Optimizing CO_2 Emission Forecasting from Vehicles Using Deep Learning and Football Optimization Algorithm

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

Accurate prediction of CO_2 emissions from vehicles is essential for environmental regulation and sustainable transport design. Existing models often suffer from limited accuracy due to suboptimal hyperparameter configurations. This study aims to enhance CO_2 emission forecasting by combining deep learning with advanced metaheuristic optimization. An attention-based Encoder LSTM (EALSTM) model is trained on Canadian vehicle emissions data, with hyperparameters tuned using a novel Football Optimization Algorithm (FbOA), inspired by cooperative team dynamics in football. Comparative evaluation against eight other optimizers shows that FbOA achieves the best performance. The optimized EALSTM model yields an RMSE of 0.00349, MAE of 0.00010, and R^2 of 0.984, outperforming all alternatives. These results demonstrate the effectiveness of domain-inspired metaheuristics in improving prediction accuracy. The proposed FbOA-EALSTM framework offers a scalable, accurate solution for emissions modeling and supports data-driven environmental policy and intelligent vehicle technologies.

Keywords: CO_2 Emissions; Metaheuristic Optimizatio; Football Optimization Algorithm (FbOA); Encoder LSTM (EALSTM); Transportation Sustainability

1 Introduction

Global warming and climate change remain at the forefront of environmental challenges confronting contemporary society, primarily driven by the continuous rise in atmospheric concentrations of greenhouse gases

DOI: https://doi.org/10.54216/JAIM.100203

Received: April 10, 2025 Revised: July 01, 2025 Accepted: September 08, 2025

(GHGs) [1, 2]. Among these, carbon dioxide (CO_2) occupies a central role due to its extensive emission volume, long atmospheric lifetime, and its dominant share in human-induced global warming. Road transportation, in particular, constitutes one of the most significant anthropogenic sources of CO_2, contributing heavily to national and global GHG inventories. The continual increase in private vehicle ownership, rapid urbanization, expansion of road infrastructure, and economic reliance on logistics and mobility services have collectively intensified the environmental burden posed by vehicular CO_2 emissions. This upward trend is particularly evident in emerging economies, but even developed nations continue to face challenges in curbing emissions from personal and commercial transport fleets.

As the urgency to address climate change deepens, there has been a growing impetus to develop data-driven, intelligent systems capable of understanding [3], quantifying [4], and ultimately mitigating the impact of vehicular emissions. Accurate and scalable CO_2 emissions modeling is a cornerstone in this endeavor. On the policy front, emission prediction models inform environmental regulations, support compliance with carbon targets, and assist in setting vehicle-level emissions standards. For automotive manufacturers, such models provide a computational means to evaluate the emissions performance of vehicles during the design and prototyping stages. For consumers and urban planners alike, emission estimates support vehicle eco-labeling programs, congestion pricing, low-emission zones, and emissions-based tolling schemes [5]. Moreover, in a broader sustainability context, accurate forecasting models are indispensable for simulating the future impact of alternative fuel technologies, hybrid and electric drivetrains, and evolving urban mobility patterns. By enabling lifecycle assessments and scenario analyses, these models allow stakeholders to explore decarbonization pathways, evaluate energy policy impacts, and integrate predictive emissions control into intelligent transportation systems [6].

To address the technical complexities of such prediction tasks, the field has increasingly turned to machine learning (ML) methods, which have shown exceptional promise in modeling complex, nonlinear phenomena from large-scale, heterogeneous datasets. Unlike traditional regression approaches that rely on linearity assumptions, handcrafted features, or simplified input-output mappings, ML models—particularly those based on deep learning—are capable of autonomously discovering latent patterns and temporal dependencies. In emission modeling contexts, this flexibility is invaluable, as vehicle emissions are influenced by dynamic operational variables such as speed, acceleration, load, fuel type, engine capacity, and environmental conditions. Recurrent neural network (RNN) architectures, and their variants such as Long Short-Term Memory (LSTM) networks [7], Gated Recurrent Units (GRU) [8], and Bidirectional RNNs (BiRNN), have been widely adopted in modeling such time-dependent relationships. Recent advances have further augmented these models with attention mechanisms, leading to the development of Encoder-Attention LSTM (EALSTM) architectures, which selectively weight input sequences based on their relevance, thereby enhancing predictive fidelity.

