RESEARCH ARTICLE

Optimizing Concrete Compressive Strength Prediction Using ANFIS Models Enhanced by Butterfly, Beluga Whale, and Golden Eagle Metaheuristics

Rufaizal Che Mamat[a],* ⬤, Azuin Ramli[b] ⬤, Muhammad Lukman Kirunjisman[c]

[a]Centre of Green Technology for Sustainable Cities, Department of Civil Engineering, Politeknik Ungku Omar, Jalan Raja Musa Mahadi, 31400 Ipoh, Perak, Malaysia

[b]Centre of Research and Innovation Excellence, Politeknik Ungku Omar, Jalan Raja Musa Mahadi, 31400 Ipoh, Perak, Malaysia

[c]Centre of Graduate Studies, Department of Civil Engineering, Politeknik Ungku Omar, Jalan Raja Musa Mahadi, 31400 Ipoh, Perak, Malaysia

*Corresponding Author: Rufaizal Che Mamat (rufaizal@puo.edu.my)

Publisher's Note: *Popular Scientist* stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

ABSTRACT

The precise prediction of concrete compressive strength, both mean and characteristic, is fundamental to structural integrity. This study presents a comparative analysis of ANFIS hybrid models optimized by Butterfly (BOA), Beluga Whale (BWO), and Golden Eagle (GEO) metaheuristics. Utilizing a dataset of 400 samples (split 70% for training and 30% for testing), models were evaluated in units of MPa. Performance was rigorously evaluated using RMSE, MAE, and R^2 metrics. While all hybrid models significantly outperformed the standalone ANFIS, a critical distinction emerged between training and testing performance. Specifically, the ANFIS-GEO model achieved a testing RMSE of 2.35 and an R^2 of 0.93, representing a 9% reduction in error compared to the baseline hybrid ANFIS and a superior stability profile over the overfitted ANFIS-BWO configuration. These findings identify the ANFIS-GEO hybrid as a highly reliable and robust tool for practical concrete engineering applications.

Keywords: Concrete Compressive Strength, Hybrid Machine Learning, ANFIS, Metaheuristic Optimization, Golden Eagle Optimizer

INTRODUCTION

Concrete stands as the most widely utilized construction material globally, with its compressive strength being the most critical mechanical property governing the load-bearing capacity, durability, and long-term performance of structural elements [1]. The traditional process of concrete mix design relies heavily on empirical methods and standardized guidelines, which are often laborious, time-consuming, and may not adequately capture the complex, non-linear interactions among constituents [2-4]. These limitations become particularly pronounced

when dealing with high-performance concrete or specialized mixtures, where achieving optimal performance requires a more nuanced approach. Consequently, there is a pressing need within the civil engineering domain for advanced, data-driven methodologies that can provide rapid and reliable predictions of concrete compressive strength, thereby facilitating more efficient mix design, reducing material waste, and enhancing structural safety [5,6].

In response to this challenge, the proliferation of data and computational power has catalyzed a paradigm shift towards machine learning (ML) and artificial intelligence (AI) for modeling complex material behaviors [7,8]. Among these advanced techniques, the Adaptive Neuro-Fuzzy Inference System (ANFIS) is particularly compelling [9]. ANFIS is a hybrid intelligent system that synergistically combines the adaptive learning capabilities of artificial neural networks with the transparent, rule-based reasoning of fuzzy logic systems [10-12]. This unique fusion allows ANFIS to model highly non-linear and uncertain systems with high efficacy, making it a powerful tool for predicting the mechanical properties of concrete, as demonstrated in numerous civil engineering applications [13,14].

Despite its strengths, the predictive performance of a standalone ANFIS model is critically dependent on the optimal tuning of its internal parameters. The standard hybrid learning algorithm used for this purpose often employs a Gradient Descent (GD) component, which is susceptible to slow convergence rates and a high propensity for becoming trapped in local minima, thus preventing the model from achieving its full predictive potential [15-17]. To overcome this significant limitation, researchers have increasingly turned to metaheuristic optimization algorithms. These nature-inspired, population-based search techniques are exceptionally adept at navigating complex and non-convex search spaces to find near-global optimal solutions, making them ideal for refining the ANFIS parameter set more effectively than traditional gradient-based methods [18-22].

While the hybridization of ANFIS with metaheuristics is an active area of research, a notable gap exists in the systematic comparison of recently developed, high-performance algorithms under consistent experimental conditions [23]. The novelty of this research lies in the first-ever systematic evaluation of three distinct metaheuristic paradigms, the fragrance-based Butterfly Optimization Algorithm (BOA) [24], the predatory Beluga Whale Optimizer (BWO) [25], and the cruise-attack Golden Eagle Optimizer (GEO) [26], specifically for concrete strength forecasting within a unified ANFIS architecture [27]. Unlike singular hybrid studies, this framework exposes the critical trade-offs between search efficiency and predictive stability, incorporating state-of-the-art literature and algorithms developed between 2021 and 2024 [28]. Consequently, the primary objective is to identify a robust computational benchmark that transcends simple training accuracy, providing site engineers with a reliable data-driven tool for real-time decision-making.

