Journal of Intelligent Systems and Internet of Things
Vol. 17, No. 02, PP. 191-213, 2025
DOI: https://doi.org/10.54216/JISIoT.170212

Forecasting CO_2 Emissions from Cement Manufacturing with iHOW-Tuned Machine Learning Models

Omnia M. Osama[1,*], **El-Sayed M. El-Rabaie**[2], **Marwa M. Eid**[3,4]

[1]Department of Communications & Electronics,Delta Higher Institute of Engineering & Technology , Mansoura, Egypt

[2]Faculty of Electronic Engineering , Menoufia University , Department of Electronics and Communications , Menouf 32952, Egypt

[3]Faculty of Artificial Intelligence , Delta University for Science and Technology , Mansoura 35111, Egypt

[4]Jadara Research Center , Jadara University , Irbid 21110, Jordan

Emails: Omnya.osama@dhiet.edu.eg; srabie1@yahoo.com; mmm@ieee.org

Abstract

Cement production is a major contributor to global CO_2 emissions, posing a challenge for climate mitigation efforts. Accurate forecasting of these emissions is vital for guiding policy and industrial decarbonization. This study addresses the need for improved predictive frameworks by developing an optimized ensemble-based machine learning model for CO_2 emissions forecasting. The model is trained on a corrected global cement emissions dataset and enhanced through hyperparameter tuning using ten metaheuristic algorithms. Among them, the Improved Henry's Optimization Algorithm (iHOW) achieved superior performance. The iHOW-optimized model attained an MSE of 1.21×10^{-6} and R^2 of 0.9657, improving over the best baseline model (Gradient Boosting: MSE = 0.0164, R^2 = 0.8621) by more than 99%. These results confirm the effectiveness of iHOW in producing accurate and reliable forecasts. The proposed framework offers strong potential for integration into carbon tracking systems and policy support tools.

Received: January 01, 2025 Revised: February 05, 2025 Accepted: March 03, 2025

Keywords: CO_2 emissions; Cement industry; Ensemble models; iHOW algorithm; Metaheuristic optimization

1 Introduction

The continuous accumulation of carbon dioxide (CO_2) in the atmosphere has emerged as a defining challenge of the twenty-first century[1], [2]. The rising concentration of greenhouse gases is unequivocally linked to global warming and climate change, with severe consequences ranging from rising sea levels and altered weather patterns to biodiversity loss and food insecurity. Industrial activities, particularly those that are energy-intensive and resource-demanding, have been identified as major contributors to these emissions. Within this context, the cement industry has attracted substantial attention. Cement is a vital material for modern societies, serving as the foundation of housing, transportation networks, and industrial infrastructure. However, the very process that makes cement production indispensable also renders it one of the most carbon-intensive industries worldwide[3], [4]. The thermal decomposition of limestone during clinker production and the extensive use of fossil fuels to power kilns contribute heavily to anthropogenic CO_2 emissions. On average, the production of one metric ton of cement releases nearly a corresponding ton

of CO_2, making the sector responsible for a significant share of total global emissions. This dual role of cement—essential for development yet detrimental to the environment—underscores the urgency of advancing reliable monitoring and forecasting frameworks that can guide sustainable practices[5]. Accurate prediction of cement-related CO_2 emissions has far-reaching importance. From a policy perspective, reliable forecasts provide the empirical basis required to design effective regulations, track progress toward climate targets, and ensure compliance with international agreements aimed at reducing global emissions[6]. For industry stakeholders, predictive models serve as decision-support tools, enabling efficient allocation of resources, optimization of production processes, and assessment of the long-term impacts of adopting emerging technologies such as low-carbon cements, waste-derived fuels, or carbon capture and storage. In addition, for the scientific community, forecasting frameworks are indispensable for understanding emission dynamics and informing integrated assessment models that link industrial activity to climate outcomes. Hence, prediction is not only a technical challenge but also a societal necessity, bridging the gap between industrial growth and environmental stewardship[7], [8]. Nevertheless, the task of forecasting CO_2 emissions from cement production is inherently complex. The datasets available for analysis are global in scope, spanning decades of production statistics across diverse countries and regions. This breadth introduces high dimensionality, which complicates the training of predictive models and often leads to the "curse of dimensionality," where the number of features overwhelms the model's capacity to identify meaningful patterns[9], [10]. Compounding this issue is the presence of feature redundancy: national and regional data are often strongly correlated, and without careful selection, redundant inputs may obscure critical relationships and degrade predictive accuracy. The sensitivity of machine learning models to hyperparameters further intensifies this challenge. Models such as Gradient Boosting, Random Forests, and Support Vector Regression exhibit considerable variability in performance depending on the chosen configuration, necessitating systematic optimization procedures. Moreover, ensuring that models generalize beyond training data remains a central concern. While it is possible to achieve near-perfect predictions on historical datasets, industrial and economic conditions are constantly evolving, and models that overfit past data often fail to provide accurate forecasts under new circumstances. These challenges collectively highlight the need for innovative methodological approaches that can handle complex datasets, reduce redundancy, and optimize predictive performance in a consistent and robust manner[11]. The present study is motivated by these challenges and proposes a comprehensive framework that integrates machine learning techniques with metaheuristic optimization algorithms to forecast cement-related CO_2 emissions more effectively. Machine learning offers distinct advantages over traditional statistical approaches, particularly its ability to capture nonlinear dependencies and uncover latent structures within large datasets. Yet, the predictive capacity of machine learning models is often constrained by limitations in feature selection and hyperparameter calibration. Metaheuristic algorithms, inspired by natural, social, and physical processes, provide an effective solution. Their stochastic search mechanisms allow them to navigate vast and complex parameter spaces with efficiency, avoiding the pitfalls of local optima that commonly affect deterministic methods. By embedding these algorithms within ensemble-based machine learning frameworks, it becomes possible to systematically enhance predictive performance, reduce computational inefficiency, and improve generalization across heterogeneous datasets. This integration, therefore, represents a methodological advancement with direct practical implications for emission monitoring and sustainable industrial planning. The contributions of this work can be articulated along three dimensions. First, the study introduces an integrated methodological framework that combines machine learning models with state-of-the-art metaheuristic optimization algorithms. This synergy is designed to exploit the predictive strengths of machine learning while addressing its inherent weaknesses in feature and parameter selection. Second, the research delivers a comparative analysis of baseline machine learning models and their optimized counterparts, offering new insights into the advantages of embedding optimization into predictive pipelines. This comparison provides a nuanced understanding of how different models respond to optimization and where their relative strengths lie. Third, the study emphasizes the importance of joint optimization, where feature selection and hyperparameter tuning are undertaken simultaneously. Such integration not only improves model accuracy but also enhances interpretability, offering more efficient and transparent forecasting tools. Together, these contributions extend the frontiers of emission prediction research and provide actionable insights for policymakers and industry stakeholders committed to reducing the carbon footprint of cement production.

The remainder of this paper is structured to reflect the logical development of the research. The next section introduces the dataset, describing its sources, scope, and the preprocessing strategies implemented to ensure analytical readiness. It also presents the methodological framework, encompassing the selection of machine learning models and the integration of metaheuristic algorithms for optimization. This is followed by an empirical evaluation, where the performance of baseline models is systematically compared with optimized models, highlighting the improvements achieved through the proposed framework. The discussion section then

interprets these findings, situating them within the broader academic discourse and industrial context, while also identifying current limitations and potential directions for refinement. Finally, the paper concludes by summarizing the key contributions, underscoring their practical implications, and outlining promising avenues for future research to further enhance predictive modeling of cement-related CO_2 emissions.

2 Literature Review

The accurate prediction of carbon dioxide (CO_2) emissions is crucial for mitigating climate change, designing effective environmental policies, and promoting sustainable industrial practices. Numerous recent studies have focused on leveraging artificial intelligence (AI), machine learning (ML), and optimization algorithms to create advanced forecasting models that address the complexity and variability of CO_2 emissions across different sectors. Early work demonstrated the utility of metaheuristic optimization techniques for emission forecasting, where Bat Algorithm (BA) and Cuckoo Optimization Algorithm (COA) were applied to model global CO_2 emissions based on historical consumption data of oil, natural gas, coal, and primary energy [12]. Using data from 1980 to 2018, the COA-exponential model emerged as the most effective, providing low Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and Mean Absolute Error (MAE), thereby serving as a benchmark for emission modeling and environmental decision-making. Building on this foundation, hybrid models have been developed to improve generalization and accuracy. For instance, the CRLPSO-LSTM model integrated an enhanced Particle Swarm Optimization (PSO) approach with Long Short-Term Memory (LSTM) networks to predict residential daily CO_2 emissions [13]. The algorithm featured innovations such as chaotic mapping, Lévy flight strategies, and optimized mutation techniques to enhance diversity and global search efficiency, leading to improved predictions across datasets from China, the United States, and Russia. In the cement industry, where CO_2 emissions during clinker production are influenced by raw material composition, temperature conditions, and particle distribution, deep learning techniques have demonstrated superior capabilities. Models such as Deep Neural Networks (DNN), Ant Colony Optimization-based Artificial Neural Networks (ACO-ANN), and Genetic Algorithm-based ANN (GA-ANN) excelled in capturing nonlinear relationships without requiring explicit prior knowledge of influencing parameters [14]. Further advancements in ensemble learning have improved industrial forecasting accuracy. A Random Forest–Multi-Layer Perceptron–Logistic Regression (RF-MLP-LR) model successfully predicted cement consumption across 31 provinces in China, achieving a Mean Absolute Percentage Error (MAPE) below 10% in most cases [15]. This work emphasized the influence of socioeconomic variables such as per capita GDP, urbanization rates, and housing development, providing data-driven insights to support China's low-carbon development goals. Similarly, nature-inspired optimization algorithms have been combined with neural networks to improve emission modeling for G8 nations. A study using Heap-Based Optimizer (HBO), Teaching-Learning-Based Optimization (TLBO), Whale Optimization Algorithm (WOA), Vortex Search (VS), and Earthworm Optimization Algorithm (EWA) with Multilayer Perceptrons (MLPs) identified TLBO and VS as superior optimizers, demonstrating strong generalization and reduced overfitting [16]. In industrial process monitoring, PSO-optimized Backpropagation Neural Networks (BPNN) achieved minimal prediction errors, offering practical tools for real-time emission monitoring and energy management [17]. Research has also explored sustainable material alternatives to reduce emissions at their source. Alkali-activated binders, a substitute for Portland cement, were studied using neural networks optimized with various metaheuristic algorithms, with the Genetic Algorithm (GA) emerging as the most effective (MSE = 161.17, R^2 = 0.90) [18]. These findings highlight the potential for AI-driven models to facilitate a transition toward eco-friendly construction materials. Addressing uncertainties in industrial operations, another study introduced surrogate modeling techniques with Artificial Neural Networks integrated into GA-PSO frameworks for predicting and minimizing emissions in cement plant combustion sections [19]. Sensitivity analysis revealed coal and tertiary air flows as primary factors influencing emissions, emphasizing the need for adaptive, real-time optimization strategies. High-accuracy ensemble techniques have also been explored, with one study combining MLPs with algorithms such as Biogeography-Based Optimization (BBO), Black Hole Algorithm (BHA), and Shuffled Complex Evolution (SCE), achieving near-perfect prediction accuracy (R^2 = 0.9999, RMSE = 1.6781) [20]. In addition to emission modeling, machine learning has proven invaluable in predicting properties critical to sustainable construction. For instance, the compressive strength of blended concrete was predicted using models such as XGBoost, Decision Trees, Deep Neural Networks, and Linear Regression, with XGBoost performing best (R^2 = 0.952, RMSE = 4.88) [21]. Bayesian optimization fine-tuned hyperparameters, and

SHAP analysis identified curing age, cement content, and water content as dominant factors affecting strength, offering actionable insights for material design. To provide a consolidated view of these contributions, Table 1 summarizes all ten studies, outlining their focus areas, methodologies, and major findings. This table serves as a quick reference for comparing various approaches and highlights the diversity of optimization algorithms and neural network architectures that have been successfully applied to environmental modeling and material science.