While these deep learning architectures offer substantial benefits, their deployment in CO_2 emission modeling presents nontrivial challenges. A major hurdle lies in the high dimensionality of vehicular datasets, which often encompass categorical and numerical variables related to powertrain configurations, transmission systems, drivetrain types, fuel specifications, and manufacturer-based differences. Many of these features are interrelated and, if not properly handled, may introduce noise or redundancy, leading to unstable learning or reduced generalization capability [9]. Compounding this issue is the inherent sensitivity of deep learning models to hyperparameter selection. Critical parameters such as learning rate, number of hidden layers, batch size, dropout ratio, and activation functions significantly influence model behavior. Inadequate tuning can lead to underfitting, overfitting, or poor convergence, especially when models are applied across

DOI: https://doi.org/10.54216/JAIM.100203

Received: April 10, 2025 Revised: July 01, 2025 Accepted: September 08, 2025

varying vehicle classes or temporal segments. Furthermore, the dynamic evolution of the vehicle fleet—with the introduction of electric vehicles, hybrid systems, and alternative fuels—necessitates that emission models be not only accurate but also adaptable and generalizable across diverse and shifting input spaces [10].

Research Gap: Although a growing number of studies have applied ML and deep learning techniques to CO_2 emission prediction, the majority have overlooked the importance of systematic, algorithmic hyperparameter optimization. Existing works tend to rely on manual tuning or simplistic grid search procedures, which are computationally inefficient and prone to convergence toward local optima. These approaches also lack scalability, particularly in complex architectures with high-dimensional hyperparameter spaces. Moreover, very few studies have evaluated such models using real-world, nationally representative datasets that include a wide array of vehicle specifications, fuel types, and drivetrain architectures over multiple years. As a result, the development of scalable, optimized, and generalizable CO_2 emission prediction frameworks remains an open and pressing challenge in both academic research and applied environmental informatics.

Objectives of the Study: This study seeks to bridge the aforementioned research gap by developing a robust, generalizable deep learning framework for vehicle-level CO_2 emission prediction. Specifically, the study pursues three primary objectives. First, it performs a comparative assessment of several state-of-the-art deep learning architectures—LSTM, GRU, BiRNN, ANN, and an enhanced EALSTM model—each evaluated for its ability to capture nonlinear and temporal patterns from raw vehicle data. Second, it applies a suite of advanced metaheuristic optimization algorithms for automated hyperparameter tuning, including the novel Football Optimization Algorithm (FbOA), which draws inspiration from cooperative strategies observed in team sports. Third, it validates the proposed framework using a large-scale, real-world dataset extracted from Canadian government open data portals, which captures a diverse range of vehicles, fuel types, and operational conditions over a span of seven years. Collectively, these objectives aim to yield a reproducible modeling pipeline that can serve as a foundational tool for data-driven policy analysis, vehicle benchmarking, and environmental forecasting.

Research Contributions: The contributions of this study are multifaceted and advance the field in several important ways. Firstly, the study provides a rigorous comparative analysis of multiple deep learning architectures, offering empirical benchmarks for CO_2 emission modeling. Secondly, it introduces a unified optimization framework based on metaheuristic search, enabling efficient, architecture-agnostic tuning of hyperparameters that significantly influence model performance. Thirdly, the study leverages a national-scale, multi-year vehicle dataset, enhancing the generalizability and real-world relevance of the proposed approach. Fourthly, the integration of the Football Optimization Algorithm (FbOA) as a novel, domain-inspired metaheuristic presents an original contribution to the field, showcasing its effectiveness in improving convergence speed, stability, and predictive accuracy. Finally, the research addresses a previously underexplored area by evaluating the interaction between model architecture and hyperparameter sensitivity, offering insights that can inform the design of more resilient and adaptable forecasting systems.

Structure of the Paper: The remainder of this paper is organized as follows. Section 2 presents a detailed overview of the dataset, its preprocessing, and the modeling methodologies, including both deep learning architectures and metaheuristic optimizers. Section 3 reports and analyzes the experimental results of both baseline and optimized models. Section 4 provides a critical discussion of the findings, their implications for current literature, and acknowledged limitations. Section 5 concludes the paper and outlines avenues for future research in the pursuit of intelligent, adaptive, and policy-relevant CO_2 emissions modeling frameworks.

DOI: https://doi.org/10.54216/JAIM.100203

Received: April 10, 2025 Revised: July 01, 2025 Accepted: September 08, 2025

2 Literature Review

Predicting CO_2 emissions from the transportation sector has emerged as a critical research focus due to its direct implications for climate policy, regulatory frameworks, and sustainable development strategies. A wide range of computational and hybrid modeling strategies has been proposed in the literature to improve the precision, adaptability, and interpretability of CO_2 emission forecasts under diverse national and regional contexts.