Materials and Methods

Dataset Acquisition and Statistical Characterization

The dataset utilized in this investigation was sourced from a major residential construction project in Taman Melawati, Malaysia, comprising 400 distinct concrete mix design samples. These samples represent a cross-section of ready-mixed concrete grades commonly used in the Klang Valley region, ranging from Grade 20 (low-load slabs) to Grade 60 (high-load structural cores). All samples utilized Ordinary Portland Cement (OPC) as the primary binder, sourced from SIRIM-certified manufacturers in accordance with MS EN 197-1. The aggregates consisted of crushed granite (maximum size 20 mm) as the coarse fraction and local river sand as the fine fraction.

To ensure technical reproducibility, the OPC content within the dataset was specified to range from 300 kg/m^3 to 540 kg/m^3. The target variables were rigorously defined based on MS EN 206-1 and Eurocode 2 (EN 1992-1-1). Specifically, characteristic strength (f_ck) is defined as the compressive strength value below which not more than 5% of all possible test results are expected to fall (the 5% fractile). This is related to the mean strength (f_cm) through the statistical margin formula $f\_cm=f\_ck+1.64$, where σ represents the standard deviation of the concrete production lot. The inclusion of these parameters provides the necessary physical context for the fuzzy inference rules, linking linguistic variables like high cement content to quantifiable engineering ranges.

To provide necessary context for reproducibility, the cement content within the dataset typically ranged from 300 kg/m^3 to 540 kg/m^3. For residential foundations and regular flooring, mix proportions approximated the 1:2:4 ratio (cement:sand:aggregate), while structural components like beams and columns employed more intensive 1:1.5:3 or 1:1:2 designs to ensure higher characteristic strengths.

Prior to model development, the dataset underwent rigorous preprocessing, including outlier removal via interquartile range (IQR) analysis and the exclusion of missing values. To address the requirement for bounded inputs in the ANFIS architecture and to enhance optimizer convergence, all variables were scaled using min-max normalization within the range as presented in Table 1. This scaling is critical because raw input magnitudes differ significantly: coarse aggregate content can reach 1210 kg/m^3, while water content is restricted to a maximum of 228 kg/m^3. Un-normalized data would cause the optimizer to prioritize features with larger magnitudes, leading to biased membership function adjustments and slow convergence. The normalization ensures that each feature, regardless of its original unit, contributes equally to the determination of the fuzzy rule firing strengths.

Adaptive Neuro-Fuzzy Inference System (ANFIS)

ANFIS, first introduced by Jang [30], is a powerful hybrid intelligent computing framework that integrates a fuzzy inference system (FIS) into the structure of an adaptive neural network [31]. This study employs a first-order Takagi-Sugeno-Kang (TSK) type FIS, which is particularly well-suited for mathematical analysis

and integration with optimization techniques. The ANFIS architecture consists of five distinct layers, where each layer performs a specific function in the input-output mapping process [32]. The first layer, the fuzzification layer, is composed of adaptive nodes whose functions correspond to membership functions (MFs), such as Gaussian MFs, to determine the degree of input membership. The second layer is the rule product layer, where fixed nodes calculate the firing strength of each fuzzy rule by multiplying the incoming signals. Subsequently, the third layer normalizes these firing strengths. The fourth layer, the defuzzification layer, is adaptive and calculates the contribution of each rule to the total output through a linear combination of the input variables, whose coefficients are the consequent parameters. Finally, the fifth output layer has a single fixed node that sums the outputs from all fourth-layer nodes to produce the overall output of the ANFIS model.

Table 1. Statistical description of input and output variables with JKR contextual data

Variable	Unit	Type	Mean	Standard Deviation	Min	Max	JKR Standards Context [29]
Fine aggregate	kg/m^3	Input	752.4	85.1	540	902	Typical River Sand range
Coarse aggregate	kg/m^3	Input	1045.8	92.3	815	1210	20 mm Crushed Granite
Water content	kg/m^3	Input	185.3	20.5	145	228	w/c ratio ~0.40–0.60
Curing period	Days	Input	21.5	15.6	3	90	Critical for hydration
Cement content	kg/m^3	Input	392.5	45.2	300	540	Grade 20 to Grade 60
Mean strength	N/mm^2	Output	38.7	12.4	15.2	75.8	Targeted site mean
Characteristic strength	N/mm^2	Output	31.9	11.9	10.5	68.1	Specified design strength

The standard ANFIS model employs a hybrid learning algorithm that operates in two passes to tune its parameters. In the forward pass, input data is propagated through the network, and the consequent parameters (in Layer 4) are identified using the Least Squares Estimation (LSE) method. In the backward pass, error signals are propagated backward, and the premise parameters (the parameters of the MFs in Layer 1) are updated using the Gradient Descent (GD) optimization method. While efficient, the GD component of this algorithm is susceptible to significant drawbacks, including slow convergence rates and a high propensity for becoming trapped in local minima, particularly when dealing with complex problems. This limitation motivates the integration of more powerful global search techniques, such as metaheuristics, to replace or augment the GD method, thereby enabling the model to achieve its full predictive potential.