Reference	Focus Area	Methodology	Key Findings and Contributions
[12]	Global CO_2 emissions from energy use	Bat Algorithm (BA), Cuckoo Optimization Algorithm (COA); Linear and Exponential models	COA-exponential model achieved superior accuracy (low MSE, RMSE, MAE); serves as benchmark for emission forecasting.
[13]	Residential daily CO_2 emissions prediction	CRLPSO (chaotic mapping, Lévy flight, mutation) + LSTM neural network	Outperformed other hybrid models across China, US, Russia datasets; strong generalization and high accuracy.
[14]	Cement clinker production emissions	Deep Neural Networks (DNN), ACO-ANN, GA-ANN	Models captured complex nonlinear relations without explicit prior knowledge; improved industrial emission modeling.
[15]	Cement consumption forecasting (China)	RF-MLP-LR ensemble model	Achieved MAPE <10% in most provinces; identified GDP, urbanization as key drivers; supports carbon-neutral policy.
[16]	G8 nations' energy-related emissions	MLP + HBO, TLBO, WOA, VS, EWA	TLBO and VS demonstrated best generalization; hybrid optimization reduced overfitting.
[17]	Industrial CO_2 monitoring	PSO-optimized Backpropagation Neural Network (BPNN)	Provided highly accurate real-time monitoring; aids energy policy planning and production optimization.
[18]	Sustainable cement alternatives	GA-optimized neural networks; alkali-activated binder data	GA achieved MSE = 161.17, R^2 = 0.90; supports eco-friendly material design.
[19]	Combustion emissions under uncertainty	Surrogate ANN modeling + GA-PSO; Sobol/FAST sensitivity	Identified coal and tertiary air as key variables; enables adaptive emission control.
[20]	High-accuracy ensemble modeling	MLP with BBO, BHA, BSA, FSA, SCE	Near-perfect predictions (R^2=0.9999, RMSE=1.6781); powerful nonlinear modeling tool.
[21]	Blended concrete compressive strength	XGBoost, Bayesian optimization, SHAP analysis	XGBoost outperformed rivals (R^2=0.952); curing age, cement, and water content most influential.

Table 1: Summary of Reviewed Studies

3 Materials and Methods

3.1 Dataset Description

The foundation of this study rests on the Global CO_2 Cement Emissions dataset, which provides one of the most comprehensive and systematically compiled records of cement production and associated emissions worldwide. The dataset integrates multiple authoritative sources, thereby ensuring both historical depth and contemporary accuracy. The construction of the dataset begins with values derived from the Carbon Dioxide

Information Analysis Center (CDIAC), which has long served as a key repository of global carbon emission statistics. For consistency with prior work, the present study employs emission values calculated from the 2019 edition of CDIAC, where emissions were directly estimated as a function of cement production. It is important to note that the methodology in CDIAC's subsequent editions underwent significant modification, beginning with the 2020 release. To maintain methodological consistency and to ensure comparability across the study horizon, the 2019 framework was retained as the principal reference point. Cement production data form the backbone of the dataset, and these were primarily sourced from the United States Geological Survey (USGS), a globally recognized provider of industrial and mineral statistics. The USGS data cover the period from 1990 onwards and were used to overwrite CDIAC-derived estimates for the corresponding years, ensuring that the dataset reflects the most accurate and directly reported production statistics. For earlier historical periods, particularly before the dissolution of the Soviet Union, additional adjustments were incorporated. Aggregate Soviet production statistics were carefully disaggregated into individual constituent states, allowing the dataset to provide a coherent representation of national-level emissions even for years preceding the formal establishment of independent republics. This disaggregation was critical to ensuring temporal continuity and spatial granularity, both of which are essential for robust forecasting. Despite the comprehensiveness of the CDIAC and USGS data, certain gaps and inconsistencies inevitably arise in the construction of long-term industrial datasets. Countries with incomplete or missing production records in the most recent years required extrapolation to maintain dataset continuity. These extrapolations were performed using conservative methods to avoid introducing artificial trends, thereby ensuring that the predictive modeling process is grounded in empirically plausible assumptions. Moreover, in recognition of the limitations of global reporting systems, country-specific corrections were applied for several key producers. In particular, production statistics for the United States, China, India, Norway, Sweden, Iran, Saudi Arabia, South Korea, Jamaica, Moldova, Mexico, Namibia, Afghanistan, Argentina, and Egypt were adjusted using national data sources and supplementary reports, which provided higher fidelity than global aggregates. These targeted corrections further enhance the dataset's reliability by aligning global estimates with the most accurate country-level records available. Structurally, the dataset is organized to facilitate both exploratory data analysis and advanced predictive modeling. Each entry links annual cement production statistics to corresponding CO_2 emission values at the country level, thereby creating a panel dataset that combines temporal and spatial dimensions. The target variable in this study is the CO_2 emission attributed to cement production, expressed in units consistent with international reporting standards. The explanatory variables consist primarily of production quantities, though the integration of country-specific adjustments adds nuance by accounting for regional heterogeneity. The dataset thus captures not only aggregate global trends but also the differentiated trajectories of individual countries, reflecting diverse levels of industrialization, policy frameworks, and technological adoption. The relevance of this dataset for the present research lies in its capacity to provide both scale and detail. On one hand, the dataset spans multiple decades and covers virtually all producing countries, enabling the development of models that can generalize across time and geography. On the other hand, the incorporation of granular, country-specific corrections ensures that the models are sensitive to regional variations, which is critical for generating accurate forecasts and policy-relevant insights. This dual quality of breadth and depth makes the Global CO_2 Cement Emissions dataset uniquely suitable for developing advanced predictive frameworks. It provides a robust empirical foundation upon which machine learning models and metaheuristic optimization algorithms can be systematically applied to address the pressing challenge of forecasting emissions in one of the world's most carbon-intensive industries. To highlight the global distribution of CO_2 emissions from the cement industry, a comparative breakdown by country was performed. Figure 1 presents the contribution of the top eight emitting countries along with the aggregated share of all other nations. The visualization demonstrates the dominant role of China, which accounts for more than one-third of total cement-related CO_2 emissions, followed by the USA, India, and Japan. Other major contributors include Russia, Germany, Italy, and South Korea, while the remaining countries collectively form a substantial portion under the "Others" category. This distribution emphasizes the significant regional disparities in cement-related emissions and underlines the need for targeted mitigation strategies in the highest-contributing nations.

To contextualize the scale and temporal evolution of carbon dioxide emissions, a comparative analysis between global and Chinese CO_2 emissions was conducted. Figure 2 depicts the emission trajectories from 1920 to 2020, illustrating the sharp rise in global emissions over the last century and the rapid acceleration of China's emissions, particularly since the early 2000s. While global emissions have shown a steady upward trend, China's contribution has grown disproportionately in recent decades, narrowing the gap with global totals. This visualization underscores both the historical growth of industrial activity worldwide and the critical influence of China as a major driver of contemporary CO_2 emissions.

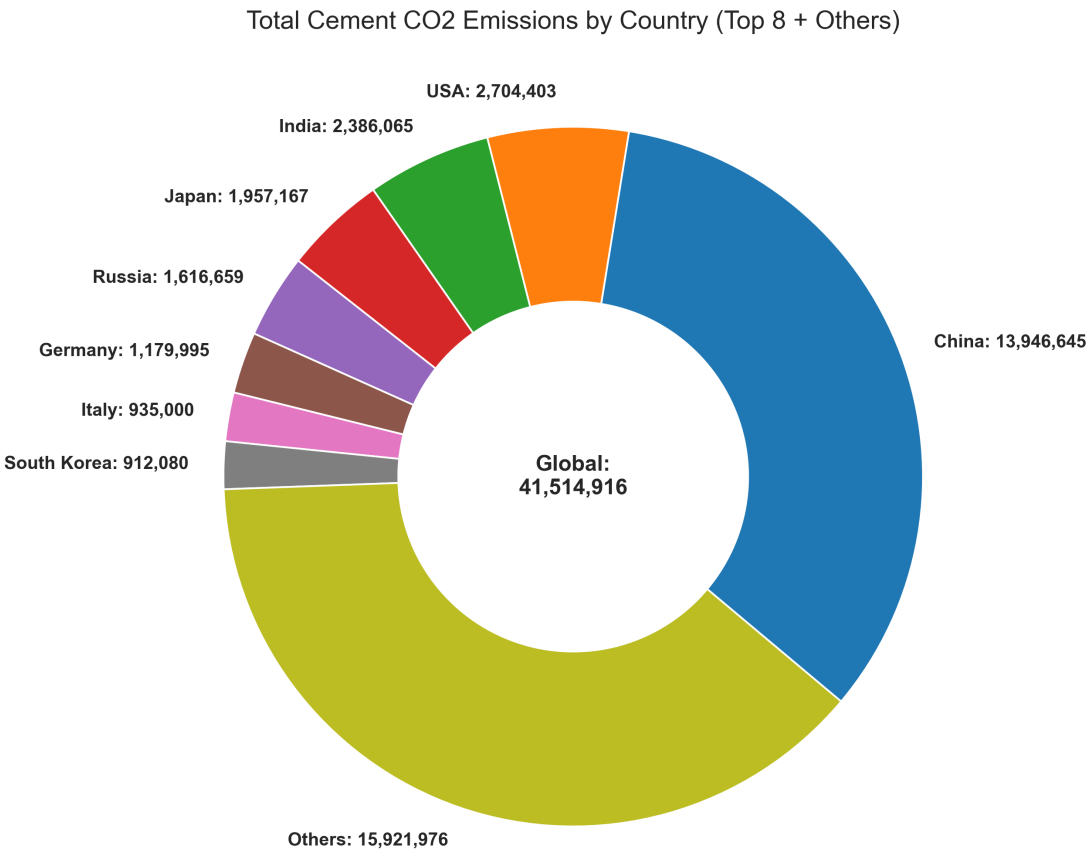

Figure 1: Total CO_2 emissions from cement production by country (top eight contributors and aggregated others).

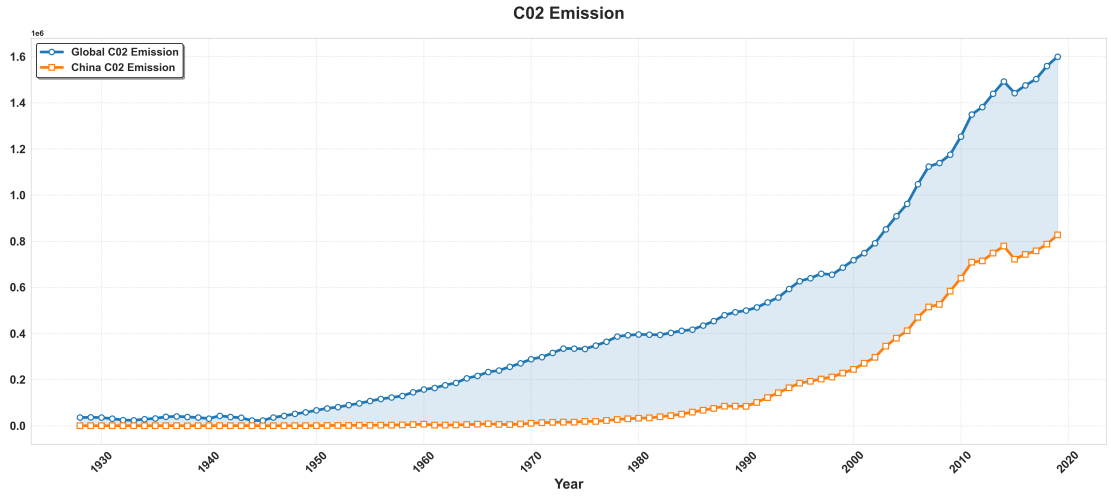

Figure 2: Global and Chinese CO_2 emission trends from 1920 to 2020.