In Cameroon, Sapnken et al. [11] introduced an Optimized Wavelet Transform Hausdorff Grey Model (OWTHGM(1,N)) to address the inherent limitations of the classical GM(1,N) models, particularly their stability and sensitivity to input noise. The model integrates wavelet-based filtering, fractional Hausdorff derivatives, and a Rao-3 algorithm for parameter optimization. This multi-stage enhancement led to significantly improved forecasting accuracy of road transport emissions, offering practical insights for national-level emission planning.

In Thailand, Sujjaviriyasup and Pitiruek [12] proposed a hybrid model combining Maximal Overlap Discrete Wavelet Transform (MODWT), Support Vector Regression (SVR), and Differential Evolution (DE) to capture emission fluctuations. Their approach showed notable superiority over traditional benchmarks such as ARIMA and Holt models, highlighting the potential of wavelet decomposition and metaheuristics for capturing both short- and long-term emission dynamics.

Beyond purely predictive frameworks, some studies have sought to merge machine learning with environmental efficiency analysis. In Iran, Zadmirzaei et al. [13] employed a fuzzy undesirable slacks-based measure (FUNSBM) within a Data Envelopment Analysis (DEA) framework, integrating it with artificial intelligence tools such as Artificial Neural Networks (ANN), Particle Swarm Optimization (PSO), and Artificial Immune Systems (AIS). Among these, the FUNSBM-AIS configuration achieved the most substantial CO_2 reduction potential, underscoring the value of combining efficiency diagnostics with intelligent optimization for sustainable forestry and land-use management.

In the United Kingdom, interpretability was prioritized in a multi-stage machine learning framework enhanced by SHapley Additive exPlanations (SHAP) [14]. The study revealed that road carbon intensity was the most influential variable in national-level CO_2 predictions, surpassing demographic and economic indicators like population and GDP per capita. This approach not only achieved strong predictive accuracy but also improved model transparency—a crucial feature for policy translation.

At the vehicle level, real-time emission monitoring was addressed by a Long Short-Term Memory (LSTM) model utilizing On-Board Diagnostics (OBD-II) data [15]. By bypassing the need for direct emission sensors, the model supports scalable and cost-effective deployment, especially for fleet monitoring and intelligent transportation systems.

In China, a hybrid Extreme Learning Machine (ELM) model optimized via Manta Ray Foraging Optimization (MRFO) was proposed to forecast national transport sector emissions [16]. The model provided insights into future peak emission scenarios, emphasizing vehicle electrification as a key mitigation strategy and estimating peak periods between 2035 and 2043, contingent on policy intensity.

A separate study in Canada combined mathematical modeling with machine learning to forecast both energy demand and CO_2 emissions in the transportation sector [17]. The framework projected a 50.02% rise

DOI: https://doi.org/10.54216/JAIM.100203

Received: April 10, 2025 Revised: July 01, 2025 Accepted: September 08, 2025

in emissions by 2048, identifying oil consumption as the most influential factor based on sensitivity analysis. This work illustrated the power of hybrid models in long-range strategic planning.

For broader geopolitical contexts, a study across G8 nations utilized multilayer perceptrons (MLPs) enhanced with nature-inspired optimization algorithms such as Teaching-Learning-Based Optimization (TLBO) and Vortex Search (VS) [18]. TLBO in particular achieved the best tradeoff between training and testing error, showcasing strong generalization across heterogeneous national contexts.

Forecasting India's greenhouse gas emissions, another study employed recent metaheuristic algorithms like Marine Predators Algorithm (MPA) and Lightning Search Algorithm (LSA) [19]. The findings predicted a 2.87-fold increase in CO_2 emissions by 2050, with MPA outperforming its counterparts in terms of predictive accuracy.

Similarly, a Red-Billed Blue Magpie Optimizer (RBMO) was used in Indonesia to optimize an ELM model for national-level CO_2 forecasting [20]. The study identified social globalization and ecological footprint as the dominant explanatory variables, confirming the interconnected nature of economic integration and environmental impact.

Expanding into regional analyses, hybrid models such as Moth Flame Optimization (MFO) coupled with MLP, Genetic Programming (GP), and Random Forest (RF) were applied to the OECD Asia-Pacific region [21]. These models incorporated environmental, socioeconomic, and meteorological variables, achieving near-perfect R^2 values (above 0.999), thereby offering exceptional accuracy for policy-oriented forecasting.