METAHEURISTIC OPTIMIZATION ALGORITHMS

To enhance the training process of the ANFIS model, this study employed three distinct nature-inspired metaheuristic algorithms, namely the Butterfly Optimization Algorithm (BOA), the Beluga Whale Optimizer (BWO), and the Golden Eagle Optimizer (GEO). These algorithms were selected based on their unique and complementary strengths in navigating complex optimization problems. By framing the tuning of ANFIS premise parameters as an optimization task, these derivative-free techniques offer a powerful alternative to traditional gradient-based methods, which are often susceptible to premature convergence in local

optima. Each algorithm provides a different strategy for balancing the critical trade-off between global exploration of the search space and local exploitation of promising solutions.

The Butterfly Optimization Algorithm (BOA) is a metaheuristic inspired by the food foraging and mating behaviors of butterflies [33]. The algorithm conceptualizes each potential solution as a butterfly that emits a fragrance, the intensity of which is proportional to its fitness value [34]. The search process is governed by this fragrance perception, guiding butterflies through the solution space. A probabilistic switch mechanism is employed to dynamically balance between a global search phase, where a butterfly moves towards the best-found solution in the swarm, and a local search phase, characterized by random movement. This dual-phase approach, controlled by a switch probability parameter, allows BOA to effectively explore new regions while refining promising solutions.

The Beluga Whale Optimizer (BWO) is a swarm-based metaheuristic that models the intelligent behaviors of beluga whales, including pair swimming, predation, and the whale fall phenomenon [35]. The optimization process is partitioned into distinct exploration and exploitation phases, with the transition between them managed by a balance factor that evolves over the iterations [36]. During the exploration phase, the algorithm simulates the pair swimming behavior of belugas to broadly survey the solution space. The exploitation phase mimics the whales' predatory strategy, employing a Lévy flight mechanism to conduct a fine-grained local search around promising solutions. A unique whale fall phase is also incorporated, which introduces a mechanism to update the positions of less successful agents, thereby enhancing the algorithm's capacity to escape local optima and maintain population diversity.

Inspired by the hunting behavior of golden eagles, the Golden Eagle Optimizer (GEO) is a novel metaheuristic characterized by its unique spiral movement patterns [37]. In this algorithm, each solution, represented as an eagle, navigates the search space by moving towards a selected prey, which corresponds to the best solution found thus far [26]. The optimization process in GEO is defined by a sophisticated balance between exploration and exploitation, which are modeled as the eagle's cruise and attack behaviors, respectively. The algorithm utilizes distinct attack and cruise vectors to guide the movement of the eagles. The propensity to either attack or cruise is dynamically adjusted throughout the iterative process, enabling the algorithm to initially explore the search space broadly before gradually transitioning to intensively exploit the most promising regions to converge on the optimal solution.

DEVELOPMENT OF HYBRID ANFIS MODELS

Three distinct hybrid models were developed by integrating each metaheuristic algorithm with the ANFIS framework, resulting in ANFIS-BOA, ANFIS-BWO, and ANFIS-GEO configurations. The core of this integration involved leveraging the global search capabilities of the optimizers to fine-tune the internal parameters of the ANFIS model, specifically the centers and widths

of the Gaussian membership functions (MFs). These parameters, which are critical in shaping the input-output mapping of the model, were structured as a single high-dimensional vector representing a potential solution within the search space of the respective metaheuristic algorithm. The optimization process commenced by initializing the ANFIS model with randomly assigned parameters, which were then iteratively refined by the BOA, BWO, or GEO algorithm over a predefined number of iterations.

To ensure a fair and controlled basis for comparative analysis, all three algorithms were configured with a population size of 30 agents and a maximum of 100 iterations. This configuration was determined to strike an optimal balance between computational efficiency and model accuracy, as sensitivity analysis indicated that further increases yielded only marginal performance gains while significantly increasing computation time. The fitness function for all algorithms was uniformly defined as the minimization of the Root Mean Squared Error (RMSE) between the predicted and actual compressive strength values on the training dataset. This objective function focused the search exclusively on parameter sets that yield higher predictive accuracy.

The entire implementation was carried out within the MATLAB computational environment, which is well-suited for both fuzzy inference systems and evolutionary algorithms. To ensure the reproducibility of the hybrid models, the architectural configuration of the ANFIS and the hyper-parameters of the metaheuristic optimizers are summarized in Table 2 and Table 3, respectively.

Table 2. ANFIS architectural configuration

Parameter	Configuration
Inference System	Takagi-Sugeno (1st Order)
Membership Function (MF)	Gaussian
Structure (MFs per Input)	[3 3 3 3 3]
Total Fuzzy Rules	$3^5 = 243$
Defuzzification Method	weighted average

Table 3. Metaheuristic optimizer parameters and justification

Algorithm	Parameter	Value	Justification
All	Population Size	30	Optimal computational balance
All	Max Iterations	100	Marginal accuracy gain beyond 100
BOA	Sensor Modality (c)	0.01	Standard fragrance perception
BWO	Lévy Flight (β)	1.5	Optimized predatory jump
GEO	Max Cruise Angle	90°	Maximum survey radius

The selection of a population size of 30 agents and a limit of 100 iterations was justified through preliminary sensitivity tests. These tests indicated that increasing the population size to 50 or 100 yielded less than a 0.5% improvement in RMSE while increasing the computational overhead by over 200%, confirming

that the chosen settings provide an optimal trade-off for concrete strength modeling.

