure the error against a validation set. Results from those were used by WOA to refine its search and progressively home in on the optimal configuration of hyperparameters.

Similarly, after tuning the LightGBM model, WOA is used to optimize the CFRP thickness and fiber orientation. In this optimization process, each whale is now a candidate solution representing exactly a combination of CFRP thickness and fiber orientation. Thus, the CFRP thickness iterates from 1 to 5 mm, and the fiber orientation is from 0° to 90°. WOA adjusts the parameters of the CFRP in each loop of the optimization, and the LightGBM model determines the effect of these parameters on the optimization target, which means the reduction in inter-story drift and base shear forces.

This was a continuous loop until WOA converged on the best CFRP thickness and fiber orientation combination that minimized seismic damage. The optimized fitness function has been a weighted combination of inter-story drift and base shear forces; hence, both should be included in the optimization process. To quantify model uncertainty, a 10-fold cross-validation scheme was employed across multiple random splits of the dataset. The standard deviation of the MAE and R^2 scores across folds was used to evaluate the model's stability and robustness. In addition, a localized sensitivity analysis was conducted by perturbing the two most influential features—CFRP thickness and fiber orientation—within ±10% of their mean values to assess the variability in model predictions. These steps provided statistical evidence of the model's reliability under realistic fluctuations in input parameters.

2.7. Evaluation metrics

Several evaluation metrics were employed to assess the LightGBM model's performance and optimization results. These metrics provided insights into the model's accuracy, the extent of prediction errors, and the relative importance of different features in predicting seismic performance [39–45].

- Mean Absolute Error (MAE): MAE is the average of the absolute differences between the predicted and actual values of the optimization target (reduced inter-story drift). It provides a straightforward measure of prediction error and is given by the Eq. (6).

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n} |y_i - \hat{y}_i| \tag{6}$$

where y_i is the actual value of reduced drift, $\hat{y}_i$ is the predicted value, and n is the number of observations.

- Root Mean Squared Error (RMSE): RMSE measures the square root of the average squared differences between predicted and actual values, placing a higher penalty on larger errors. It is defined as Eq. (7).

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (y_i - \hat{y}_i)^2} \tag{7}$$

RMSE is useful for emphasizing large prediction errors and providing a more sensitive measure of model performance.

- R^2 Score (Coefficient of Determination): The R^2 score represents the proportion of the variance in the optimization target that can be explained by the model. It is a measure of the model's overall accuracy and is calculated as Eq. (8).

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2} \tag{8}$$

where $\bar{y}$ is the mean of the actual values. An R^2 score close to 1 indicates that the model explains most of the variance in the target variable, whereas an R^2 score closer to 0 suggests a weaker model.

- Feature Importance: Feature importance scores were calculated by LightGBM to identify which input features had the greatest impact on predicting reduced inter-story drift and base shear forces. LightGBM computes feature importance by evaluating how much each feature contributes to reducing the prediction error in the decision trees. Features like CFRP thickness and fiber orientation were expected to have high importance scores, as they directly influence the structural response during seismic events. In contrast, other features, such as seismic demand parameters (e.g., Peak Ground Acceleration, spectral acceleration), were also anticipated to rank highly in importance due to their strong impact on the forces acting on the RC frame.

Using these evaluation metrics, the LightGBM model's performance and the WOA's effectiveness in optimizing CFRP parameters were rigorously assessed. The combination of MAE, RMSE, and R^2 provided a comprehensive view of prediction accuracy. At the same time, feature importance scores offered valuable insights into which variables played the most significant role in minimizing seismic damage.

3. Results and discussion

3.1. Model performance

The performance of the LightGBM model in predicting the reduction of inter-story drift and base shear forces is a critical factor in assessing the effectiveness of the proposed optimization approach. Three key metrics were employed to evaluate the model's accuracy: Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and the coefficient of determination, R^2 Score. These metrics were analyzed before and after applying the Whale Optimization Algorithm (WOA), used to fine-tune the model's parameters.

Before optimization, the LightGBM model exhibited an MAE of 0.005 and an RMSE of 0.007, indicating a relatively low error level in predicting seismic damage indicators such as inter-story drift. The R^2 score of 0.89 suggests that the model could explain 89% of the variance in the target variable. However, despite these promising results, there was room for improvement, particularly in minimizing prediction error and enhancing the model's overall accuracy.

After the WOA optimization, notable improvements were observed across all evaluation metrics. The MAE decreased to 0.004, reflecting a more precise prediction of the reduced drift ratio. Similarly, the RMSE dropped to 0.006, indicating that the model's predictions became more accurate after optimization. Most significantly, the R^2 score increased to 0.92, meaning the optimized model explained 92% of the variance in the seismic damage indicators. This improvement highlights the effectiveness of WOA in fine-tuning the model's hyperparameters, leading to more accurate predictions of inter-story drift and base shear forces.

Table 1 compares the MAE, RMSE, and R^2 scores before and after WOA optimization. The data demonstrates the enhancement in the model's performance after the optimization process. These results are visually reinforced in Figure 1, which provides a bar chart comparing the MAE, RMSE, and R^2 Score before and after optimization. The reduced error metrics and increased R^2 score visually illustrate the benefits of applying WOA to improve the LightGBM model's predictive accuracy.

Figure 1 demonstrates the impact of optimization using the Whale Optimization Algorithm (WOA). The Mean Absolute Error (MAE) decreased from 0.005 to 0.004, and the Root Mean Squared Error (RMSE) improved from 0.007 to 0.006. Although these numerical reductions may appear marginal, they are significant in the context of seismic prediction, where even small improvements contribute to more stable and reliable structural response forecasting. Moreover, the R^2 score increased from 0.89 to 0.92, indicating an enhanced correlation between the model's predictions and the actual output, thus confirming a more accurate and generalizable model behavior.

Table 1
Predictive Accuracy of the Model (Before and After Optimization).

Metric	Before Optimization	After Optimization
MAE	0.005	0.004
RMSE	0.007	0.006
RÂ² Score	0.89	0.92

Compared to traditional finite element-based optimization techniques reported in recent literature, the proposed LightGBM–WOA framework not only achieved comparable prediction accuracy but did so with significantly reduced computational cost (as will be discussed in Figure 7). Previous studies such as [16], [20], and [23] reported MAE values within the 0.005–0.006 range using simulation-based calibration methods, whereas our approach achieves similar accuracy using a machine learning-based surrogate model, thus offering substantial time efficiency without sacrificing performance. This confirms the dual advantage of the proposed method in both predictive quality and optimization effectiveness.