A recent study in China proposed a hybrid framework using Least Squares Support Vector Machine (LSSVM) optimized with an improved Sparrow Search Algorithm (SSA), incorporating Sin chaos mapping, adaptive inertia, and Cauchy-Gaussian perturbation [22]. This model demonstrated strong predictive capability even on small datasets, emphasizing its practical viability in data-sparse environments.

Beyond transport, emissions-aware AI modeling has extended to low-carbon material design. A study using ANN models optimized by Whale Optimization Algorithm (WOA), Cuckoo Search Optimization (CSO), and MFO aimed to predict compressive strength in sustainable concrete [23]. The ANN-WOA model delivered the most accurate results, supporting SDG-aligned infrastructure development.

Production-routing systems have also integrated emissions considerations. A mixed-integer linear programming (MILP) model, solved with PSO, was proposed to minimize both cost and CO_2 emissions [24]. The inclusion of logistics outsourcing and route optimization highlighted the trade-offs between operational efficiency and environmental performance.

At the provincial level in China, a Grey Convolution Model based on the UNAGO operator was designed for time-series prediction [25]. Its adaptability across different temporal scales allowed it to outperform conventional models, yielding Mean Absolute Percentage Errors (MAPE) below 10%, thus proving useful for sub-national planning.

Deep learning-based studies have also improved via advanced preprocessing. A Dual-Path Recurrent Neural Network (DPRNN) optimized using a novel Nigerian Optimization Algorithm (NiOA) integrated

DOI: https://doi.org/10.54216/JAIM.100203

Received: April 10, 2025 Revised: July 01, 2025 Accepted: September 08, 2025

Principal Component Analysis (PCA) and Blind Signal Separation (BSS) [26]. The model achieved an R^2 of 0.9736, confirming its predictive power in complex datasets.

Central European CO_2 emissions were addressed using MLP models tuned by multiple optimizers including Biogeography-Based Optimization (BBO), Shuffled Complex Evolution (SCE), and Firefly Swarm Algorithm (FSA) [27]. Among them, SCE and BBO consistently achieved the highest R^2 scores and the lowest RMSE values, demonstrating high adaptability across urban data streams.

Finally, in Tamil Nadu, India, the Chaotic Ant Lion Optimization Technique (CALOT) was applied to joint energy demand forecasting and capacity expansion planning [28]. This study integrated emission constraints into long-term energy strategy design, aiding state-level decarbonization initiatives.

Table 1 provides a structured summary of these reviewed studies, outlining their methodological contributions, target domains, and key insights. The diversity of modeling approaches—from hybrid AI systems and wavelet-enhanced models to metaheuristic optimization frameworks—underscores the growing complexity and interdisciplinary nature of CO_2 emissions forecasting research.

Ref.	**Focus Area**	**Methodology**	**Key Findings and Contributions**
[11]	Road transport emissions in Cameroon	OWTHGM(1,N) with wavelet filtering and Hausdorff derivative	Improved prediction accuracy; supports policy planning in Cameroon
[12]	Emission monitoring in Thailand	MODWT-SVR-DE hybrid model	Outperforms ARIMA, Holt, SES; suitable for emission control policies
[13]	Environmental efficiency in Iran	FUNSBM + ANN, PSO, AIS	FUNSBM-AIS reduces CO_2 emissions by 53%; evaluates inefficiency in DMUs
[14]	UK transportation sector	SHAP-enhanced ML model	Highlights road carbon intensity; improves model interpretability
[15]	Real-time vehicular emission estimation	LSTM with OBD-II data	Scalable emission monitoring without onboard sensors
[16]	China transport emission peaks	MRFO-ELM hybrid model	Predicts emission peaks under different scenarios; stresses electrification
[17]	Canada transport emission forecast	Mathematical + ML hybrid	Predicts 50.02% emission rise by 2048; oil most impactful
[18]	CO_2 prediction in G8 nations	MLP with nature-inspired optimization	TLBO and VS show best accuracy and generalization
[19]	India's GHG emissions trajectory	MPA, LSA, SOS, BSA	CO_2 to rise 2.5–2.87x by 2050; MPA most accurate
[20]	Indonesia's CO_2 emissions	RBMO-ELM	Social globalization and ecological footprint key drivers
[21]	Emissions in OECD Asia-Pacific	MFO/SHO/GOA + MLP, GP, RF	MFO-MLP achieves top accuracy; supports policy design