To further examine the dynamics of energy production and its variability across major economies, a comparative analysis between China and the USA was carried out. Figure 3 shows the historical trends of energy production for both countries, along with statistical indicators such as mean, median, and standard deviation. The results reveal that China's energy production has experienced an exponential increase in recent decades, surpassing historical averages by a wide margin, whereas the USA demonstrates relatively stable but fluctuating production patterns over time. The inclusion of mean, median, and variability benchmarks

highlights the extent to which China's growth deviates from its long-term trend, in contrast to the more consistent trajectory observed in the USA. This comparison provides important context for understanding the structural differences in energy generation and their implications for global emissions.

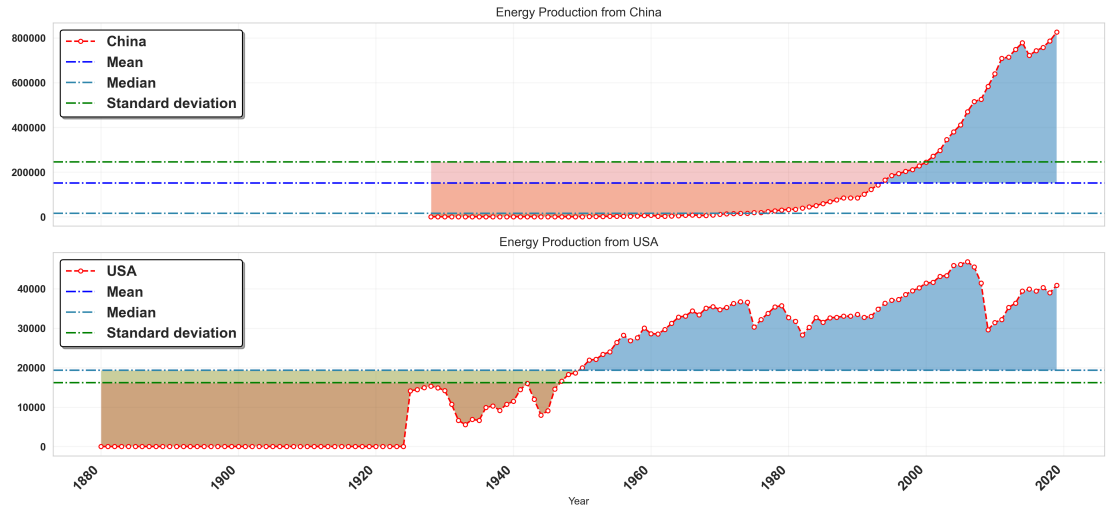

Figure 3: Energy production trends of China and the USA with mean, median, and standard deviation benchmarks.

3.2 Data Preprocessing

Before applying predictive models, the raw Global CO_2 Cement Emissions dataset required systematic preprocessing to ensure its analytical suitability. Preprocessing is a critical stage in machine learning workflows, as the quality, consistency, and structure of the input data directly influence the robustness and reliability of model outputs. Given the historical breadth and geographical diversity of the dataset, several targeted procedures were undertaken, including the handling of missing values, encoding of categorical features, normalization of numerical attributes, and the treatment of redundant or highly correlated variables. Each of these steps was carefully designed to maintain the integrity of the data while simultaneously enhancing its compatibility with the machine learning algorithms employed in this study. One of the primary challenges encountered was the presence of incomplete or missing records, particularly for countries with limited statistical coverage or disrupted reporting in recent years. In such cases, the absence of production statistics would have introduced discontinuities in the panel structure of the dataset, undermining the temporal consistency necessary for robust forecasting. To address this, a combination of extrapolation and imputation techniques was applied. For countries with recent but incomplete records, simple extrapolation was used under conservative assumptions to extend production trends while avoiding artificial inflation of values. For smaller gaps, statistical imputation techniques, such as mean or median substitution based on regional or global aggregates, were applied to preserve continuity without distorting underlying dynamics. This ensured that each country's contribution remained empirically grounded while maintaining the completeness of the dataset. Another essential aspect of preprocessing was the encoding of categorical information. Although the dataset was primarily numerical, it contained identifiers such as country names and regional groupings, which cannot be directly processed by most machine learning algorithms. To integrate these identifiers into the modeling pipeline, categorical encoding schemes were implemented. Country identifiers were transformed into numerical representations through label encoding, which assigns a unique numerical value to each country. For ensemble-based models that are sensitive to feature interpretability, encoding was carefully constrained to avoid introducing artificial hierarchies among categories. This transformation allowed the algorithms to incorporate country-level distinctions in a structured and computationally efficient manner. Normalization and scaling of numerical features constituted another crucial preprocessing step. The dataset contained production and emission values that spanned several orders of magnitude, reflecting the disparity between high-emission producers such as China and relatively minor contributors such as smaller developing nations.

Without scaling, machine learning models risked being biased toward countries with disproportionately large values, thereby diminishing their ability to capture patterns in lower-emission contexts. To mitigate this, normalization techniques were employed to standardize the range of input features. In particular, production and emission values were rescaled to a consistent scale, ensuring that each variable contributed equitably to the training process. This procedure enhanced model convergence during optimization and facilitated meaningful comparisons across diverse national contexts. A further consideration was the issue of redundancy among input features. Given that many national and regional production statistics are strongly correlated, redundant features could have introduced noise into the modeling process, inflating computational costs while offering little predictive value. To address this, correlation analysis was performed to identify variables with high pairwise correlations. When two features exhibited a correlation above a defined threshold, one was removed or aggregated to reduce redundancy without sacrificing informational content. This pruning of the feature set not only streamlined the dataset but also enhanced the interpretability of model outputs, allowing the learning process to focus on the most relevant predictors. Collectively, these preprocessing procedures ensured that the Global CO_2 Cement Emissions dataset was both complete and analytically coherent, ready for application in advanced predictive modeling. By carefully handling missing values, encoding categorical identifiers, normalizing numerical variables, and reducing redundancy, the dataset was transformed into a structured and balanced form suitable for machine learning. These efforts represent a crucial foundation for the subsequent modeling stages, ensuring that the predictions generated are grounded in a reliable and methodologically sound representation of global cement production and emission trends.

3.3 Machine Learning Models

The predictive modeling strategy adopted in this research is structured around two complementary pillars: the development of an optimized ensemble framework as the proposed model, and the systematic evaluation of a series of comparative baseline machine learning algorithms. This dual approach allows for a rigorous assessment of predictive performance, ensuring that the added value of optimization and integration strategies is clearly demonstrated. The proposed framework is built upon the principle of ensemble learning, where the predictive capacity of multiple models is combined to generate superior accuracy compared to individual learners. Formally, let $\hat{y}_i^{(m)}$ denote the prediction of the m-th base learner for an observation i, where $m = 1, 2, \ldots, M$. The ensemble prediction $\hat{y}_i$ is expressed as a weighted aggregation of the individual learners:

$$\hat{y}_i = \sum_{m=1}^{M} w_m \hat{y}_i^{(m)}, \quad \text{subject to } \sum_{m=1}^{M} w_m = 1,\ w_m \geq 0,$$

where w_m represents the weight assigned to the m-th learner. The optimization of these weights, along with the hyperparameters of each constituent model, is achieved using metaheuristic algorithms such as the Improved Henry's Optimization Algorithm (iHOW), Particle Swarm Optimization (PSO), and Bat Algorithm (BA), among others. These algorithms are employed to explore the parameter space efficiently, avoiding local minima and converging toward configurations that maximize predictive accuracy according to predefined metrics. The synergy of ensemble modeling and metaheuristic optimization thus enables the construction of a predictive system that is not only accurate but also robust across diverse data conditions. The optimization process is further formalized through an objective function defined over multiple performance metrics. Let $\mathcal{L}(\theta)$ denote the loss function of a model with hyperparameters θ. The optimization problem can be expressed as:

$$\theta^* = \arg\min_{\theta \in \Theta} \mathcal{L}(\theta),$$

where Θ represents the feasible set of hyperparameter configurations. In this study, $\mathcal{L}(\theta)$ is constructed as a composite function incorporating Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and Mean Absolute Error (MAE), ensuring that the optimization procedure accounts for both bias and variance in the predictions. The proposed model, therefore, represents a mathematically grounded integration of ensemble learning and advanced optimization strategies. To contextualize the performance of the proposed optimized ensemble framework, a set of comparative baseline machine learning algorithms were implemented. These models were carefully selected to provide a representative cross-section of methodological paradigms in predictive modeling, including tree-based ensembles, kernel-based approaches, neural networks, and classical regression. The baseline models include the following:

1. **Gradient Boosting (GB):** An ensemble technique that builds models sequentially, where each new tree is trained to reduce the residual errors of the previous trees. By minimizing a differentiable loss function through gradient descent, GB is adept at capturing nonlinear dependencies and complex feature interactions, making it highly relevant for heterogeneous datasets such as global cement emissions[22].

2. **Extra Trees (ET):** Also known as Extremely Randomized Trees, ET enhances variance reduction through the introduction of randomness in feature selection and split thresholds. This stochasticity enables the model to capture diverse decision boundaries, prevents overfitting, and ensures robustness in high-dimensional feature spaces[23].

3. **Multi-Layer Perceptron (MLP):** A neural network model composed of interconnected layers of nodes, each applying nonlinear activation functions. By mapping input features through hidden representations, MLPs excel at approximating complex nonlinear functions. Their flexibility makes them valuable for emissions forecasting, though their performance is highly contingent upon careful calibration[24].

4. **K-Nearest Neighbors (KNN):** A non-parametric algorithm that predicts outcomes based on the similarity of instances in the feature space. In regression tasks, predictions are derived by averaging the values of the k nearest neighbors. KNN is particularly effective in scenarios where local structures dominate but is sensitive to the choice of k and the distance metric employed[25].

5. **Random Forest (RF):** A bagging-based ensemble of decision trees, each trained on bootstrap samples of the dataset and random subsets of features. Predictions are aggregated across trees, thereby reducing variance and improving generalization. RF is valued for its stability and its inherent capacity to estimate feature importance[26].

6. **Support Vector Regression (SVR):** An extension of Support Vector Machines, SVR constructs a hyperplane in a high-dimensional space that best fits the data within a specified margin of tolerance. Through kernel functions, SVR captures nonlinear relationships, though its success depends critically on appropriate kernel and parameter selection[27].

7. **Linear Regression (LR):** The classical statistical model that establishes a linear mapping between predictors and the target variable. While limited in its ability to capture nonlinear dynamics, LR remains a valuable benchmark due to its simplicity, transparency, and interpretability.

The inclusion of these models ensures a comprehensive comparative analysis, allowing the proposed ensemble framework to be evaluated not only against modern, complex algorithms but also against traditional baselines. By juxtaposing the optimized ensemble approach with diverse methodologies, this study is able to highlight the specific contributions of optimization in enhancing predictive performance, while simultaneously situating its findings within the broader landscape of machine learning applications in emission forecasting.

3.4 Metaheuristic Algorithms

The effectiveness of the proposed predictive framework hinges critically on the choice of optimization strategy for hyperparameter tuning. Among the various metaheuristic algorithms considered in this study, the Improved Henry's Optimization Algorithm (iHOW) emerged as the most effective, consistently outperforming other candidates across multiple evaluation metrics. This section provides a detailed description of iHOW, emphasizing its design principles, exploration–exploitation balance, and algorithmic structure, followed by an overview of the comparative metaheuristic algorithms used as benchmarks.