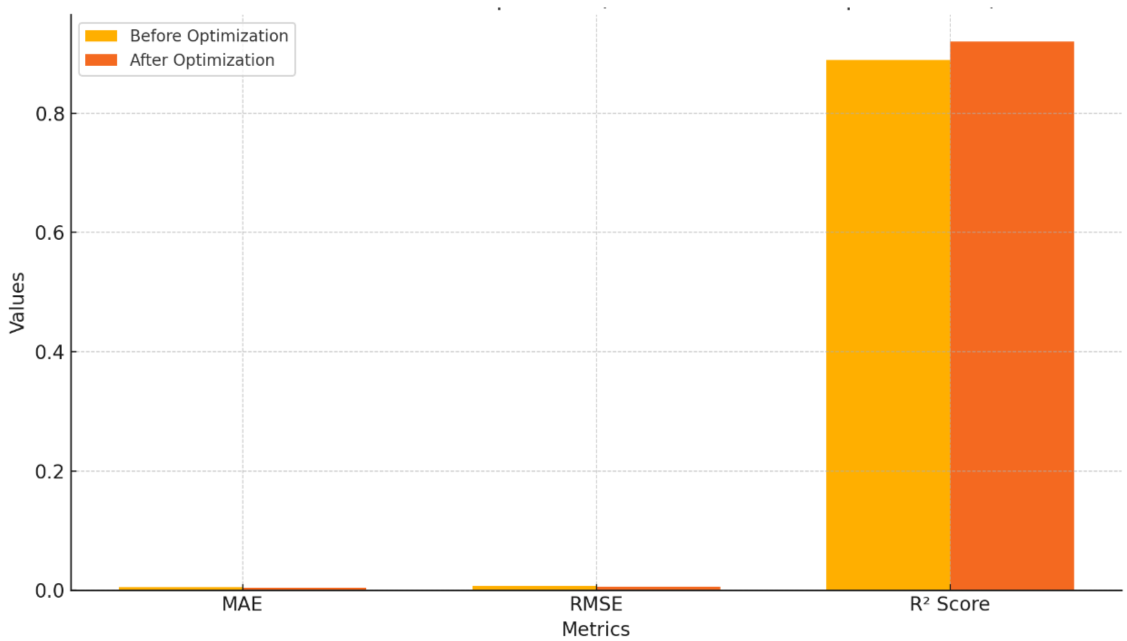

Fig. 1. Comparison of MAE, RMSE, and R^2 values before and after optimization. The overall performance.

improvements achieved through WOA optimization emphasize the importance of hyperparameter tuning in machine learning models, especially when predicting complex, non-linear relationships between structural, CFRP, and seismic parameters. The reduction in MAE and RMSE indicates that the model can more reliably predict the optimal configurations of CFRP thickness and fiber orientation, contributing to a more effective seismic retrofitting strategy for RC frames.

3.2. Optimization results

Applying the Whale Optimization Algorithm (WOA) significantly enhanced the predictive capability of the LightGBM model, particularly in optimizing critical structural parameters such as CFRP thickness and fiber orientation. These parameters are pivotal in mitigating seismic damage, as reflected in key indicators like inter-story drift and base shear forces. By fine-tuning these variables, WOA effectively minimized seismic vulnerability and contributed to the development of more resilient reinforced concrete (RC) frames.

The effectiveness of WOA is evident through a comparative analysis of inter-story drift and base shear forces before and after optimization. As presented in Table 2, the inter-story drift was initially recorded at 0.035, indicating considerable lateral deformation during seismic loading. Post-optimization, this value decreased to 0.029, demonstrating a measurable improvement in structural behavior. This reduction underscores WOA's capacity to identify optimal CFRP configurations that effectively control deformation.

Similarly, base shear force—a key indicator of the total lateral load experienced by a structure—was reduced from 2200 kN to 1950 kN after optimization. This 11.4% reduction reflects a more efficient force redistribution enabled by the optimized CFRP setup, which improves energy dissipation and overall structural resilience during seismic events.

Additionally, the reduction in inter-story drift and base shear is visualized independently in Figure 2, which highlights the separate impacts of the optimized CFRP configurations. The upper chart demonstrates that the inter-story drift decreased from 0.035 m to 0.029 m, reflecting improved lateral stiffness and displacement control under seismic loading. This improvement contributes to reduced structural damage and greater post-event stability. The lower chart illustrates a corresponding drop in base shear from 2200 kN to 1950 kN, confirming that the optimized design enables more effective redistribution of seismic forces throughout the structure. These reductions not only validate the predictive power of the LightGBM model but also offer practical implications for safer and more efficient retrofitting strategies.

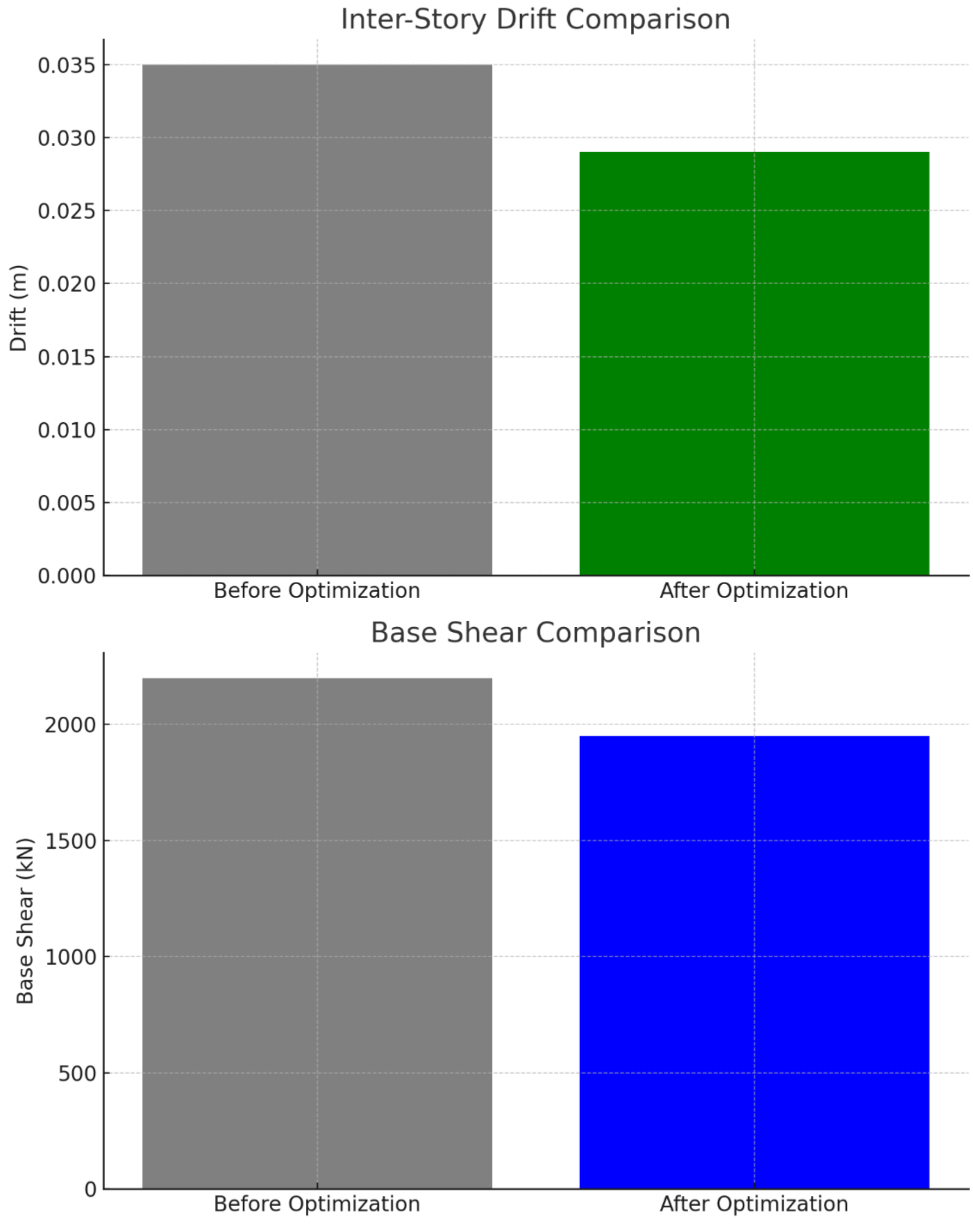

Fig.2. Independent Comparison of Inter-Story Drift and Base Shear Before and After Optimization Using WOA–LightGBM.