DOI: https://doi.org/10.54216/JAIM.100203

Received: April 10, 2025 Revised: July 01, 2025 Accepted: September 08, 2025

Ref.	Focus Area	Methodology	Key Findings and Contributions
[22]	China carbon forecasting	Improved SSA + LSSVM	Accurate predictions with limited data; strong model stability
[23]	Concrete sustainability	ANN with WOA, CSO, MFO	ANN-WOA most accurate; supports SDG 13 through low-carbon design
[24]	Green production-routing	MILP + PSO	Optimizes cost and emissions; PSO outperforms VNS
[25]	Provincial CO_2 emissions in China	Grey convolution model with UNAGO	Outperforms traditional models; robust under multiple horizons
[26]	High-accuracy CO_2 prediction	NiOA-DPRNNs with PCA, BSS	R^2 = 0.9736; strong specificity and robustness
[27]	Urban CO_2 in Central Europe	MLP + BBO, SCE, FSA	SCE and BBO produce best metrics (R^2 = 0.9999)
[28]	Power planning in Tamil Nadu	CALOT for DFP and CEP	Integrates emissions and cost into long-term expansion plans

Table 1: Summary of Reviewed Studies

3 Materials and Methods

In order to construct a reliable and scalable predictive framework for vehicular CO_2 emissions, this study utilizes a robust methodology that integrates real-world data acquisition, data preprocessing, model formulation, and hyperparameter optimization. The modeling pipeline is rooted in a data-driven architecture where machine learning models are employed to capture the underlying nonlinear relationships between vehicle specifications and the resulting CO_2 output. The modeling process is enhanced through the application of metaheuristic algorithms, which facilitate the automated tuning of hyperparameters, thereby improving model stability and generalization. The approach is systematic and reproducible, designed to ensure that findings can be extended to other datasets and emission contexts. This section presents the dataset used in the study and outlines the methodological approach adopted for data processing and model development.

3.1 Dataset Description

The dataset employed in this study originates from the official open data repository of the Government of Canada, specifically curated to reflect CO_2 emissions from passenger vehicles. It provides a comprehensive and longitudinal view of vehicle characteristics and their associated emissions over a period of seven consecutive years. The dataset contains a total of 7,385 instances, each corresponding to a unique vehicle model configuration, with 12 distinct attributes characterizing various physical, mechanical, and fuel-related properties. These features serve as potential predictors in the machine learning models developed for CO_2 estimation.

The key attributes in the dataset include information about the vehicle's drivetrain type, transmission system, fuel type, engine displacement, number of cylinders, city and highway fuel consumption rates

DOI: https://doi.org/10.54216/JAIM.100203

Received: April 10, 2025 Revised: July 01, 2025 Accepted: September 08, 2025

(measured in liters per 100 kilometers), and the combined CO_2 emissions output (measured in grams per kilometer). The dataset further incorporates various domain-specific abbreviations, such as drivetrain configurations denoted by `4WD/4X4` (Four-Wheel Drive), `AWD` (All-Wheel Drive), `SWB` (Short Wheelbase), `LWB` (Long Wheelbase), and `EWB` (Extended Wheelbase). Transmission systems are encoded as `A` (Automatic), `M` (Manual), `AM` (Automated Manual), `AS` (Automatic with Select Shift), and `AV` (Continuously Variable), followed by a numerical suffix indicating the number of gears, typically ranging from 3 to 10.

Fuel types are classified using single-letter codes: `X` indicates regular gasoline, `Z` denotes premium gasoline, `D` corresponds to diesel, `E` represents ethanol blends such as E85, and `N` signifies natural gas. These categorical variables are encoded using appropriate transformation techniques during preprocessing. Fuel consumption values are provided separately for city and highway conditions, and the combined consumption is presented both in L/100 km and miles per gallon (mpg), though the L/100 km metric is used for model training due to its consistency with the CO_2 target variable.

The target variable, CO_2 emissions, represents the tailpipe emission rate for each vehicle under combined city and highway driving conditions, weighted at 55% city and 45% highway usage. It is expressed in grams per kilometer (g/km) and constitutes a direct measurement that aligns with national emission standards and environmental regulatory reporting.