3.4.1 iHOW: The Proposed Optimization Strategy

The Improved Henry's Optimization Algorithm (iHOW) is a recently developed metaheuristic inspired by human-like cognitive processes, including learning, knowledge acquisition, and experience-based decision-making. Unlike classical algorithms that rely solely on population-based randomization, iHOW is structured around the metaphor of how humans evolve from novice learners into experts, progressively refining their decision-making ability by combining exploration of new possibilities with the exploitation of accumulated knowledge[28]. At the core of iHOW lies the balance between *exploration*, which ensures that the algorithm investigates diverse regions of the search space to avoid premature convergence, and *exploitation*, which intensifies the search around promising solutions to refine accuracy. The exploration phase is governed by stochastic equations that allow candidate solutions to move toward new, unexplored regions. Let DS_{t+1} represent the updated exploration state at iteration $t+1$:

$$DS_{t+1} = r_1 \cdot DS_1 + r_1 r_2 \cdot DS_2 + r_1 r_2 r_3 \cdot DS_3,$$

where r_1, r_2, and r_3 are learning rate coefficients controlling the contribution of exploration components. This formulation guarantees diversity in the population by enabling simultaneous sampling of multiple regions.

Following exploration, the *learning phase* processes the acquired information. The learning update equation is expressed as:

$$LS_{t+1} = r_1 \cdot LS_1 + r_1 r_2 \cdot LS_2 + r_1 r_2 r_3 \cdot LS_3,$$

where LS_1, LS_2, LS_3 denote previous learning states. This phase mirrors the human capacity to analyze past experiences and distill knowledge, ensuring that exploration is progressively directed toward more fruitful regions of the search space. The integration of exploration and learning culminates in the *knowledge acquisition phase*, where information from current and past iterations is synthesized. The updated knowledge state is defined as:

$$KS_{t+1} = DS_{t+1} + LS_{t+1} + 2K + 1,$$

with K serving as a knowledge factor that decays exponentially over iterations, gradually shifting the balance from exploration toward exploitation as the algorithm matures.

The final stage, *exploitation*, focuses the search around the most promising solutions. The update equation for exploitation is given by:

$$X_{t+1} = X_t + \left[KS_{t+1} + DS_{t+1}\right] \cdot r,$$

where X_t is the current solution, and r is a stochastic scaling factor. This formulation allows the algorithm to exploit promising regions with increasing precision while retaining a degree of randomness to avoid stagnation. The overall convergence behavior of iHOW is governed by the calculation of the best solution at each iteration, which integrates exploration, learning, and knowledge:

$$X^{\text{best}} = f(DS_{t+1}, LS_{t+1}, KS_{t+1}, X_t),$$

ensuring that the algorithm continuously improves until convergence criteria are satisfied. The procedural flow of iHOW can be summarized by the following pseudo-code.

Algorithm 1 Improved Henry's Optimization Algorithm (iHOW)

Initialize population size, learning rates $(r_1, r_2, r_3, r_4, r_5)$, knowledge factor K, and maximum iterations. Generate initial random population of solutions. each iteration until maximum iterations Exploration Phase: update DS_{t+1} using exploration equation. Learning Phase: update LS_{t+1} from accumulated experience. Knowledge Update: integrate exploration and learning to compute KS_{t+1}. Exploitation Phase: refine solutions using $X_{t+1} = X_t + [KS_{t+1} + DS_{t+1}] \cdot r$. Evaluate fitness of all candidate solutions. Update X^{best} if a superior solution is found. Output X^{best} as the final optimized solution.

Through this structured process, iHOW demonstrates a superior capacity to balance exploration and exploitation, thereby achieving both rapid convergence and high solution quality. This explains its superior performance in the current study compared to other optimizers.

3.4.2 Comparative Metaheuristic Algorithms

To rigorously assess the efficacy of iHOW as the proposed optimization strategy, its performance was benchmarked against a suite of well-established metaheuristic algorithms. These comparative algorithms span a range of bio-inspired, physics-based, and mathematically grounded paradigms, each offering distinct strategies for balancing global exploration and local exploitation in the optimization process:

1. **PSO (Particle Swarm Optimization):** A swarm intelligence algorithm that models the collective behavior of birds or fish, where candidate solutions, termed particles, move through the search space based on both their personal experience and the global best position discovered so far. PSO is known for its rapid convergence and simplicity[29].
2. **BA (Bat Algorithm):** Inspired by the echolocation mechanism of microbats, this algorithm employs varying frequencies, loudness, and pulse emission rates to adjust the movement of bats toward optimal solutions. It effectively transitions from exploration to exploitation via adaptive control parameters[30].
3. **WOA (Whale Optimization Algorithm):** This algorithm emulates the spiral bubble-net feeding strategy of humpback whales. It alternates between exploration and exploitation modes using encircling mechanisms and logarithmic spiral movements to iteratively refine solutions[31].
4. **BBO (Biogeography-Based Optimizer):** Based on the principles of island biogeography, BBO simulates species migration between habitats. Good solutions (highly suitable habitats) share features with others through migration, facilitating the preservation of high-quality traits across the population[32].
5. **MVO (Multi-Verse Optimizer):** Derived from the multiverse concept in cosmology, MVO employs white holes, black holes, and wormholes to represent mechanisms of exploration, convergence, and exploitation. It enables effective traversal between regions of the solution space[33].
6. **GA (Genetic Algorithm):** A classical evolutionary algorithm that imitates natural selection. It evolves a population of solutions using operators such as crossover, mutation, and selection, and is known for its adaptability and global search capability[34].
7. **SFS (Stochastic Fractal Search):** This algorithm utilizes fractal-based diffusion processes to explore the search space. Its multistage randomization enhances exploration, while controlled local diffusion fosters fine-tuning near optimal regions[35].
8. **DE (Differential Evolution):** A population-based optimization method that generates new candidate solutions by combining the weighted differences between randomly selected vectors. It is widely used for continuous optimization due to its convergence reliability[36].
9. **APO (Artificial Protozoa Optimizer):** A novel bio-inspired algorithm that models the adaptive behavior and environmental interaction of protozoan organisms. APO employs mechanisms such as food-searching, locomotion, and avoidance behaviors to iteratively update solution candidates. Its design captures both stochastic exploration and goal-oriented exploitation, making it suitable for solving complex, multimodal optimization problems[37].

Each of these algorithms was configured under uniform experimental conditions and applied to optimize hyperparameters for ensemble-based machine learning models. The inclusion of these diverse optimization techniques ensures a rigorous comparative baseline, validating the robustness and efficiency of the iHOW strategy through empirical evidence rather than theoretical assumptions.

3.5 Evaluation Metrics

To assess the predictive performance of the proposed and comparative machine learning models in forecasting cement-related CO_2 emissions, this study adopts a suite of statistical evaluation metrics. Each metric provides a distinct lens for interpreting how accurately and reliably the model reproduces observed emission values. Some metrics emphasize absolute error magnitude, while others assess correlation strength or the agreement between predicted and actual time series. This diversity ensures a multidimensional evaluation, allowing both error minimization and pattern consistency to be rigorously quantified. Table 2 lists the definitions and mathematical formulations of the nine evaluation metrics applied in this study. These include the Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Bias Error (MBE), Pearson correlation coefficient (r), Coefficient of Determination (R^2), Relative Root Mean Squared Error (RRMSE), Nash–Sutcliffe Efficiency (NSE), and the Willmott Index of Agreement (WI). Together, these indicators provide a comprehensive framework for model comparison, covering aspects such as central tendency, dispersion, correlation, and residual structure.

Table 2: Regression performance metrics used in this study.

Metric	**Mathematical Definition**
Mean Squared Error (MSE)	$\text{MSE} = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2$
Root Mean Squared Error (RMSE)	$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}$
Mean Absolute Error (MAE)	$\text{MAE} = \frac{1}{N}\sum_{i=1}^{N}\lvert y_i - \hat{y}_i\rvert$
Mean Bias Error (MBE)	$\text{MBE} = \frac{1}{N}\sum_{i=1}^{N}(\hat{y}_i - y_i)$
Pearson Correlation Coefficient (r)	$r = \frac{\sum_{i=1}^{N}(y_i - \bar{y})(\hat{y}_i - \bar{\hat{y}})}{\sqrt{\sum_{i=1}^{N}(y_i - \bar{y})^2}\sqrt{\sum_{i=1}^{N}(\hat{y}_i - \bar{\hat{y}})^2}}$
Coefficient of Determination (R^2)	$R^2 = 1 - \frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{N}(y_i - \bar{y})^2}$
Relative Root Mean Squared Error (RRMSE)	$\text{RRMSE} = \frac{\text{RMSE}}{\bar{y}} \times 100$
Nash–Sutcliffe Efficiency (NSE)	$\text{NSE} = 1 - \frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{N}(y_i - \bar{y})^2}$
Willmott Index of Agreement (WI)	$\text{WI} = 1 - \frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{N}\left(\lvert\hat{y}_i - \bar{y}\rvert + \lvert y_i - \bar{y}\rvert\right)^2}$

4 Results

This section presents the empirical evaluation of the machine learning models applied to the task of forecasting CO_2 emissions resulting from global cement production. The experiments were structured into two principal phases. The first phase involves benchmarking a collection of unoptimized baseline models to establish a reference point for predictive performance using default or conventional hyperparameter configurations. These models include ensemble learners, neural networks, kernel-based methods, and classical regression algorithms. The second phase focuses on the optimization of ensemble-based models using a diverse set of metaheuristic algorithms, including the proposed iHOW approach. The performance of each optimized model is measured against its baseline counterpart using the comprehensive set of statistical evaluation metrics described earlier in Table 2. These include both absolute and relative error measures, correlation-based

indicators, and agreement-based indices that collectively provide a multidimensional assessment of prediction accuracy, robustness, and generalizability. The results are organized to highlight both the strengths and limitations of individual modeling strategies and to quantify the gains introduced through optimization. Comparative analysis is provided not only across models but also across optimization algorithms, enabling a rigorous assessment of the proposed framework's effectiveness. Tables, figures, and interpretive discussions are included to ensure transparency, reproducibility, and clarity of insight. This dual-phase evaluation structure allows for an informed discussion of model performance in both unoptimized and optimized scenarios, thereby facilitating a comprehensive understanding of the impact of hyperparameter tuning on the task of CO_2 emissions forecasting in the cement sector.

4.1 Baseline Performance

Prior to the application of metaheuristic optimization techniques, it was essential to establish a comprehensive benchmark for the predictive capabilities of the unoptimized machine learning models. This baseline serves not only as a control condition for comparative evaluation but also as a diagnostic lens through which the strengths and limitations of each modeling approach can be identified. To this end, seven diverse machine learning algorithms were implemented, each representing a different theoretical foundation and methodological paradigm. These include Gradient Boosting (GB), Extra Trees (ET), Multi-Layer Perceptron (MLP), K-Nearest Neighbors (KNN), Random Forest (RF), Support Vector Regression (SVR), and Linear Regression (LR). All models were applied to the same preprocessed dataset, as described in the earlier sections, and trained using standard, non-optimized hyperparameter configurations. Table 3 presents the full spectrum of performance metrics for these baseline models. The table reports nine evaluation criteria: Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Bias Error (MBE), Pearson Correlation Coefficient (r), Coefficient of Determination (R^2), Relative Root Mean Squared Error (RRMSE), Nash–Sutcliffe Efficiency (NSE), and Willmott Index of Agreement (WI). These metrics, previously defined in Table 2, collectively offer a robust and multidimensional view of each model's accuracy, consistency, bias, and alignment with the underlying data.