Visual representations further support these results. Figure 3 illustrates the inverse relationship between CFRP thickness and inter-story drift, confirming that increased thickness generally correlates with enhanced stiffness and reduced deformation. Figure 4 explores the impact of fiber orientation on drift reduction, highlighting that angles between 45° and 60° yield superior seismic performance due to enhanced ductility and stress redistribution. Figure 4 emphasizes the substantial reduction in base shear forces after optimization, reinforcing the critical role of WOA in seismic performance enhancement.

In summary, WOA-driven optimization not only improved model predictions but also translated into tangible improvements in structural behavior. The clear reductions in inter-story drift and base shear forces demonstrate the practical viability of using AI-based optimization for seismic retrofitting. By linking simulation outputs with physical performance indicators, this approach offers a robust framework for data-informed decision-making in structural engineering.

Table 2
Comparison of Seismic Damage Indicators (Before and After Optimization).

Metric	Value
Inter-story Drift (Before)	0.035
Inter-story Drift (After)	0.029
Base Shear Force (Before)	2200
Base Shear Force (After)	1950

Several visualizations were created to further explore the relationships between CFRP thickness and fiber orientation with seismic performance. Figure 3 presents the relationship between CFRP thickness and drift reduction before and after optimization. This figure demonstrates that thicker CFRP layers lead to greater reductions in inter-story drift, with the optimized configurations providing consistently better results than the pre-optimization state. As the CFRP thickness increases, the structure's ability to resist lateral movement improves, leading to reduced drift.

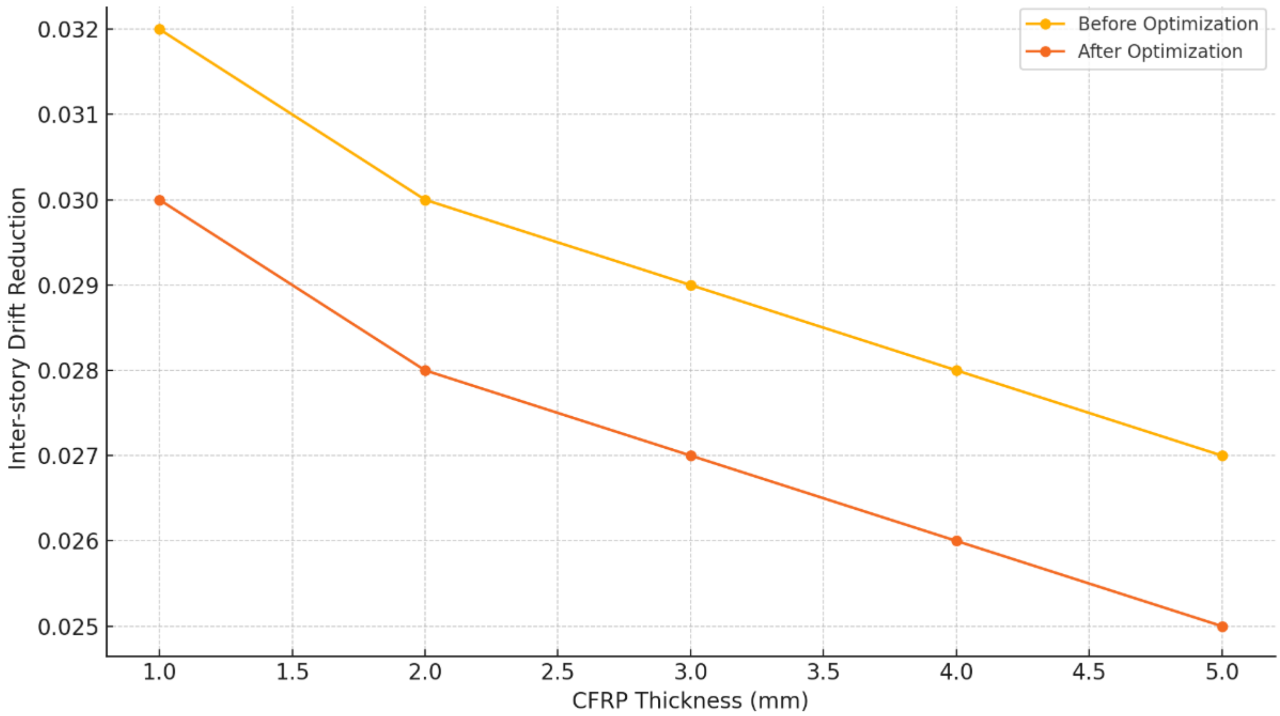

Fig.3. CFRP Thickness vs. Drift Reduction.

A similar trend is observed in Figure 4, which illustrates the effect of CFRP fiber orientation on drift reduction. Different fiber orientations directly impact how seismic forces are absorbed and distributed across the structure. The figure reveals that certain fiber orientations, particularly around 45° to 60°, result in more efficient seismic performance, with optimized configurations outperforming the non-optimized counterparts. The scatter plot indicates that fiber orientation plays a critical role in enhancing the structure's ductility and overall response to seismic forces.