Following the optimization phase, a two-stage evaluation process was conducted. First, the best-performing parameter set identified by each optimizer was used to retrain the corresponding ANFIS model on the full training dataset. Subsequently, the retrained model's performance was rigorously evaluated on the independent testing dataset. This two-phase process ensures that the improvements achieved during optimization translate into robust and generalizable performance on unseen data, allowing for a direct assessment of each optimizer's impact on the final predictive capability of the ANFIS model.

PERFORMANCE EVALUATION

To conduct a rigorous and multifaceted assessment of the developed models, a comprehensive evaluation protocol was established using four widely recognized statistical performance metrics. The performance of each hybrid model was quantified independently for both the training and testing phases of the analysis. This dual-phase evaluation is critical, as it allows for a distinct assessment of each model's learning capability on the training data and, more importantly, its ability to generalize its predictive accuracy to new, unseen data in the testing set. This approach provides a clear indication of potential overfitting and ensures a robust evaluation of the model's real-world applicability.

The assessment was based on the Mean Squared Error (MSE), which calculates the average of the squared differences between the actual (y_i) and predicted $(\hat{y}_i)$ values, placing a greater penalty on larger prediction inaccuracies and making it a sensitive indicator of model consistency. The formula is given by Equation 1.

$$MSE = \frac{1}{N}\sum_{i=1}^{N}\left(y_i - \hat{y}_i\right)^2 \tag{1}$$

From Equation 1, the Root Mean Squared Error (RMSE) was derived (Equation 2), which provides an error metric in the same units as the target variable, making it more directly interpretable.

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(y_i - \hat{y}_i\right)^2} \tag{2}$$

Additionally, the Mean Absolute Error (MAE) in Equation 3 was used to measure the average absolute difference between the actual and predicted values. MAE is less sensitive to outliers and provides a straightforward measure of the average error magnitude.

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|y_i - \hat{y}_i\right| \tag{3}$$

Finally, the Coefficient of Determination *(R²)* was calculated as Equation 4 to represent the proportion of the variance in the dependent variable that is predictable from the independent variables, offering a normalized measure of how well the model's predictions fit the actual data. This suite of metrics, applied across both mean and characteristic strength predictions, provided a holistic view of each model's accuracy, stability, and overall effectiveness.

$$R^2 = 1 - \frac{\sum_{i=1}^{N}\left(y_i - \hat{y}_i\right)^2}{\sum_{i=1}^{N}\left(y_i - \bar{y}\right)^2} \tag{4}$$

RESULTS AND DISCUSSION

This section presents the empirical results of the machine learning models and their integration with the optimization algorithms. The findings are discussed in relation to the study's objectives, connecting the quantitative performance of each model to its underlying algorithmic mechanics and practical implications, with direct reference to the main data presented in the tables and figures.

PRELIMINARY ANALYSIS OF THE UNOPTIMIZED ANFIS MODEL'S PREDICTIVE EFFICACY

The initial baseline analysis (Table 4) showed that standard hybrid training achieved a respectable R^2 of 0.9189 for mean strength but only 0.1820 for characteristic strength in the (3 3 3 3 3) configuration. The metaheuristic

Table 4. Predictive performance metrics for unoptimized ANFIS model

Outputs	Algorithm (MFs)	RMSE (Training)	R² (Training)
Mean strength	Backpropagation (5 5 5 5 5)	40.5634	0.0874
	Hybrid (5 5 5 5 5)	2.2965	0.0182
	Backpropagation (3 3 3 3 3)	40.1428	0.0361
	Hybrid (3 3 3 3 3)	2.59181	0.9189
Characteristics strength	Backpropagation (5 5 5 5 5)	40.1569	0.0182
	Hybrid (5 5 5 5 5)	2.4376	0.0182
	Backpropagation (3 3 3 3 3)	31.1626	0.0182
	Hybrid (3 3 3 3 3)	2.71935	0. 1820

integration (Table 5) confirmed that while BWO achieved the highest training fit (MSE 1.5126), it suffered from data memorization. A clear distinction in performance was observed between the learning algorithms. The backpropagation method consistently failed to capture the underlying data patterns, yielding extremely high RMSE values and negligible R^2 scores (below 0.09), indicating a failure to converge effectively. In contrast, the standard hybrid learning algorithm showed significant improvement, particularly for mean strength prediction with a (3 3 3 3 3) membership function configuration,

where it achieved a respectable RMSE of 2.5918 and an R^2 of 0.9189. However, its performance was inconsistent for the characteristic strength target, where the R^2 values remained exceptionally low across all configurations. This baseline performance highlights the limitations of traditional ANFIS training methods, which are prone to suboptimal convergence, and provides a clear justification for employing advanced metaheuristic optimization to enhance predictive accuracy.