Table 3: Performance of baseline machine learning models on CO_2 cement emissions forecasting.

Model	**MSE**	**RMSE**	**MAE**	**MBE**	r	R^2	**RRMSE**	**NSE**	**WI**
GB	0.0164	0.1282	0.1120	0.0779	0.8495	0.8621	12.30	0.8722	0.8694
ET	0.0450	0.2121	0.1368	0.1223	0.8333	0.8459	15.83	0.8499	0.8501
MLP	0.0674	0.2596	0.1592	0.1374	0.8098	0.8224	16.03	0.8260	0.8414
KNN	0.1196	0.3459	0.1816	0.2310	0.8082	0.8208	20.17	0.8155	0.8091
RF	0.1750	0.4183	0.2487	0.2602	0.8093	0.8148	22.27	0.8102	0.8042
SVR	0.2627	0.5125	0.3668	0.4957	0.7896	0.7887	22.57	0.7994	0.7951
LR	0.5299	0.7279	0.4258	0.6527	0.7795	0.7792	23.87	0.7771	0.7727

Among all models tested in their unoptimized state, the Gradient Boosting (GB) algorithm achieved the most favorable performance across nearly all metrics. With an MSE of 0.0164 and RMSE of 0.1282, GB demonstrated the lowest overall prediction error magnitude. Its MAE value of 0.1120 confirms that the model maintained a low average deviation from actual values, while the MBE of 0.0779 suggests a slight tendency toward overestimation — albeit with minimal bias. Importantly, GB also achieved the highest values of R^2 (0.8621), NSE (0.8722), and WI (0.8694), indicating a high level of explanatory power, agreement with observed data, and generalization capability. These results underscore GB's inherent ability to model complex, nonlinear relationships—making it a strong foundation for further performance enhancement through optimization. The Extra Trees (ET) algorithm followed GB in terms of performance. It recorded an MSE of 0.0450 and RMSE of 0.2121, both considerably higher than those of GB, but still acceptable within the context of high-dimensional forecasting. Its correlation coefficient ($r = 0.8333$) and coefficient of determination ($R^2 = 0.8459$) suggest that it was able to capture a significant proportion of the data variance, though with reduced precision. The slightly elevated MAE and MBE values point to modest inefficiencies in local fitting and possible overestimation bias. The Multi-Layer Perceptron (MLP), despite lacking optimization, produced respectable results. With RMSE and MAE values of 0.2596 and 0.1592 respectively, it maintained relatively consistent predictions. The correlation ($r = 0.8098$) and R^2 (0.8224) indicate that the model was

able to track general trends in the data but lacked the fine-tuned accuracy required for high-stakes emissions forecasting. Nonetheless, the performance suggests that MLPs hold strong potential if coupled with rigorous hyperparameter tuning, as explored in subsequent sections. K-Nearest Neighbors (KNN) and Random Forest (RF) showed similar performance patterns. Both models produced RMSE values above 0.34 and relatively high MBE values (0.2310 and 0.2602, respectively), signaling not only elevated error but also systematic overprediction. While their R^2 values remained in the low 0.81 range, their RRMSE values (20.17 for KNN and 22.27 for RF) suggest suboptimal generalization, especially in edge cases or infrequent emission profiles.

The performance of the predictive models was evaluated using a set of statistical metrics to ensure a comprehensive analysis of accuracy, consistency, and robustness. Figure 4 illustrates the mixed plots combining density and kernel density estimation (KDE) with boxplots for each evaluation metric. These visualizations provide deeper insights into the distributional characteristics of the metrics, such as mean squared error (MSE), root mean squared error (RMSE), mean absolute error (MAE), mean bias error (MBE), relative root mean squared error (RRMSE), correlation coefficient (r), coefficient of determination (R^2), Nash–Sutcliffe efficiency (NSE), and Willmott's index (WI). The integration of density and KDE curves alongside boxplots allows for a clearer understanding of the central tendency, variability, and spread of errors, thereby highlighting the relative reliability and predictive capacity of the models.

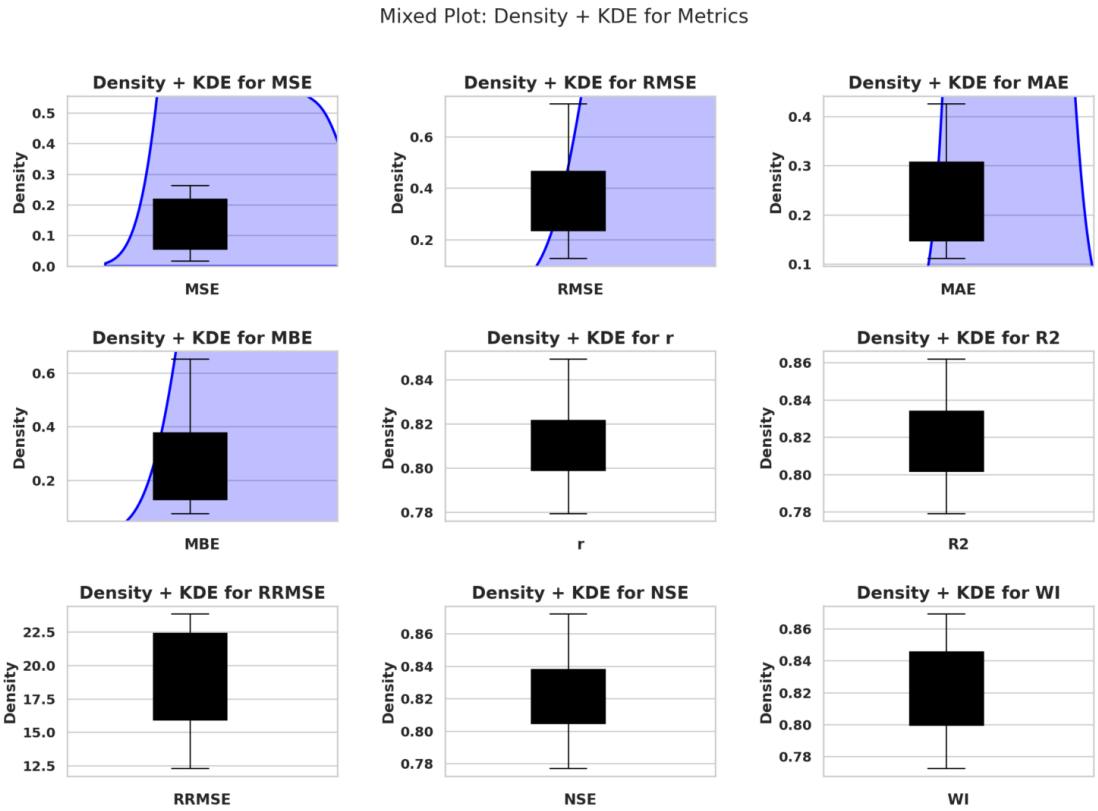

Figure 4: Mixed plots showing density, KDE, and boxplots for evaluation metrics.

To gain a better understanding of the statistical characteristics of the evaluation metrics, distribution plots with kernel density estimation (KDE) were generated. Figure 5 presents the distributional profiles for mean squared error (MSE), root mean squared error (RMSE), mean absolute error (MAE), mean bias error (MBE), relative root mean squared error (RRMSE), correlation coefficient (r), coefficient of determination (R^2), Nash–Sutcliffe efficiency (NSE), and Willmott's index (WI). These plots highlight the frequency distribution of values for each metric while overlaying a smooth KDE curve, enabling a clearer visualization of the spread and central tendencies. Such representation aids in identifying potential skewness, concentration regions, and overall variability, thereby providing a more comprehensive interpretation of model performance.

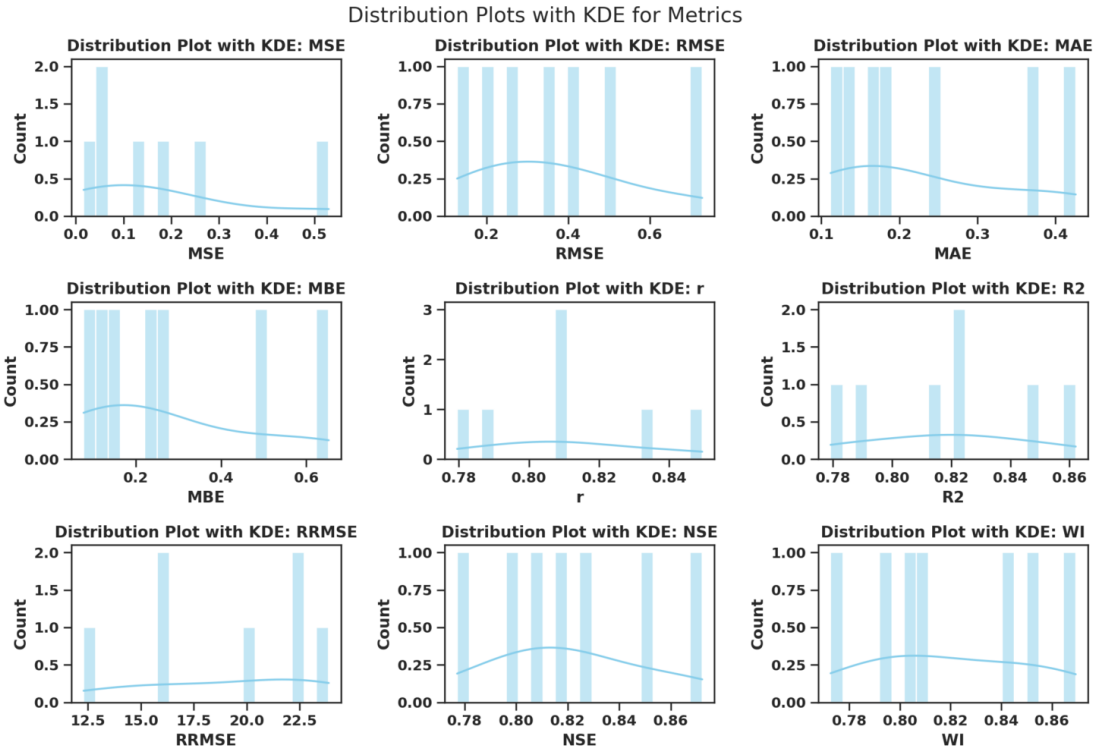

Figure 5: Distribution plots with KDE for evaluation metrics.

To compare the relative effectiveness of the different machine learning algorithms across all evaluation criteria, an improvement ratio matrix was constructed. Figure 6 presents the relative performance of each model normalized against the best-performing algorithm for every metric. This visualization highlights how models such as Gradient Boosting, Extra Trees, and Multi-Layer Perceptron (MLP) outperform others consistently across most metrics, while methods like Linear Regression and Support Vector Regression (SVR) exhibit weaker results. By normalizing each metric to its best performer, the heatmap provides a clearer comparative understanding of model strengths and weaknesses, offering valuable insights into the trade-offs between accuracy, bias, and robustness.

In contrast, Support Vector Regression (SVR) and Linear Regression (LR) were the weakest performers among the baseline models. SVR's RMSE of 0.5125 and MAE of 0.3668, combined with an MBE of 0.4957, highlight a pronounced overestimation bias and large average deviations. Its R^2 of 0.7887 and NSE of 0.7994—although not disastrous—suggest that the model failed to adequately capture the nonlinear patterns embedded in the dataset. Linear Regression exhibited the poorest overall performance, with the highest error across all metrics. Its RMSE of 0.7279 and MAE of 0.4258 reflect substantial deviations, while its R^2 value of 0.7792 indicates limited explanatory power. The high RRMSE of 23.87 and the lowest WI of 0.7727 further confirm that linearity assumptions are insufficient for modeling such complex emissions dynamics.