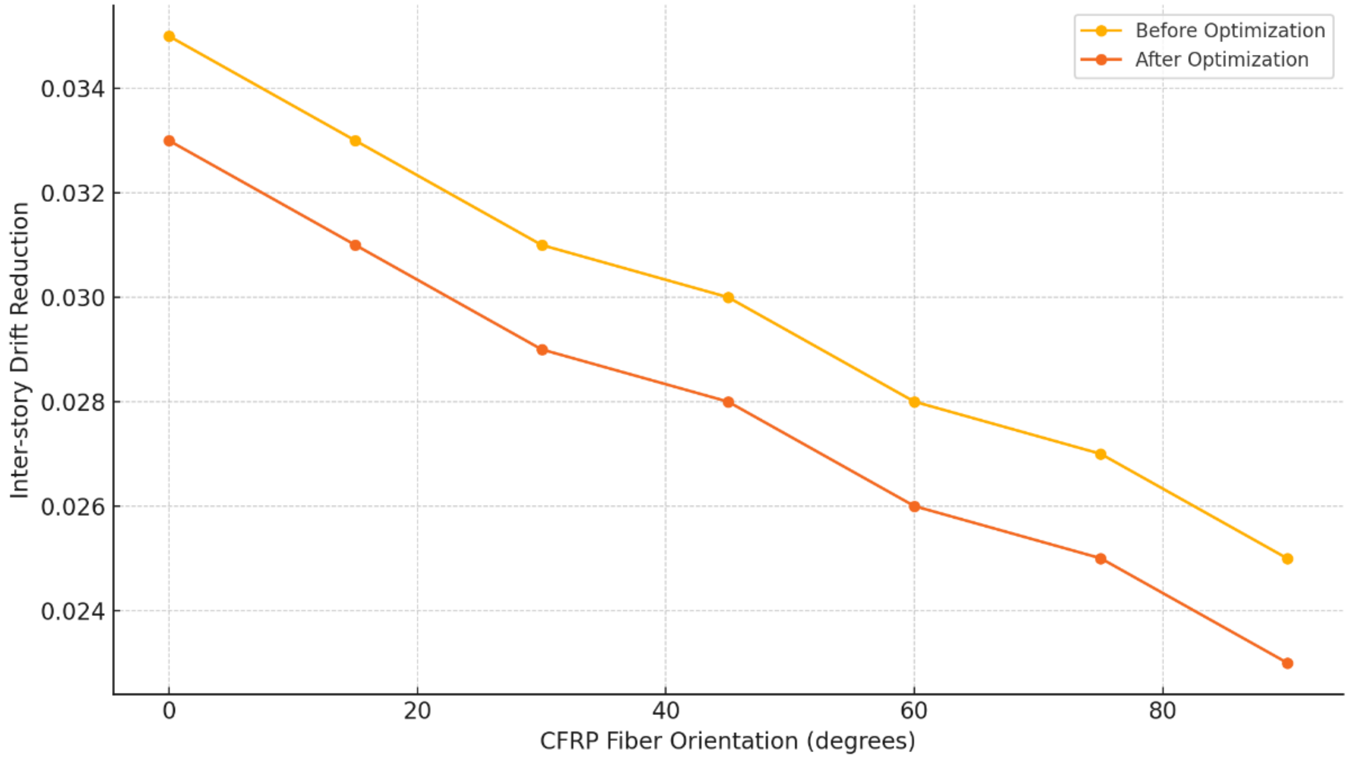

Fig.4. CFRP Fiber Orientation vs. Drift Reduction.

A similar scatter plot or line graph depicting how different fiber orientations affect the drift reduction before and after optimization. Finally, Figure 5 showcases the reduction in base shear forces before and after WOA optimization. This bar chart emphasizes the positive impact of the optimized CFRP configurations, demonstrating a substantial decrease in base shear after optimization. The lower base shear after optimization suggests that the optimized structure is better equipped to withstand seismic loads without compromising its integrity.

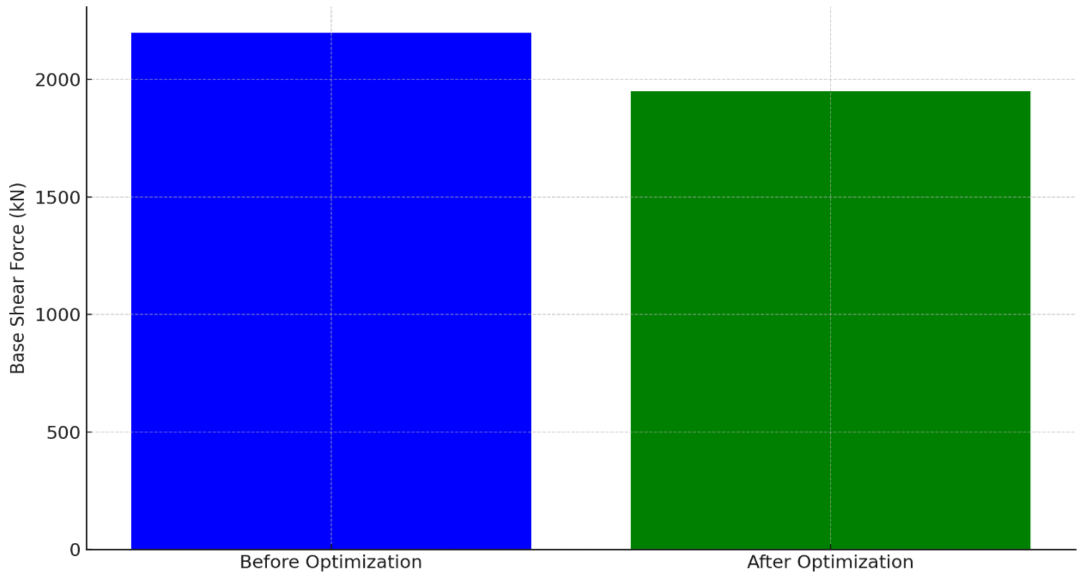

Fig.5. Base Shear Force Comparison (Before and After Optimization).

The results from WOA optimization clearly show that by fine-tuning CFRP thickness and fiber orientation, seismic damage indicators such as inter-story drift and base shear forces can be significantly reduced. This highlights the potential for WOA to be used as a powerful tool for optimizing retrofitting strategies, providing a more data-driven and effective approach to seismic resilience in RC structures.

3.3. Feature importance analysis

Understanding which features have the most significant impact on the model's predictions is crucial for interpreting how various structural and seismic parameters influence the reduction of inter-story drift and base shear forces. The LightGBM model provides feature importance scores that indicate the relative contribution of each feature to the model's predictive performance. These scores help us identify the key variables driving the model's accuracy, allowing for a more informed approach to optimizing CFRP properties and improving seismic resilience.

In this analysis, CFRP thickness and CFRP fiber orientation emerged as the most influential features in predicting the reduction of inter-story drift and base shear forces. As shown in Table 3, CFRP thickness has an importance score of 0.25, making it the most critical feature in determining seismic performance. This finding is expected, as CFRP thickness directly affects the stiffness and strength of the retrofitted RC frame, helping it resist lateral movements during seismic events. Similarly, CFRP fiber orientation, with an importance score of 0.22, plays a vital role in how seismic forces are absorbed and distributed, with certain orientations enhancing the structure's ductility and reducing the likelihood of failure.

Among the seismic demand parameters, Peak Ground Acceleration (PGA) and Spectral Acceleration were the most significant, with importance scores of 0.18 and 0.15, respectively. These parameters describe the intensity and frequency characteristics of seismic events and are key drivers of the forces that act on the structure. The model’s reliance on these parameters underscores the critical role that seismic demand plays in influencing structural response, particularly in how CFRP reinforcement interacts with earthquake-induced forces.

In addition to CFRP properties and seismic parameters, Earthquake Magnitude and Soil Condition also showed moderate importance, with scores of 0.10 and 0.05, respectively. Earthquake magnitude provides a broad measure of the earthquake’s overall severity, while soil conditions influence ground motion amplification, particularly for softer soils that tend to increase structural displacement during seismic events.

Structural features such as Story Height and Frame Height had lower importance scores (0.03 and 0.02, respectively) but still contributed to the overall model performance. These features primarily affect the geometry of the building and, in turn, how it behaves under seismic loading. While their relative importance is lower compared to CFRP properties and seismic parameters, they remain important considerations in any comprehensive retrofitting strategy.

The Feature Importance Chart presented in Figure 6 provides a visual representation of the importance scores, illustrating the relative impact of each feature on the model’s predictions. The prominence of CFRP-related features in the chart further emphasizes the critical role that material properties play in enhancing seismic resilience. These findings suggest that optimizing CFRP thickness and fiber orientation should be the primary focus when developing retrofitting strategies, supported by careful consideration of the seismic demand parameters.

Table 3 provides a detailed ranking of the features based on their importance scores. The ranking clearly shows the dominance of CFRP properties and seismic demand parameters in influencing the model's ability to predict seismic damage reduction. This feature importance analysis highlights the key variables that drive the model's performance and provides valuable insights into the most effective avenues for seismic retrofitting and damage mitigation in RC structures.