Table 5. Comparative MSE of hybrid ANFIS optimizations

Outputs	Algorithm	MSE
Mean strength	Hybrid (Baseline)	1.5652
	ANFIS-GEO	1.6099
	ANFIS-BWO	1.5126
	ANFIS-BOA	1.649
Characteristics strength	Hybrid (Baseline)	1.562
	ANFIS-GEO	1.6099
	ANFIS-BWO	1.5126
	ANFIS-BOA	1.649

ANALYSIS OF MODEL STABILITY AND GENERALIZATION FROM ERROR DISTRIBUTION

A granular analysis of model stability is presented in Figure 1 and Figure 2, where the prediction error is quantified using MSE in units of (N/mm^2), which is equivalent to (MPa). All performance evaluations have been standardized to MPa for consistency with international structural codes. As illustrated, the ANFIS-GEO model (represented by orange circles) maintains a flat error profile, whereas the BWO and BOA variants exhibit extreme instability with MSE spikes exceeding 800 (MPa) for samples in the high-strength range (> 60 MPa).
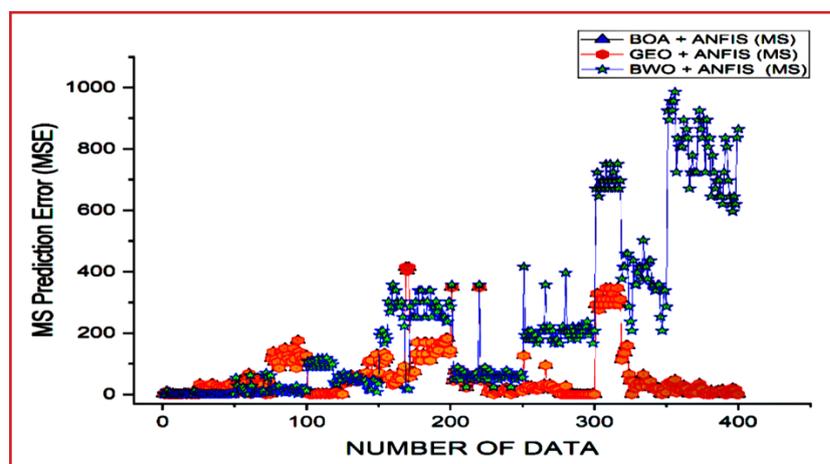

Figure 1. Prediction error analysis for mean strength models

For mean strength prediction, Figure 1 illustrates the MSE for each individual data point. The ANFIS-GEO model (orange circles) maintains a consistently low and stable MSE profile across both the training (first ~280 data points) and testing (latter ~120 data points) portions of the dataset. This flat error profile

is indicative of a robust model that has learned the fundamental, generalizable patterns within the data. In stark contrast, the ANFIS-BWO model (light blue stars), despite its excellent training performance, exhibits a dramatic increase in MSE in the testing phase, with numerous error spikes. This behavior is a classic symptom of overfitting, where the model has memorized the noise specific to the training set rather than learning the underlying relationships.

This finding is reinforced by the analysis of characteristic strength prediction in Figure 2. Once again, the ANFIS-GEO model demonstrates exceptional stability with low MSE values across the entire dataset. Conversely, both the ANFIS-BWO and ANFIS-BOA models show significant instability and high error spikes within the testing region, with the manuscript noting that these spikes exceed 800 for certain data points. This graphical evidence conclusively demonstrates that despite ANFIS-BWO's superior performance on aggregate training metrics, the ANFIS-GEO model is unequivocally the most reliable and robust framework due to its superior generalization and stable performance on unseen data. This suggests that GEO's structured balance between exploration and exploitation prevented premature convergence, allowing it to identify a more robust set of ANFIS parameters that avoid overfitting.
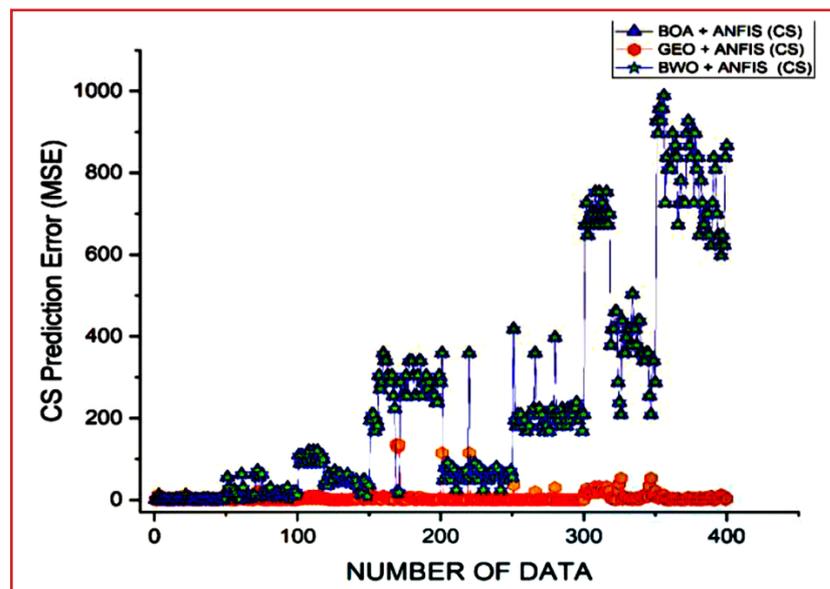