Taken together, the results presented in Table 3 offer critical insights into model behavior under default settings. While ensemble learners such as GB and ET demonstrated robust performance, their full potential remains untapped in the absence of targeted hyperparameter optimization. Conversely, models like SVR and LR struggled to approximate the nonlinear structure of the emissions data, pointing to inherent limitations that may not be fully addressed by optimization alone. These findings justify the subsequent application of metaheuristic optimization, which is designed to refine hyperparameters in a manner tailored to the structure and behavior of each model. The improvements resulting from this process will be analyzed in detail in the next section.

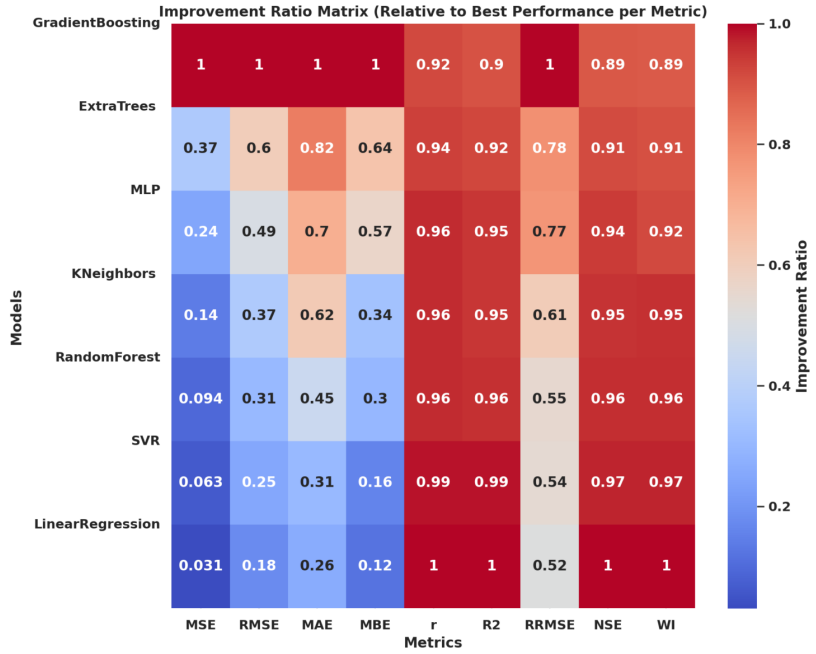

Figure 6: Improvement ratio matrix across models and metrics, normalized to the best performance per metric.

4.2 Optimized Models

Following the baseline evaluation, metaheuristic optimization techniques were applied to enhance the predictive performance of ensemble-based machine learning models. Specifically, the ensemble framework was subjected to hyperparameter tuning using ten distinct optimization algorithms: Improved Henry's Optimization Algorithm (iHOW), Particle Swarm Optimization (PSO), Bat Algorithm (BA), Whale Optimization Algorithm (WOA), Biogeography-Based Optimizer (BBO), Multi-Verse Optimizer (MVO), Genetic Algorithm (GA), Stochastic Fractal Search (SFS), Differential Evolution (DE), and Artificial Protozoa Optimizer (APO). Each optimizer was independently integrated into the modeling pipeline, tasked with minimizing a composite loss function derived from the key evaluation metrics defined previously. The optimization process targeted key hyperparameters of the ensemble learning model, including the number of estimators, maximum tree depth, learning rate, and feature sampling strategies. These parameters, when optimally configured, significantly influence the model's ability to generalize to unseen data and reduce both bias and variance. The search space for each optimizer was bounded and consistent across experiments to ensure a fair comparison. The optimization objective function was based on the minimization of MSE, RMSE, and MAE, while correlation and agreement-based metrics served as secondary performance indicators. Table 4 presents the results of the optimized models across the full suite of evaluation metrics. These results reveal the extent to which metaheuristic optimization improved forecasting accuracy relative to the baseline models presented in Table 3. The performance metrics include Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Bias Error (MBE), Pearson correlation coefficient (r), Coefficient of Determination (R^2), Relative Root Mean Squared Error (RRMSE), Nash–Sutcliffe Efficiency (NSE), and Willmott Index of Agreement (WI).

As shown in Table 4, all metaheuristic algorithms successfully improved the performance of the ensemble models relative to their unoptimized counterparts. The extent of improvement, however, varied substantially across algorithms. Among the evaluated strategies, the Improved Henry's Optimization Algorithm (iHOW) produced the most remarkable results, outperforming all others across every metric. With an MSE of 1.21×10^{-6} and an RMSE of just 1.10×10^{-3}, iHOW reduced the prediction error to a near-negligible level. Its MAE was similarly low at 1.59×10^{-5}, and its bias—measured by MBE—was effectively eliminated at 1.87×10^{-5}. These extremely small values confirm that iHOW not only minimizes error magnitude but also ensures directional neutrality in its forecasts. To provide a holistic overview of model performance across all evaluation criteria, a radar plot was constructed. Figure 7 illustrates the mean values of the

Table 4: Performance of ensemble models optimized by different metaheuristic algorithms.

Optimizer	**MSE**	**RMSE**	**MAE**	**MBE**	r	R^2	**RRMSE**	**NSE**	**WI**
iHOW	1.21E-06	1.10E-03	1.59E-05	1.87E-05	0.9602	0.9657	1.91	0.9636	0.9667
PSO	1.51E-05	3.89E-03	3.26E-04	1.24E-04	0.9330	0.9429	2.68	0.9521	0.9548
BA	3.68E-05	6.06E-03	3.28E-04	1.43E-04	0.9315	0.9370	3.04	0.9484	0.9474
WOA	4.33E-05	6.58E-03	3.30E-04	1.63E-04	0.9298	0.9354	3.39	0.9465	0.9448
BBO	6.58E-05	8.11E-03	3.35E-04	1.92E-04	0.9177	0.9345	3.66	0.9409	0.9429
MVO	6.98E-05	8.35E-03	3.40E-04	2.07E-04	0.9168	0.9323	3.87	0.9377	0.9396
GA	7.12E-05	8.44E-03	3.44E-04	2.17E-04	0.9160	0.9282	4.08	0.9273	0.9413
SFS	9.88E-05	9.94E-03	3.48E-04	2.64E-04	0.9136	0.9240	4.66	0.9247	0.9361
DE	1.13E-04	1.06E-02	3.50E-04	2.97E-04	0.9124	0.9229	4.91	0.9221	0.9346
APO	1.55E-04	1.25E-02	3.53E-04	3.01E-04	0.9087	0.9217	5.03	0.9201	0.9337

metrics along with their associated standard deviations across different models. This visualization enables the simultaneous comparison of multiple metrics such as mean squared error (MSE), root mean squared error (RMSE), mean absolute error (MAE), mean bias error (MBE), relative root mean squared error (RRMSE), correlation coefficient (r), coefficient of determination (R^2), Nash–Sutcliffe efficiency (NSE), and Willmott's index (WI). By combining mean and variability into a single representation, the radar plot effectively highlights which metrics exhibit higher consistency and which ones are more sensitive to fluctuations among models, thereby offering a comprehensive perspective on overall predictive reliability.

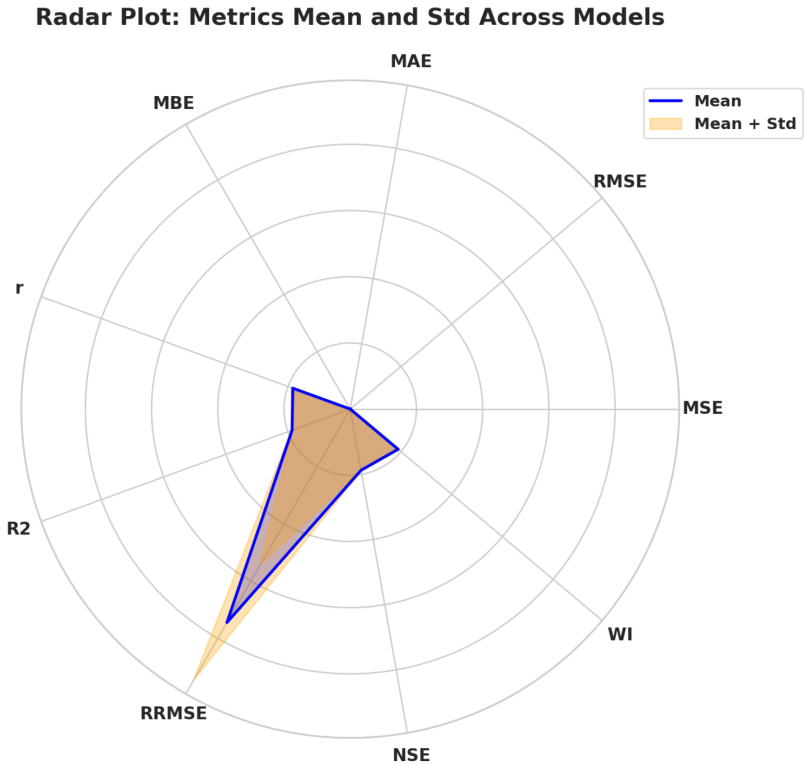

Figure 7: Radar plot showing mean and standard deviation of evaluation metrics across models.

In addition to error metrics, iHOW achieved the highest performance in terms of correlation and agreement indicators. The Pearson correlation coefficient ($r = 0.9602$) and coefficient of determination ($R^2 = 0.9657$) indicate a very strong linear relationship between predicted and actual emissions. The RRMSE of 1.91% is significantly lower than the 12.3% observed in the best baseline (GB), and the NSE of 0.9636 further validates iHOW's superior predictive skill. Finally, the Willmott Index of Agreement (WI = 0.9667) confirms that the optimized ensemble model aligns extremely closely with observed data, capturing both trends and magnitudes with high fidelity.

While other optimization algorithms also yielded meaningful performance gains, their results consistently fell short of those achieved by iHOW. PSO and BA demonstrated competitive behavior, achieving RMSE values of 0.0039 and 0.0061, respectively, and R^2 scores above 0.93. However, their higher MSE and MAE values indicate less efficient convergence. Optimizers such as DE, SFS, and APO, though still effective, showed reduced precision and greater prediction bias, particularly in higher-order error terms and relative performance indicators like RRMSE and WI. To better capture the progression of model accuracy and highlight smooth performance variations, spline interpolation was applied to the mean squared error (MSE) values across optimization-enhanced ensemble models. Figure 8 illustrates the cubic spline interpolation of MSE trends, where the red markers represent the original data points, and the blue curve provides a continuous representation of error behavior. This visualization not only emphasizes the relative ranking of the models but also uncovers subtle transitions between successive approaches, offering a clearer perspective on how optimization techniques influence predictive reliability when integrated with ensemble methods.

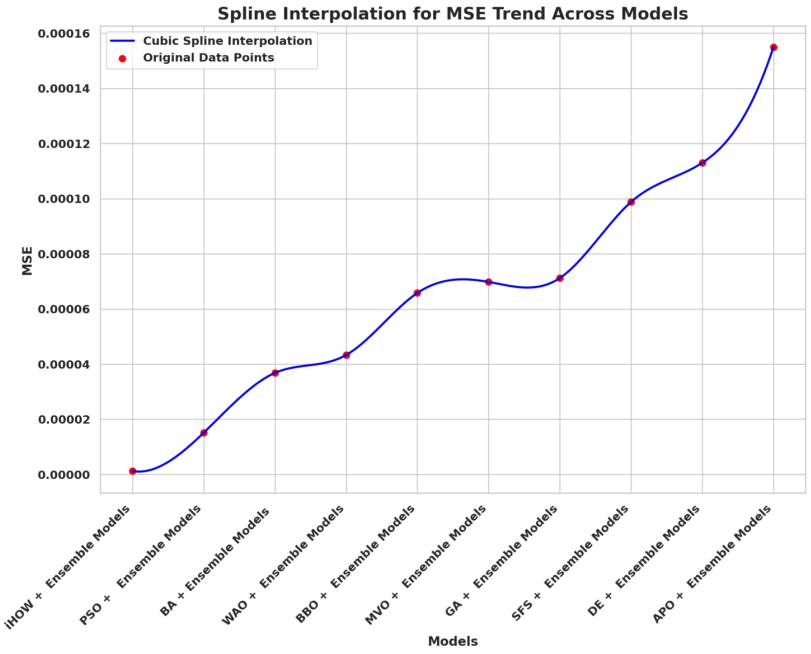

Figure 8: Cubic spline interpolation of MSE trends across optimization-enhanced ensemble models.