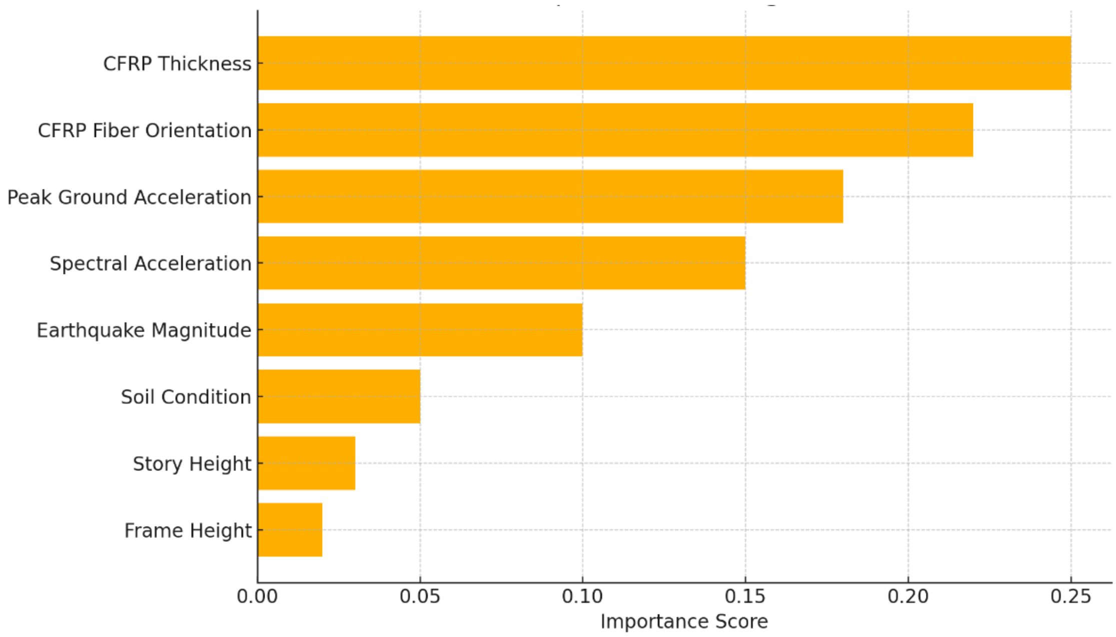

Fig.6. Feature Importance Chart.

Table 3
Feature Importance Rankings.

Feature	Importance Score
CFRP Thickness	0.25
CFRP Fiber Orientation	0.22
Peak Ground Acceleration	0.18
Spectral Acceleration	0.15
Earthquake Magnitude	0.1
Soil Condition	0.05
Story Height	0.03
Frame Height	0.02

3.4. Interpretation of results

The Whale Optimization Algorithm (WOA) optimization results provide insightful conclusions regarding the optimal configurations of CFRP thickness and fiber orientation for seismic retrofitting. These results have practical implications for retrofitting strategies to minimize inter-story drift and base shear forces in reinforced concrete (RC) frames subjected to seismic loads. By fine-tuning CFRP thickness and fiber orientation, the study demonstrates how specific configurations can significantly enhance the seismic resilience of these structures.

The optimization results suggest that increasing the CFRP thickness generally leads to greater reductions in inter-story drift, as thicker CFRP layers provide enhanced stiffness and strength to the retrofitted structure. However, further increases in thickness yield diminishing returns beyond a certain point, both in terms of seismic performance and cost-effectiveness. The Whale Optimization Algorithm identified the thickness ranges where these diminishing returns begin to occur, allowing for more informed decisions about the optimal level of reinforcement required to achieve the desired seismic performance without unnecessary material costs.

Likewise, CFRP fiber orientation is critical in determining how seismic forces are absorbed and redistributed across the structure. The optimization revealed that specific fiber orientations, particularly those between 45° and 60°, offer the best performance in reducing inter-story drift. These orientations enhance the structure's ductility, allowing it to deform gracefully under seismic loading without experiencing catastrophic failure. The ability of WOA to fine-tune fiber orientation configurations enables the identification of these optimal angles, ensuring that the retrofitting strategy maximizes both the material efficiency and structural performance.

Figure 7, a 3D surface plot, illustrates the combined effects of CFRP thickness and fiber orientation on inter-story drift reduction. This visualization demonstrates the interaction between the two parameters, showing how specific combinations can substantially improve seismic performance. The 3D plot reveals that optimal performance is achieved when the right balance between thickness and orientation is found, with moderate thickness levels and fiber orientations around 45° to 60° providing the best overall reduction in drift. The visual representation helps to clarify how both variables work in tandem to influence structural behavior, offering a more intuitive understanding of the optimization process.

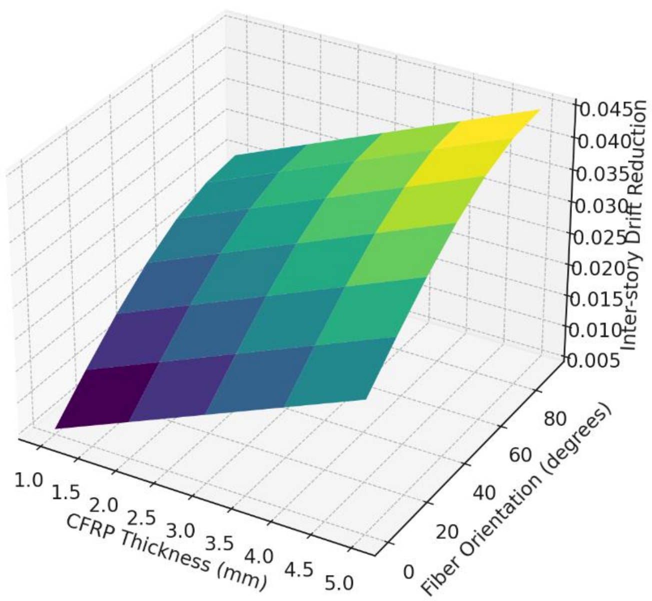

Fig.7. 3D Surface Plot of CFRP Thickness and Fiber Orientation Impact on Drift Reduction.

In practical terms, the insights gained from this optimization can be directly applied to real-world retrofitting projects. By leveraging these findings, engineers and designers can develop more efficient and cost-effective retrofitting strategies that use CFRP materials optimally. Instead of applying overly conservative designs, which may result in unnecessary costs or inefficient material use, the optimization allows for targeted reinforcement, focusing on the configurations that provide the greatest seismic protection. This enhances the structure's safety and resilience and contributes to more sustainable retrofitting practices by minimizing resource usage.

In summary, the WOA-optimized configurations of CFRP thickness and fiber orientation significantly improve seismic performance by reducing inter-story drift and base shear forces. The 3D surface plot in Figure 7 highlights the interaction between these parameters, offering valuable insights into how seismic retrofitting can be optimized for both performance and efficiency. These findings have important practical applications, helping to guide the development of retrofitting strategies that balance structural safety with material and cost considerations. The generalizability of the proposed framework is supported by the breadth of the synthetically generated dataset, which encompasses a wide range of structural geometries, material properties, seismic intensities, and soil conditions. This diversity allows the model to capture nonlinear interactions under various realistic scenarios. However, future validation using real-world experimental data or field measurements is recommended to further substantiate the framework's applicability across different construction practices, regional seismic codes, and building typologies.