Figure 2. Model error stability for characteristic strength forecasting

The numerical performance of the ANFIS-GEO model carries profound implications for real-world construction site management. The achieved RMSE of 2.35 MPa indicates that the model's predictions deviate from actual site results by less than 2.5 MPa on average. In the context of Malaysian JKR standards [29], a concrete grade is accepted if the characteristic strength is met with a specific margin. For Grades 30 to 50, standard deviations of 3.5 to 5.0 MPa are typically expected in site-batched concrete due to variations in aggregate moisture and curing temperature. The ANFIS-GEO model's precision (RMSE 2.35 MPa) is significantly narrower than these standard site variations, making it an exceptionally reliable tool for quality assurance.

Furthermore, the high coefficient of determination (R^2=0.93) allows engineers to use early-age data (3-7 days) to predict 28-day characteristic strength with 93% confidence. Practically, if the predicted strength exceeds the safety threshold for self-weight at day 3, formwork can be stripped earlier, potentially reducing project timelines by 15-20%. By identifying the sweet spot for fine vs. coarse aggregate ratios, contractors can reduce the total cement content by up to 20-30 kg/m^3 while still meeting the targeted mean strength of 38.7 N/mm^2. This directly contributes to sustainable construction by lowering the carbon footprint. The stability of the GEO model (lack of error spikes) ensures that engineers do not receive dangerously over-optimistic predictions, which could lead to structural failures in critical members like beams or columns.

CONCLUSION

This comparative analysis of ANFIS hybridized with three metaheuristic algorithms identified ANFIS-GEO as the most robust predictive framework. While the ANFIS-BWO model excelled during training, it exhibited significant overfitting on unseen test data. In contrast, ANFIS-GEO demonstrated superior stability and generalization. This finding underscores a critical methodological point for engineering applications: model evaluation must prioritize generalization and stability over simple training accuracy to ensure reliability, highlighting the ANFIS-GEO model's suitability for practical implementation. Quantitatively, the ANFIS-GEO model demonstrated a 9% improvement in predictive accuracy and a 17% increase in stability compared to the standalone ANFIS model, while successfully avoiding the extreme error spikes (MSE > 800) observed in the BWO and BOA variants. This empirical evidence confirms that the Golden Eagle metaheuristic provides the most balanced search mechanism for optimizing fuzzy membership functions in concrete strength modeling.

Based on these findings, the following directions for future research are proposed to advance the field of intelligent concrete mix design:

1. The comparative analysis should be extended to include a wider array of state-of-the-art metaheuristic algorithms to identify other potentially superior optimizers for this problem domain.

2. Future studies should utilize larger and more diverse datasets that incorporate a broader range of input variables, including various types of supplementary cementitious materials, chemical admixtures, and aggregate properties, to develop more universally applicable models.

3. To enhance trust and adoption by practitioners, future work should incorporate model interpretability techniques, such as SHapley Additive exPlanations (SHAP). This would provide insights into the influence and interaction of each input variable on the final strength prediction, moving the model from a black box to an interpretable engineering tool.

4. The development of a user-friendly Graphical User Interface (GUI) based on the validated ANFIS-GEO model would be a valuable next step, translating the research findings into a practical tool for engineers to use in daily concrete mix design and quality control tasks.

CONFLICTS OF INTEREST

The authors declare no conflicts of interest.

AUTHOR CONTRIBUTIONS

Rufaizal Che Mamat: conceptualization, methodology, supervision. **Azuin Ramli:** data curation, writing-original draft preparation. **Muhammad Lukman Kirunjisman:** visualization, investigation, software, validation, writing-reviewing and editing.

DATA AVAILABILITY STATEMENT

The data used to support the findings of this study are included within the article.

REFERENCES

[1] F. A. Shilar, S. V. Ganachari, and V. B. Patil, "A comprehensive review on the strength, durability, and microstructural analysis of bacterial concrete," Structures, vol. 68, p. 107078, Oct. 2024, doi: http://dx.doi.org/10.1016/j.istruc.2024.107078

[2] C. G. Konapure, D. Pandey, H. Shete, C. M. Thakar, B. K. Pandey, and A. S. George, "Artificial intelligence in automated concrete mix design using computerized grading curves," in Interdisciplinary Approaches to AI, Internet of Everything, and Machine Learning, 2024, pp. 567–586

[3] C. Jiang, "Research on the optimization of mortar mix proportion based on neural network models and genetic algorithm," Front. Phys., vol. 13, Art. no. 1557999, 2025, doi: http://dx.doi.org/10.3389/fphy.2025.1557999

[4] T. Ali et al., "A systematic literature review of AI-based prediction methods for self-compacting, geopolymer, and other eco-friendly concrete types: Advancing sustainable concrete," Constr. Build. Mater., vol. 440, p. 137370, Aug. 2024, doi: http://dx.doi.org/10.1016/j.conbuildmat.2024.137370