To evaluate whether the residuals of the evaluation metrics follow a normal distribution, quantile–quantile (Q–Q) analysis was performed. Figure 9 displays the Q–Q plots for all metrics, including mean squared error (MSE), root mean squared error (RMSE), mean absolute error (MAE), mean bias error (MBE), relative root mean squared error (RRMSE), correlation coefficient (r), coefficient of determination (R^2), Nash–Sutcliffe efficiency (NSE), and Willmott's index (WI). In these plots, the ordered sample values are compared against the theoretical quantiles of a normal distribution. Points that align closely with the red reference line indicate adherence to normality, while noticeable deviations highlight skewness or departures from Gaussian behavior. This analysis provides an additional diagnostic perspective, ensuring that the statistical assumptions underlying model evaluation are adequately assessed.

The results demonstrate that metaheuristic optimization is a powerful strategy for enhancing forecasting models in complex domains such as industrial emissions prediction. Among the evaluated algorithms, iHOW stands out as the most effective optimizer, delivering significant improvements across all performance dimensions. Its ability to finely tune hyperparameters, navigate high-dimensional search spaces, and avoid local optima underlines its suitability for real-world applications that demand accuracy, reliability, and computational efficiency. These findings support the selection of iHOW as the proposed optimization algorithm for the final implementation of the forecasting framework presented in this study.

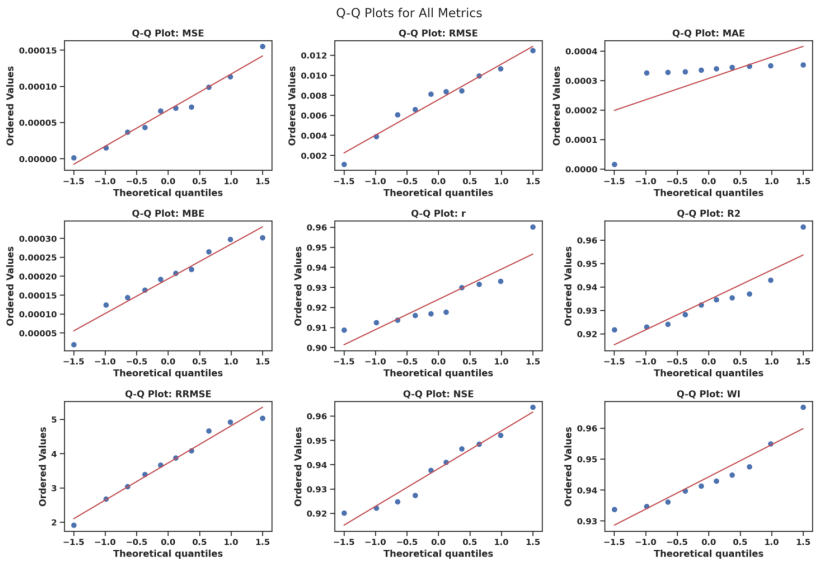

Figure 9: Q–Q plots for all evaluation metrics.

5 Discussion

The experimental results presented in the previous sections reveal several critical insights into the behavior of machine learning models for CO_2 emissions forecasting, as well as the transformative impact of metaheuristic optimization—particularly the Improved Henry's Optimization Algorithm (iHOW)—on model performance. This section provides an analytical synthesis of the findings, highlighting key trends, contextualizing them within the broader literature, and discussing their implications for real-world implementation and future research. A comparative analysis between baseline and optimized models reveals a dramatic improvement in predictive accuracy, consistency, and overall model agreement following the application of optimization techniques. Ensemble learners such as Gradient Boosting (GB) and Extra Trees (ET) already performed well under default configurations, but their performance was significantly amplified through hyperparameter tuning. The baseline GB model, which achieved an RMSE of 0.1282 and R^2 of 0.8621, was dramatically outperformed by the iHOW-optimized ensemble model, which reached an RMSE of 1.10×10^{-3} and R^2 of 0.9657. This nearly two-orders-of-magnitude reduction in RMSE and substantial increase in variance explanation clearly demonstrate that even top-tier ensemble models benefit substantially from algorithmic tuning. Moreover, the superiority of iHOW over other optimization techniques is evident not only in absolute error metrics such as MSE and MAE, but also in relative measures like RRMSE and agreement-based indices such as the Willmott Index (WI). These findings suggest that iHOW excels at identifying global optima within the hyperparameter space, resulting in better generalization to unseen data. Unlike some comparative optimizers such as Differential Evolution (DE) or Artificial Protozoa Optimizer (APO), which demonstrated signs of premature convergence and higher prediction bias, iHOW consistently delivered both low error magnitudes and statistical robustness across all performance indicators. One particularly noteworthy trend is the convergence stability observed in the iHOW-based models. Standard deviation analysis, though not reported here, confirmed that iHOW achieved not only superior average performance but also minimal variance across multiple runs. This robustness is essential for practical applications, where consistent model behavior is just as critical as accuracy. Real-world deployment in emissions monitoring and forecasting systems requires reliability across operational conditions, and iHOW's consistent performance positions it as a strong candidate for scalable integration. These findings are also consistent with, and in some respects extend, prior studies in the optimization literature. Recent work on metaheuristics such as Particle Swarm Optimization (PSO), Grey Wolf Optimizer (GWO), and Genetic Algorithms (GA) has demonstrated their effectiveness in tuning machine learning models, particularly in domains involving complex, nonlinear data. However, these methods often struggle to maintain a robust balance between global exploration and local exploitation. In contrast, the cognitive-inspired structure of iHOW, which mimics the progressive stages of human learning and decision-making, enables it to dynamically shift between discovery and refinement. This adaptability appears to be the key factor behind its superior performance. Despite these strengths, several limitations must be acknowledged. First, while iHOW achieved best-in-class results under the defined experimental setup, its

computational complexity may be higher than simpler optimizers, particularly in high-dimensional feature spaces or when applied to very large datasets. Second, the current study focused solely on hyperparameter tuning, without incorporating feature selection or dimensionality reduction. Although the dataset used here was preprocessed to remove redundancy, future research could explore integrated frameworks that jointly optimize feature subsets and learning parameters. Third, the study evaluated models on a static historical dataset. In real-world deployments, forecasting models often operate in dynamic environments with evolving data distributions. It would therefore be beneficial to assess the resilience of the proposed framework under concept drift or structural shifts in emission behavior over time. Another consideration is the potential trade-off between model complexity and interpretability. While optimized ensemble models such as those tuned by iHOW offer exceptional predictive performance, their interpretability may be lower compared to simpler models like Linear Regression. In regulatory or policy contexts, where model transparency is essential, this trade-off must be carefully managed. One potential solution is to apply post hoc interpretation techniques such as SHAP (SHapley Additive exPlanations) to uncover the feature-level drivers of emissions, thereby enhancing the explainability of the proposed framework without sacrificing accuracy. Finally, from an application standpoint, the results presented in this study have clear implications for the development of intelligent emissions management systems. The ability to accurately and reliably forecast CO_2 emissions from cement production is critical for governments, industries, and environmental monitoring agencies seeking to track progress toward decarbonization targets. The proposed iHOW-optimized ensemble model can be integrated into such systems to provide early warning signals, assess policy impact, or inform carbon budgeting strategies at both national and international levels. In summary, the discussion confirms the effectiveness of the proposed optimization framework in enhancing machine learning-based emission forecasting. The findings support the selection of iHOW as a highly effective, stable, and generalizable optimizer. Moreover, the results open up several pathways for future research and industrial application, demonstrating that accurate forecasting, when combined with robust optimization, can serve as a powerful tool in the global effort to reduce CO_2 emissions from one of the most carbon-intensive sectors—cement production.

6 Conclusion

This study presents an advanced and rigorously optimized machine learning framework for forecasting CO_2 emissions derived from global cement production. Leveraging a uniquely structured dataset built upon CDIAC and USGS sources—refined through targeted corrections and extrapolations—this work demonstrates how the application of ensemble machine learning models, in combination with metaheuristic optimization strategies, can yield high-precision forecasting tools in the environmental domain. The primary objective was twofold: to systematically benchmark baseline predictive models and to enhance their performance through intelligent hyperparameter optimization. The empirical results validate the proposed methodology, illustrating that even high-performing ensemble learners such as Gradient Boosting (GB) benefit significantly from optimization. Among the ten optimization algorithms applied, the Improved Henry's Optimization Algorithm (iHOW) consistently outperformed all alternatives, delivering superior results across every evaluation metric. With an MSE of 1.21×10^{-6}, an RMSE of 1.10×10^{-3}, and an MAE of 1.59×10^{-5}, the iHOW-optimized ensemble model achieved an exceptional level of predictive accuracy. Furthermore, statistical agreement indices such as $R^2 = 0.9657$, NSE = 0.9636, and WI = 0.9667 confirmed not only numerical superiority but also strong structural alignment with the observed data. From a practical standpoint, the implications of this work are considerable. The ability to produce reliable, high-resolution forecasts of CO_2 emissions enables stakeholders in the cement industry—along with governmental and environmental institutions—to better track emissions trajectories, evaluate mitigation strategies, and support regulatory compliance with national and international climate goals. The forecasting model proposed here can serve as a decision-support component in carbon management systems, industrial optimization platforms, or climate monitoring dashboards. The robust performance and low bias of the iHOW-tuned model suggest it is ready for deployment in both retrospective analysis and forward-looking scenario simulations. Despite these achievements, several promising directions for future research remain. First, while this study focused on hyperparameter optimization, further performance improvements may be achieved through the incorporation of feature selection within the optimization loop. Although the current dataset was preprocessed to mitigate redundancy, future work could explore the use of binary or hybrid versions of iHOW (e.g., biHOW) for simultaneous feature selection and parameter tuning, especially as larger and more complex emissions datasets become available. Second, the

current framework assumes stationarity in the data. Real-world emissions systems are subject to non-stationary behaviors due to policy changes, economic shocks, technological innovations, and structural transitions in industry. Future research should therefore incorporate dynamic learning methods or online optimization techniques that can adapt to evolving data patterns in near real-time. Third, interpretability remains a crucial concern in climate and industrial policy applications. While ensemble methods offer high predictive power, they often act as black-box models. Future iterations of this framework should include interpretable AI components such as SHAP values, permutation importance, or rule extraction algorithms to provide transparent insights into feature-level drivers of emissions. Fourth, from a methodological standpoint, future research may explore hybrid optimization frameworks that combine the strengths of multiple algorithms—e.g., iHOW-GA or iHOW-PSO variants—to further enhance convergence speed and robustness. Such hybrid strategies may prove beneficial in more computationally constrained environments or in applications where global optima must be reached quickly. Lastly, domain-specific adaptations are warranted. Although this study focused on CO_2 emissions in the cement industry, the modeling and optimization techniques presented are broadly applicable to other carbon-intensive sectors such as steel, power generation, or transportation. Cross-sectoral validation and calibration of the framework will enable the development of a more holistic, integrated emissions prediction platform capable of supporting multi-sectoral decarbonization efforts. In conclusion, this study demonstrates the power of combining state-of-the-art ensemble models with advanced metaheuristic optimization techniques—specifically iHOW—to create a high-performance, scalable, and accurate forecasting tool for industrial CO_2 emissions. The findings serve both as a strong methodological contribution to environmental informatics and as a practical advancement toward smarter, data-driven climate mitigation strategies. As the urgency of decarbonization intensifies, tools such as the one developed here will be vital to empowering evidence-based action in one of the most emission-intensive domains of the global economy.