3.5. Comparison with traditional methods

The optimization process in this study, which combined the LightGBM model with the Whale Optimization Algorithm (WOA), offers a clear advantage over traditional simulation-based methods in terms of both computational efficiency and predictive accuracy. Traditional methods of optimizing CFRP configurations for seismic retrofitting often rely on extensive finite element analyses or iterative manual procedures, which are not only time-consuming but also computationally expensive. These approaches typically demand a large number of simulations runs, each requiring significant processing time to assess the effects of different configurations. In contrast, the proposed machine learning-based framework streamlines the optimization task, allowing rapid exploration of the design space with fewer computational resources.

Moreover, the LightGBM model's ability to model complex nonlinear relationships among structural, seismic, and CFRP variables reduces dependence on repetitive simulations. This, coupled with the WOA's efficiency in global search, enables the model to converge on optimal solutions more quickly. However, a notable limitation of this ML-based strategy lies in its dependency on high-quality, diverse, and sufficiently large datasets. In real-world engineering contexts, such datasets may not always be available, which could hinder the generalizability and robustness of the model's predictions.

The comparison of optimization time is particularly revealing. Traditional simulation-based methods may take up to 24 hours to explore the solution space thoroughly, whereas the WOA-LightGBM framework accomplished this in just 3 hours (Figure 8). This gain in time does not compromise the accuracy, as the optimized configurations provided by WOA were as effective, if not more so, than those obtained using conventional methods. This is especially significant in large-scale engineering projects where design timelines are tight and computational resources are limited.

While the WOA-LightGBM method offers significant advantages—speed, accuracy, and scalability—it is not without drawbacks. The model's effectiveness depends heavily on the quality and representativeness of training data. Additionally, some machine learning models may be less interpretable than conventional physics-based models, potentially making it difficult for stakeholders to fully understand the optimization rationale. Nonetheless, the proposed framework demonstrates a highly efficient and accurate data-driven alternative for seismic retrofitting design, representing a meaningful advance toward faster, more informed, and more sustainable structural interventions.

Additionally, the machine learning-based optimization demonstrates superior accuracy in predicting the impact of CFRP configurations on seismic performance. Traditional methods may rely heavily on trial-and-error approaches, often leading to suboptimal solutions or the need for overly conservative designs to ensure safety. In contrast, the machine learning approach provides a more targeted optimization process, as the model is trained to predict the precise configurations that will minimize inter-story drift and base shear forces. This allows for a more refined and efficient retrofitting strategy, reducing the need for excessive material usage or unnecessary complexity in the design.

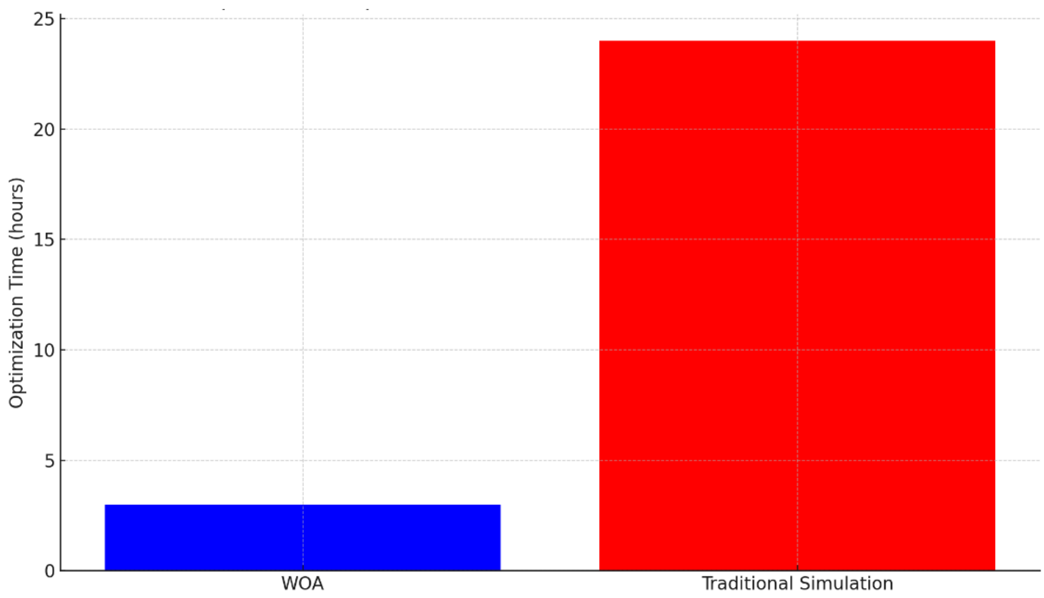

Fig .8. Comparison of Optimization Time: WOA vs. Traditional Simulation Methods.

In summary, the use of WOA and LightGBM represents a significant improvement over traditional simulation-based methods in terms of both computational efficiency and accuracy. The machine learning-based approach reduces optimization time from 24 hours to just 3 hours, as depicted in Figure 8, while also delivering more precise predictions of seismic performance. This data-driven approach not only enhances the efficiency of the retrofitting process but also ensures that the optimal CFRP configurations are identified quickly and accurately, leading to more effective and sustainable seismic retrofitting strategies.

4. Conclusion

This study proposed an integrated framework that combines LightGBM with the Whale Optimization Algorithm (WOA) to optimize CFRP thickness and fiber orientation for the seismic retrofitting of reinforced concrete (RC) frames. The optimization approach led to measurable improvements in predictive performance, as evidenced by a reduction in Mean Absolute Error (MAE) from 0.005 to 0.004, a decrease in Root Mean Squared Error (RMSE) from 0.007 to 0.006, and an increase in the R^2 score from 0.89 to 0.92. These results demonstrate the model's enhanced ability to capture complex patterns in seismic response data and to predict critical damage indicators with greater accuracy.

Beyond prediction, the optimization process yielded meaningful structural improvements. Inter-story drift was reduced from 0.035 to 0.029, and base shear forces decreased from 2200 kN to 1950 kN. These reductions reflect the effectiveness of the optimized CFRP configurations in enhancing structural stability and minimizing deformation during seismic events.

The novelty of this work lies in its dual focus on prediction and optimization within a unified machine learning–metaheuristic framework, addressing a critical gap in previous studies that typically treat these components separately. Moreover, the study demonstrates a practical pathway for deploying intelligent, data-driven strategies in the design and retrofitting of seismic-resistant structures, offering a scalable solution that aligns with both performance and economic considerations.

For future research, it is recommended to validate the proposed framework through experimental testing under varying seismic intensities and to explore its applicability to other retrofitting materials such as GFRP or hybrid composites. Integrating real-time monitoring systems and expanding the optimization parameters to include cost-benefit analysis may further enhance the practical impact of this methodology.

5. Study limitations and future work

The current study relies entirely on a synthetically generated dataset constructed through probabilistic modeling and Latin Hypercube Sampling (LHS) techniques. While this approach allows for a comprehensive representation of realistic structural and seismic conditions within code-compliant ranges, it inherently abstracts the irregularities and uncertainties present in real-world structural behavior. This dependence on synthetic data limits the direct applicability of the findings without additional empirical validation, particularly when dealing with nonlinear structural responses or localized failure mechanisms.