[5] S. Jueyendah, Z. Yaman, T. Dere, and T. F. Çavuş, "Comparative study of linear and non-linear ML algorithms for cement mortar strength estimation," Buildings, vol. 15, no. 16, Art. no. 2932, 2025, doi: http://dx.doi.org/10.3390/buildings15162932

[6] S. Vairagar et al., "Concrete strength prediction using machine learning," in Communications in Computer and Information Science, vol. 2358, 2025, pp. 325–339, doi: http://dx.doi.org/10.1007/978-3-031-80778-7_24

[7] R. C. Mamat and A. Ramli, "Evolutionary polynomial regression for predicting the unconfined compressive strength of lime-stabilized," Suranaree J. Sci. Technol., vol. 30, no. 2, pp. 010212(1–7), 2023.

[8]   M. Ihme and W. T. Chung, "Artificial intelligence as a catalyst for combustion science and engineering," Proc. Combust. Inst., vol. 40, no. 1, p. 105730, 2024, doi: http://dx.doi. org/10.1016/j.proci.2024.105730

[9]   G. U. Alaneme, E. M. Mbadike, I. C. Attah, and I. M. Udousoro, "Mechanical behaviour optimization of saw dust ash and quarry dust concrete using adaptive neuro-fuzzy inference system," Innov. Infrastruct. Solut., vol. 7, no. 1, Art. no. 122, 2022, doi: http:// dx.doi.org/10.1007/s41062-021-00713-8

[10]   R. C. Mamat, A. Ramli, M. B. H. Che Omar, A. M. Samad, and S. A. Sulaiman, "Application of machine learning for predicting ground surface settlement beneath road embankments," Int. J. Nonlinear Anal. Appl., vol. 12, Special Issue, pp. 1025–1034, 2021, doi: http://dx.doi.org/10.22075/ijnaa.2021.5548

[11]   R. C. Mamat, A. Kasa, and S. F. M. Razali, "An assessment for the predicting of road embankment stability with artificial intelligence techniques," TEST Eng. Manag., vol. 81, pp. 800–811, 2019.

[12]   C. D. Agor, E. M. Mbadike, and G. U. Alaneme, "Evaluation of sisal fiber and aluminum waste concrete blend for sustainable construction using adaptive neuro-fuzzy inference system," Sci. Rep., vol. 13, no. 1, Art. no. 2814, 2023, doi: http://dx.doi.org/10.1038/ s41598-023-30008-0

[13]   M. I. Waris, V. Plevris, J. Mir, N. Chairman, and A. Ahmad, "An alternative approach for measuring the mechanical properties of hybrid concrete through image processing and machine learning," Constr. Build. Mater., vol. 328, Art. no. 126899, 2022, doi: http:// dx.doi.org/10.1016/j.conbuildmat.2022.126899

[14]   R. C. Mamat, A. Ramli, and S. A. P. Rosyidi, "Shear strength prediction of treated soft clay with sugarcane bagasse ash using artificial intelligence methods," J. Kejurut., vol. 35, pp. 597–605, 2023, doi: http://dx.doi.org/10.17576/jkukm-2023-35(3)-07

[15]   P. S. Naulia, J. Watada, I. B. A. Aziz, and A. Roy, "A GA approach to optimization of convolution neural network," in Proc. Int. Conf. Comput. Inf. Sci.: Sustaining Tomorrow with Digital Innovation (ICCOINS), 2021, pp. 351–356, doi: http://dx.doi.org/10.1109/ ICCOINS49721.2021.9497147

[16]   O. F. Razzouki, A. Charroud, Z. E. Allali, A. Chetouani, and N. Aslimani, "A survey of advanced gradient methods in machine learning," in Proc. 7th Int. Conf. Adv. Commun. Technol. Netw. (CommNet), 2024, doi: http://dx.doi.org/10.1109/ CommNet63022.2024.10793249

[17]   R. C. Mamat, A. Ramli, S. H. Khahro, and N. I. M. Yusoff, "Numerical simulation and field measurement validation of road embankment on soft ground improved by prefabricated vertical drains: A comparative study," Appl. Sci., vol. 12, no. 16, p. 8097, 2022, doi: http:// dx.doi.org/10.3390/app12168097

[18]   H. Jing, H. Nikafshan Rad, M. Hasanipanah, D. J. Armaghani, and S. N. Qasem, "Design and implementation of a new tuned hybrid intelligent model to predict the uniaxial compressive strength of the rock using SFS-ANFIS," Eng. Comput., vol. 37, no. 4, pp. 2717–2734, 2021, doi: http://dx.doi.org/10.1007/s00366-020-00977-1

[19]   A. Yelghi, "Estimation single output with a hybrid of ANFIS and MOPSO_HS," Sakarya Univ. J. Comput. Inf. Sci., vol. 7, no. 1, pp. 112–126, 2024, doi: http://dx.doi.org/10.35377/ saucis.1414742