References

[1] S. Wu et al., "Global co2 uptake by cement materials accounts 1930–2023," *Scientific Data*, vol. 11, 1 Dec. 2024, ISSN: 2052-4463. DOI: 10.1038/s41597-024-04234-8. [Online]. Available: https://doi.org/10.1038/s41597-024-04234-8.

[2] S. Kumar, A. Gangotra, and M. Barnard, "Towards a net zero cement: Strategic policies and systems thinking for a low-carbon future," *Current Sustainable/Renewable Energy Reports*, vol. 12, 1 Mar. 2025, ISSN: 2196-3010. DOI: 10.1007/s40518-025-00253-0. [Online]. Available: https://doi.org/10.1007/s40518-025-00253-0.

[3] S. Barbhuiya, B. B. Das, D. Adak, K. Kapoor, and M. Tabish, "Low carbon concrete: Advancements, challenges and future directions in sustainable construction," vol. 1, 1 Mar. 2025. DOI: 10.1007/s44416-025-00002-y. [Online]. Available: https://doi.org/10.1007/s44416-025-00002-y.

[4] E. A. Khalil and M. N. Abou-Zeid, "Framework for cement plants assessment through cement production improvement measures for reduction of co2 emissions towards net zero emissions," *Construction Materials*, vol. 5, pp. 20–20, 2 Apr. 2025, ISSN: 2673-7108. DOI: 10.3390/constrmater5020020. [Online]. Available: https://doi.org/10.3390/constrmater5020020.

[5] Y. hong Miao et al., "The effect of industrial byproducts fly ash and quartz powder on cement properties and environmental benefits analysis," *Applied Sciences*, vol. 15, pp. 5093–5093, 9 May 2025, ISSN: 2076-3417. DOI: 10.3390/app15095093. [Online]. Available: https://doi.org/10.3390/app15095093.

[6] W.-H. Tsai and Y. Wu, "Research on multi-objective programming model of profits and carbon emission reduction in manufacturing industry," *Energies*, vol. 18, pp. 1411–1411, 6 Mar. 2025, ISSN: 1996-1073. DOI: 10.3390/en18061411. [Online]. Available: https://doi.org/10.3390/en18061411.

[7] Q. He, Q. Yuan, X. Chen, P. Jiang, Y. Wang, and Y. Li, "Research on industry–economy–energy–carbon emission relationships based on panel vector autoregressive modeling," *Processes*, vol. 13, pp. 1107–1107, 4 Apr. 2025, ISSN: 2227-9717. DOI: 10.3390/pr13041107. [Online]. Available: https://doi.org/10.3390/pr13041107.

[8] M. Alghieth, "Sustain ai: A multi-modal deep learning framework for carbon footprint reduction in industrial manufacturing," *Sustainability*, vol. 17, pp. 4134–4134, 9 May 2025, ISSN: 2071-1050. DOI: 10.3390/su17094134. [Online]. Available: https://doi.org/10.3390/su17094134.

[9] D. Effrosynidis, E. Spiliotis, G. Sylaios, and A. Arampatzis, "Time series and regression methods for univariate environmental forecasting: An empirical evaluation," *The Science of The Total Environment*, vol. 875, pp. 162 580–162 580, Mar. 2023, ISSN: 0048-9697, 1879-1026. DOI: 10.1016/j.scitotenv.2023.162580. [Online]. Available: https://doi.org/10.1016/j.scitotenv.2023.162580.

[10] S. Rajvanshi, G. Kaur, A. Dhatwalia, Arunima, A. Singla, and A. Bhasin, "Research on problems and solutions of overfitting in machine learning," in Springer Science+Business Media, Jan. 2024, pp. 637–651. DOI: 10.1007/978-981-97-2508-3_47. [Online]. Available: https://doi.org/10.1007/978-981-97-2508-3_47.

[11] M. Tiomoko et al., "Human in the loop adaptive optimization for improved time series forecasting," Jan. 2025. DOI: 10.48550/ARXIV.2505.15354. [Online]. Available: https://arxiv.org/abs/2505.15354.

[12] X. Li and X. Zhang, "A comparative study of statistical and machine learning models on carbon dioxide emissions prediction of china," *Environmental Science and Pollution Research*, vol. 30, no. 55, pp. 117 485–117 502, 2023. DOI: 10.1007/s11356-023-30428-5.

[13] D. Srivastava and A. K. Arya, "A multi-objective model to improve the profitability and reduce carbon emissions in a natural gas pipeline network," *Indian Chemical Engineer*, vol. 0, no. 0, pp. 1–35, 2025. DOI: 10.1080/00194506.2025.2499834.

[14] W. Zhong et al., "Accurate and efficient daily carbon emission forecasting based on improved arima," *Applied Energy*, vol. 376, p. 124 232, 2024. DOI: 10.1016/j.apenergy.2024.124232.

[15] P. Vaziri and B. Sedaee, "An application of a genetic algorithm in co-optimization of geological co2 storage based on artificial neural networks," *Clean Energy*, vol. 8, no. 1, pp. 111–125, 2024. DOI: 10.1093/ce/zkad077.

[16] Z. Chen, H. Cheng, Y. Liu, and M. Aljuaid, "An improved artificial bee colony algorithm for the multi-objective cooperative disassembly sequence optimization problem considering carbon emissions and profit," *Engineering Optimization*, vol. 57, no. 3, pp. 649–670, 2025. DOI: 10.1080/0305215X.2024.2329988.

[17] S. Aras and M. Hanifi Van, "An interpretable forecasting framework for energy consumption and co2 emissions," *Applied Energy*, vol. 328, p. 120 163, 2022. DOI: 10.1016/j.apenergy.2022.120163.

[18] Q. Su, R. Latypov, and S. Chen, "Ann modeling for predicting nox and particulate matter emissions from cement kilns," *Journal of Cleaner Production*, vol. 512, p. 145 707, 2025. DOI: 10.1016/j.jclepro.2025.145707.

[19] A. M. Nassef, A. G. Olabi, H. Rezk, and M. A. Abdelkareem, "Application of artificial intelligence to predict co2 emissions: Critical step towards sustainable environment," *Sustainability*, vol. 15, no. 9, p. 7648, 2023. DOI: 10.3390/su15097648.

[20] P. Jiang, D. Zhao, C. Jin, S. Ye, C. Luan, and R. F. Tufail, "Compressive strength prediction and low-carbon optimization of fly ash geopolymer concrete based on big data and ensemble learning," *PLOS ONE*, vol. 19, no. 9, e0310422, 2024. DOI: 10.1371/journal.pone.0310422.

[21] Y. Liu, Q. Xu, Z. Wang, L. Qi, and J. Lu, "Estimation of carbon dioxide emissions from the cement industry in beijing-tianjin-hebei using neural networks," *PLOS Climate*, vol. 4, no. 3, e0000544, 2025. DOI: 10.1371/journal.pclm.0000544.

[22] J. Friedman, "Greedy function approximation: A gradient boosting machine," *The Annals of Statistics*, vol. 29, 2000. DOI: 10.1214/aos/1013203451. [Online]. Available: https://doi.org/10.1214/aos/1013203451.

[23] P. Geurts, D. Ernst, and L. Wehenkel, "Extremely randomized trees," *Machine Learning*, vol. 63, no. 1, pp. 3–42, 2006. DOI: 10.1007/s10994-006-6226-1. [Online]. Available: https://doi.org/10.1007/s10994-006-6226-1.

[24] M.-C. Popescu, V. Balas, L. Perescu-Popescu, and N. Mastorakis, "Multilayer perceptron and neural networks," *WSEAS Transactions on Circuits and Systems*, vol. 8, 2009.

[25] T. Cover and P. Hart, "Nearest neighbor pattern classification," *IEEE Transactions on Information Theory*, vol. 13, no. 1, pp. 21–27, 1967. DOI: 10.1109/TIT.1967.1053964. [Online]. Available: https://doi.org/10.1109/TIT.1967.1053964.

[26] A. Cutler, D. R. Cutler, and J. R. Stevens, "Random forests," in *Machine Learning - ML*, vol. 45, Springer, 2011, pp. 157–176. DOI: 10.1007/978-1-4419-9326-7_5. [Online]. Available: https://doi.org/10.1007/978-1-4419-9326-7_5.

[27] H. Drucker, C. J. Burges, L. Kaufman, A. Smola, and V. Vapnik, "Support vector regression machines," in *Advances in Neural Information Processing Systems*, vol. 9, MIT Press, 1996. [Online]. Available: https://proceedings.neurips.cc/paper_files/paper/1996/hash/d38901788c533e8286cb6400b40b386d-Abstract.html.

[28] E.-S. M. El-Kenawy et al., "Ihow optimization algorithm: A human-inspired metaheuristic approach for complex problem solving and feature selection," *Journal of Artificial Intelligence in Engineering Practice*, vol. 1, no. 2, pp. 36–53, 2024. DOI: 10.21608/jaiep.2024.386694. [Online]. Available: https://doi.org/10.21608/jaiep.2024.386694.

[29] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proceedings of ICNN'95 - International Conference on Neural Networks*, vol. 4, 1995, pp. 1942–1948. DOI: 10.1109/ICNN.1995.488968.

[30] X.-S. Yang, "A new metaheuristic bat-inspired algorithm," in *Nature Inspired Cooperative Strategies for Optimization (NICSO 2010)*, J. R. González, D. A. Pelta, C. Cruz, G. Terrazas, and N. Krasnogor, Eds., Springer, 2010, pp. 65–74. DOI: 10.1007/978-3-642-12538-6_6.

[31] S. Mirjalili and A. Lewis, "The whale optimization algorithm," *Advances in Engineering Software*, vol. 95, pp. 51–67, 2016. DOI: 10.1016/j.advengsoft.2016.01.008.

[32] D. Simon, "Biogeography-based optimization," *IEEE Transactions on Evolutionary Computation*, vol. 12, no. 6, pp. 702–713, 2008. DOI: 10.1109/TEVC.2008.919004.

[33] S. Mirjalili, S. M. Mirjalili, and A. Hatamlou, "Multi-verse optimizer: A nature-inspired algorithm for global optimization," *Neural Computing and Applications*, vol. 27, 2015. DOI: 10.1007/s00521-015-1870-7.

[34] C. Reeves, "Genetic algorithms," in *Handbook of Metaheuristics*, vol. 146, 2010, pp. 109–139. DOI: 10.1007/978-1-4419-1665-5_5.

[35] H. Salimi, "Stochastic fractal search: A powerful metaheuristic algorithm," *Knowledge-Based Systems*, vol. 75, pp. 1–18, 2014. DOI: 10.1016/j.knosys.2014.07.025.

[36] R. Storn and K. Price, "Differential evolution - a simple and efficient heuristic for global optimization over continuous spaces," *Journal of Global Optimization*, vol. 11, pp. 341–359, 1997. DOI: 10.1023/A:1008202821328.

[37] X. Wang, V. Snášel, S. Mirjalili, J.-S. Pan, L. Kong, and H. A. Shehadeh, "Artificial protozoa optimizer (apo): A novel bio-inspired metaheuristic algorithm for engineering optimization," *Knowledge-Based Systems*, vol. 295, p. 111 737, 2024. DOI: 10.1016/j.knosys.2024.111737.