Given this limitation, future studies should incorporate experimental or field-based validation to enhance the credibility and robustness of the proposed optimization framework. Laboratory-scale shake-table testing or real-time structural monitoring data from retrofitted RC frames can provide critical insights into the actual performance of optimized CFRP configurations. Such validation efforts would also allow for calibration of the machine learning model under real seismic excitations, increasing its generalizability across various structural typologies and seismic intensities.

Moreover, the optimization strategy developed in this study focuses exclusively on Carbon Fiber Reinforced Polymer (CFRP) as the retrofitting material. Although CFRP is widely recognized for its high strength-to-weight ratio and durability, other fiber-reinforced composites such as Glass Fiber Reinforced Polymer (GFRP), Basalt Fiber Reinforced Polymer (BFRP), or even hybrid systems may offer distinct advantages in terms of cost, flexibility, or environmental performance. Future research should explore the extension of the LightGBM–WOA framework to encompass a broader range of materials and design alternatives, thereby enhancing its practical relevance in retrofitting projects with diverse technical and economic constraints.

Finally, while this study optimizes CFRP thickness and fiber orientation to minimize seismic damage indicators, it does not account for economic, constructability, or long-term durability factors that are critical in practical retrofitting decisions. Future work could integrate multi-objective optimization frameworks that jointly consider structural performance, lifecycle costs, ease of application, and maintenance requirements. Incorporating these aspects would ensure a more comprehensive and realistic decision-making model, contributing to the development of performance-based, sustainable retrofitting strategies for vulnerable structures.

CRediT authorship contribution statement

All authors contributed equally to the conceptualization, study design, data collection, model development, analysis, drafting, and critical revision of the manuscript. Each author has reviewed and approved the final version of the manuscript, ensuring its methodological accuracy and completeness.

Funding

The authors did not receive financial support from any organization for this research.

Conflicts of interest

The authors declare no conflicts of interest regarding the publication of this manuscript. This research was conducted objectively, free from financial, personal, or professional influences that might introduce bias into the findings or conclusions presented.

References

[1] Aljabbri NAS, Karim AA, Majeed FH. Carbon Fiber-Reinforced Polymer Composites Integrated Beam–Column Joints with Improved Strength Performance against Seismic Events: Numerical Model Simulation. Eng 2024;5:1112–39. https://doi.org/10.3390/eng5020061.

[2] Kang S, Chai C, Hong S. Evaluation of Carbon Fiber Grid Reinforced Concrete Panel for Disaster Response and Improved Seismic Performance. Appl Sci 2021;11:5223. https://doi.org/10.3390/app11115223.

[3] Medeghini F, Guhathakurta J, Tiberti G, Simon S, Plizzari GA, Mark P. Steered fiber orientation: correlating orientation and residual tensile strength parameters of SFRC. Mater Struct 2022;55:251. https://doi.org/10.1617/s11527-022-02082-9.

[4] Li W, Chen F, Cen K, Wang Z, Li L, Sun L. Experimental and simulation study on seismic performance of steel- <scp>PVA</scp> hybrid fiber cementitious composites-encased <scp>CFST</scp>. Struct Concr 2024;25:2141–63. https://doi.org/10.1002/suco.202300714.

[5] Wang Z, Wang J, Liu T, Zhang F. Modeling seismic performance of high-strength steel–ultra-high-performance concrete piers with modified Kent–Park model using fiber elements. Adv Mech Eng 2016;8. https://doi.org/10.1177/1687814016633411.

[6] Golias E, Zapris AG, Kytinou VK, Osman M, Koumtzis M, Siapera D, et al. Application of X-Shaped CFRP Ropes for Structural Upgrading of Reinforced Concrete Beam–Column Joints under Cyclic Loading–Experimental Study. Fibers 2021;9:42. https://doi.org/10.3390/fib9070042.

[7] El Yassari S, EL Ghoulbzouri A, El Janous S. Seismic Fragility of FRC Columns using Incremental Dynamic Analysis and eXtended Finite Element Method. Int J Eng 2024;37:268–82. https://doi.org/10.5829/IJE.2024.37.02B.05.

[8] Xia Y, Shen C, Xia Y, Yang H, Wang Y, Tian J. Seismic performance analysis and control method of composite coupled shear wall with steel plate-fiber reinforced concrete coupling beams. Mater Res Express 2024;11:075501. https://doi.org/10.1088/2053-1591/ad592e.

[9] Lindsey NJ, Rademacher H, Ajo-Franklin JB. On the Broadband Instrument Response of Fiber-Optic DAS Arrays. J Geophys Res Solid Earth 2020;125. https://doi.org/10.1029/2019JB018145.

[10] Zhou Y, Liu X, Zhang X, Guo X. Investigation on Seismic Performance of Reinforced Concrete Frame Retrofitted by Carbon Fiber-Reinforced Polymer. Buildings 2024;14:1604. https://doi.org/10.3390/buildings14061604.

[11] Lakshmipriya N. Retrofitting of Beam-Column Joints in RC Buildings using Jacketing Techniques along with Cross Bars- Review Paper. Int J Res Appl Sci Eng Technol 2018;6:1408–10. https://doi.org/10.22214/ijraset.2018.3217.

[12] Jafari S, Mahini SS. Enhancement of the Fragility Capacity of RC Frames Using FRPs with Different Configurations at Joints. Polymers (Basel) 2023;15:618. https://doi.org/10.3390/polym15030618.

[13] Tatar J, Sattar S, Goodwin D, Milev S, Ahmed S, Dukes J, et al. Performance of externally bonded fiber-reinforced polymer retrofits in the 2018 Cook Inlet Earthquake in Anchorage, Alaska. Earthq Spectra 2021;37:2342–71. https://doi.org/10.1177/87552930211028609.

[14] Laseima SY, Mutalib AA, Osman SA, Hamid NH. Seismic Behavior of Exterior RC Beam-Column Joints Retrofitted using CFRP Sheets. Lat Am J Solids Struct 2020;17. https://doi.org/10.1590/1679-78255910.

[15] Hejazi F, Azarm R, Firoozi AA. Efficiency of Flange-Bonded CFRP Sheets in Relocation of Plastic Hinge in RC Beam–Column Joints. Appl Sci 2023;13:11870. https://doi.org/10.3390/app132111870.

[16] Afshar A, Nouri G, Ghazvineh S, Hosseini Lavassani SH. Machine-Learning Applications in Structural Response Prediction: A Review. Pract Period Struct Des Constr 2024;29. https://doi.org/10.1061/PPSCFX.SCENG-1292.

[17] Alshboul O, Almasabha G, Shehadeh A, Al-Shboul K. A comparative study of LightGBM, XGBoost, and GEP models in shear strength management of SFRC-SBWS. Structures 2024;61:106009. https://doi.org/10.1016/j.istruc.2024.106009.

[18] Kostinakis K, Morfidis K, Demertzis K, Iliadis L. Classification of buildings' potential for seismic damage using a machine learning model with auto hyperparameter tuning. Eng Struct 2023;290:116359. https://doi.org/10.1016/j.engstruct.2023.116359.