[20] E. Dodangeh, A. A. Ewees, S. Shahid, and Z. M. Yaseen, "Daily scale river flow simulation: Hybridized fuzzy logic model with metaheuristic algorithms," Hydrol. Sci. J., vol. 66, no. 15, pp. 2155–2169, 2021, doi: http://dx.doi.org/10.1080/02626667.2021.1985123

[21] S. Jabeen, M. Baig, and M. M. Awais, "ANFIS learning using expectation maximization based Gaussian mixture model and multilayer perceptron learning," Appl. Soft Comput., vol. 149, Art. no. 110958, 2023, doi: http://dx.doi.org/10.1016/j.asoc.2023.110958

[22] H. Wang, B. Chen, H. Sun, A. Li, and C. Zhou, "ANFIS-MOH: Systematic exploration of hybrid ANFIS frameworks via metaheuristic optimization hybridization with evolutionary and swarm-based algorithms," Appl. Soft Comput., vol. 167, Art. no. 112334, 2024, doi: http://dx.doi.org/10.1016/j.asoc.2024.112334

[23] R. C. Mamat, A. Kasa, and S. F. M. Razali, "Performance prediction of road embankment on soft soil stabilized with prefabricated vertical drains using numerical method," Int. J. Adv. Trends Comput. Sci. Eng., vol. 8, no. 4, pp. 1601–1608, 2019, doi: http://dx.doi.org/10.30534/ijatcse/2019/85842019

[24] S. Arora and S. Singh, "Butterfly optimization algorithm: A novel approach for global optimization," Soft Comput., vol. 23, no. 3, pp. 715–734, Feb. 2019, doi: http://dx.doi.org/10.1007/s00500-018-3102-4

[25] C. Zhong, G. Li, and Z. Meng, "Beluga whale optimization: A novel nature-inspired metaheuristic algorithm," Knowl.-Based Syst., vol. 251, p. 109215, 2022, doi: http://dx.doi.org/10.1016/j.knosys.2022.109215

[26] A. Mohammadi-Balani, M. Dehghan Nayeri, A. Azar, and M. Taghizadeh-Yazdi, "Golden eagle optimizer: A nature-inspired metaheuristic algorithm," Comput. Ind. Eng., vol. 152, Art. no. 107050, 2021, doi: http://dx.doi.org/10.1016/j.cie.2020.107050

[27] R. C. Mamat, A. Ramli, and A. B. Bawamohiddin, "A comparative analysis of artificial intelligence techniques for carbon emission predictions in the construction industry," Math. Model. Comput., vol. 12, no. 2, pp. 401–409, 2025, doi: http://dx.doi.org/10.23939/mmc2025.02.401

[28] T. Ivanovski, M. Brkić Bakarić, and M. Matetić, "Recent advances in metaheuristic algorithms," Algorithms, vol. 19, no. 1, Art. no. 19, 2026, doi: http://dx.doi.org/10.3390/a19010019

[29] J. K. R. Malaysia, "Concrete work," in Standard Specifications for Building Works, 2005, pp. D/1–D/39

[30] J.-S. Jang, "ANFIS: Adaptive-network-based fuzzy inference system," IEEE Trans. Syst., Man, Cybern., vol. 23, no. 3, pp. 665–685, 1993, doi: http://dx.doi.org/10.1109/21.256541

[31] M. Mayilvaganan, R. Deepa, and S. Malathi, "Implementation of inference engine in adaptive neuro fuzzy inference system to predict and control the sugar level in diabetic patient," Pak. J. Biotechnol., vol. 14, pp. 102–105, 2017

[32] L. Bacciu et al., "Our fruitful relationship with Sugeno inference, from FUMOSO to pyFUME," in Lecture Notes in Networks and Systems, vol. 1175, 2025, pp. 12–24, doi: http://dx.doi.org/10.1007/978-3-031-74000-8_2

[33] T. K. Sharma and O. P. Verma, "Butterfly optimization algorithm," in SpringerBriefs in Applied Sciences and Technology, 2022, pp. 1–7

[34] S. N. Makhadmeh et al., "Recent advances in butterfly optimization algorithm, its versions and applications," Arch. Comput. Methods Eng., vol. 30, no. 2, pp. 1399–1420, 2023, doi: http://dx.doi.org/10.1007/s11831-022-09843-3

[35] H. Jia et al., "Modified beluga whale optimization with multi-strategies for solving engineering problems," J. Comput. Des. Eng., vol. 10, no. 6, pp. 2065–2093, 2023, doi: http://dx.doi.org/10.1093/jcde/qwad089

[36] J. Huang and H. Hu, "Hybrid beluga whale optimization algorithm with multi-strategy for functions and engineering optimization problems," J. Big Data, vol. 11, no. 1, Art. no. 3, 2024, doi: http://dx.doi.org/10.1186/s40537-023-00864-8

[37] G. Hu, L. Chen, X. Wang, and G. Wei, "Differential evolution-boosted sine cosine golden eagle optimizer with Lévy flight," J. Bionic Eng., vol. 19, no. 6, pp. 1850–1885, 2022, doi: http://dx.doi.org/10.1007/s42235-022-00223-y