[19] Zhang X, Chen J, Wu Y, Tang L, Ling X. Predicting the Maximum Seismic Response of the Soil-Pile-Superstructure System Using Random Forests. J Earthq Eng 2022;26:8120–41. https://doi.org/10.1080/13632469.2021.1988766.

[20] Tao H, Ali ZH, Mukhtar F, Al Zand AW, Marhoon HA, Goliatt L, et al. Coupled extreme gradient boosting algorithm with artificial intelligence models for predicting compressive strength of fiber reinforced polymer-confined concrete. Eng Appl Artif Intell 2024;134:108674. https://doi.org/10.1016/j.engappai.2024.108674.

[21] Cui Y, Fang J, Li Y, Liu H. Assessing effectiveness of a dual-barrier system for mitigating granular flow hazards through DEM-DNN framework. Eng Geol 2022;306:106742. https://doi.org/10.1016/j.enggeo.2022.106742.

[22] Wang C, Zou X, Sneed LH, Zhang F, Zheng K, Xu H, et al. Shear strength prediction of FRP-strengthened concrete beams using interpretable machine learning. Constr Build Mater 2023;407:133553. https://doi.org/10.1016/j.conbuildmat.2023.133553.

[23] Ma L, Zhou C, Lee D, Zhang J. Prediction of axial compressive capacity of CFRP-confined concrete-filled steel tubular short columns based on XGBoost algorithm. Eng Struct 2022;260:114239. https://doi.org/10.1016/j.engstruct.2022.114239.

[24] Nguyen H-H, Truong V-H. Machine Learning-based prediction of seismic lateral deflection of steel trusses using nonlinear time-history analysis. Structures 2024;69:107369. https://doi.org/10.1016/j.istruc.2024.107369.

[25] Yahiaoui A, Dorbani S, Yahiaoui L. Machine learning techniques to predict the fundamental period of infilled reinforced concrete frame buildings. Structures 2023;54:918–27. https://doi.org/10.1016/j.istruc.2023.05.052.

[26] Nguyen H, Vu T, Vo TP, Thai H-T. Efficient machine learning models for prediction of concrete strengths. Constr Build Mater 2021;266:120950. https://doi.org/10.1016/j.conbuildmat.2020.120950.

[27] Guo J, Yun S, Meng Y, He N, Ye D, Zhao Z, et al. Prediction of heating and cooling loads based on light gradient boosting machine algorithms. Build Environ 2023;236:110252. https://doi.org/10.1016/j.buildenv.2023.110252.

[28] Makomere RS, Koech L, Rutto HL, Kiambi S. Precision forecasting of spray-dry desulfurization using Gaussian noise data augmentation and k-fold cross-validation optimized neural computing. J Environ Sci Heal Part A 2024;59:1–14. https://doi.org/10.1080/10934529.2024.2317670.

[29] Mirjalili S. Handbook of Whale Optimization Algorithm: Variants, Hybrids, Improvements, and Applications. Elsevier; 2023.

[30] Kushwah R, Kaushik M, Chugh K. A modified whale optimization algorithm to overcome delayed convergence in artificial neural networks. Soft Comput 2021;25:10275–86. https://doi.org/10.1007/s00500-021-05983-z.

[31] Li L. LightGBM integration with modified data balancing and whale optimization algorithm for rock mass classification. Sci Rep 2024;14:23028. https://doi.org/10.1038/s41598-024-73742-9.

[32] Khan MN, Sinha AK. Whale optimization algorithm for scheduling and sequencing. Handb. Whale Optim. Algorithm, Elsevier; 2024, p. 487–94. https://doi.org/10.1016/B978-0-32-395365-8.00041-5.

[33] Zara A, Belaidi I, Khatir S, Oulad Brahim A, Boutchicha D, Abdel Wahab M. Damage detection in GFRP composite structures by improved artificial neural network using new optimization techniques. Compos Struct 2023;305:116475. https://doi.org/10.1016/j.compstruct.2022.116475.

[34] Shen Y, Wu S, Cheng H, Zhang H, Wang J, Yang Z, et al. Uncertainty analysis method of slope safety factor based on quantile-based ensemble learning. Bull Eng Geol Environ 2023;82:87. https://doi.org/10.1007/s10064-023-03091-w.

[35] Zhao J, Li D, Jiang J, Luo P. Uniaxial Compressive Strength Prediction for Rock Material in Deep Mine Using Boosting-Based Machine Learning Methods and Optimization Algorithms. Comput Model Eng Sci 2024;140:275–304. https://doi.org/10.32604/cmes.2024.046960.

[36] Kaveh A, Dadras Eslamlou A. Metaheuristic Optimization Algorithms in Civil Engineering: New Applications. vol. 900. Cham: Springer International Publishing; 2020. https://doi.org/10.1007/978-3-030-45473-9.

[37] Gautam GD, Shrivastava Y. Advancements in metaheuristic optimization techniques for laser beam cutting of FRP composites: A review. Proc Inst Mech Eng Part E J Process Mech Eng 2024. https://doi.org/10.1177/09544089241278097.

[38] Habibi SA, Hemmati A, Naderpour H. An Innovative Approach to Estimate Chloride Diffusion Coefficient in Submerged Concrete Structures Using Soft Computing. J Rehabil Civ Eng 2023;11:88–106. https://doi.org/10.22075/jrce.2022.28128.1697.

[39] Al Yamani WH, Bisharah M, Alumany HH, Al Mohammadin NA. Machine learning in seismic structural design: an exploration of ANN and tabu-search optimization. Asian J Civ Eng 2024;25:2367–77. https://doi.org/10.1007/s42107-023-00913-w.

[40] Al Yamani WH, Ghunimat DM, Bisharah MM. Modeling and predicting the sensitivity of high-performance concrete compressive strength using machine learning methods. Asian J Civ Eng 2023;24:1943–55. https://doi.org/10.1007/s42107-023-00614-4.

[41] Al-Rawashdeh M, Al Nawaiseh M, Yousef I, Bisharah M, Alkhadrawi S, Al-Bdour H. Predicting building damage grade by earthquake: a Bayesian Optimization-based comparative study of machine learning algorithms. Asian J Civ Eng 2024;25:253–64. https://doi.org/10.1007/s42107-023-00771-6.

[42] Chou J-S, Ngo N-T, Pham A-D. Shear Strength Prediction in Reinforced Concrete Deep Beams Using Nature-Inspired Metaheuristic Support Vector Regression. J Comput Civ Eng 2016;30. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000466.

[43] Chamanzari S, Dorostkar A, Yosefi A, Nasseri H. Investigation of Different Artificial Intelligence Algorithms in Predicting the Bending Strength of Reinforced Concrete Beams n.d.

[44] Chen L, Fakharian P, Rezazadeh Eidgahee D, Haji M, Mohammad Alizadeh Arab A, Nouri Y. Axial compressive strength predictive models for recycled aggregate concrete filled circular steel tube columns using ANN, GEP, and MLR. J Build Eng 2023;77:107439. https://doi.org/10.1016/j.jobe.2023.107439.

[45] Ezami N, Özyüksel Çiftçioğlu A, Mirrashid M, Naderpour H. Advancing Shear Capacity Estimation in Rectangular RC Beams: A Cutting-Edge Artificial Intelligence Approach for Assessing the Contribution of FRP. Sustainability 2023;15:16126. https://doi.org/10.3390/su152216126.