Given the nature of real-world government-sourced datasets, the integrity and consistency of the data are deemed reliable. Nevertheless, preprocessing steps are conducted to handle missing values, ensure numerical consistency, and standardize units where required. The data are randomly partitioned into training and testing subsets using an 80% to 20% split, respectively, to allow for model validation on previously unseen instances. No external datasets are integrated, and the analysis is entirely self-contained within the provided dataset, ensuring transparency and reproducibility.

The selection of this dataset is motivated by its official status, broad representational scope, and structured encoding of relevant vehicle attributes. It offers a rich ground for examining the applicability of advanced machine learning models to real-world CO_2 emission prediction and serves as an ideal testbed for evaluating the influence of model architecture and optimization on predictive performance.

3.2 Data Preprocessing

Effective data preprocessing is a critical prerequisite for building reliable and generalizable machine learning models. In this study, the raw dataset underwent a sequence of preprocessing steps to ensure consistency, enhance data quality, and facilitate model convergence. These steps were applied uniformly to both training and testing subsets to preserve model integrity and avoid data leakage. Missing values were assessed across all columns. Given the structured nature of the dataset and its government-sourced origin, the proportion of missing data was minimal. Nevertheless, any incomplete records were removed to avoid bias and preserve statistical validity. No imputation techniques were required due to the sparsity of such instances. Categorical variables—including drivetrain type, fuel type, and transmission configuration—were encoded using a combination of label encoding and one-hot encoding. Label encoding was applied where ordinal relationships were present or model architectures could leverage embedding layers, while one-hot encoding was used for nominal features to prevent introducing artificial hierarchies. Numerical features, such as engine displacement, number of cylinders, and fuel consumption metrics, were normalized using Min-Max scaling

DOI: https://doi.org/10.54216/JAIM.100203

Received: April 10, 2025 Revised: July 01, 2025 Accepted: September 08, 2025

to transform their values into a uniform range between 0 and 1. This standardization step was essential for accelerating the convergence of gradient-based learning algorithms and ensuring balanced contributions of input features. The target variable, CO_2 emissions (g/km), was not normalized to preserve interpretability of model outputs and maintain alignment with real-world reporting standards. However, it was rescaled internally during loss computation where applicable, depending on the model's architecture. Finally, the dataset was randomly partitioned into training and testing subsets using an 80% to 20% split, respectively. The random seed was fixed to ensure reproducibility of experimental results. No stratification was applied, as the emission target variable is continuous rather than categorical. The resulting training data served as input for model development and hyperparameter tuning, while the testing data was used exclusively for evaluating final model performance. The relationship between engine specifications, fuel consumption patterns, and CO_2 emissions is central to understanding vehicular environmental impacts. Correlation analysis provides a quantitative means to evaluate how these factors are interlinked and to identify potential redundancies among variables. As illustrated in Figure 1, the analysis shows that engine size and the number of cylinders are strongly correlated, both of which also demonstrate significant positive associations with CO_2 emissions. Similarly, fuel consumption metrics in city, highway, and combined driving conditions exhibit high positive correlations with emissions, while fuel efficiency measured in miles per gallon (mpg) reveals a strong negative correlation. These findings emphasize that larger engines with higher fuel consumption contribute substantially to increased emissions, whereas improvements in fuel efficiency act inversely. Such insights are critical for developing policies aimed at reducing emissions through both technological innovation and stricter fuel economy standards.

Dimensionality reduction techniques, such as Principal Component Analysis (PCA), are essential for uncovering latent structures within high-dimensional vehicular datasets. By projecting the original variables onto orthogonal components, PCA facilitates the visualization of variance patterns while retaining the majority of the informational content. As shown in Figure 2, the first two principal components (PCA1 and PCA2) reveal distinct clustering behavior across different fuel types. The distribution of PCA1 highlights notable differences between conventional fuel types, particularly gasoline (Z) and diesel (D), compared to alternative fuels such as ethanol (E) and natural gas (N). Similarly, the distribution along PCA2 further distinguishes categories, suggesting that fuel-specific characteristics capture meaningful variability in the dataset. These results indicate that PCA not only reduces dimensional complexity but also provides discriminative power for identifying structural differences between fuel categories, which is valuable for both classification tasks and emission pattern analysis.

Evaluating CO_2 emissions across different vehicle classes provides a clear perspective on how design, size, and intended usage affect environmental performance. Boxplot analysis allows for the visualization of both the central tendency and variability of emissions, making it possible to identify not only the average behavior but also the range of outliers within each class. As illustrated in Figure 3, smaller vehicle types such as station wagons, compacts, and subcompacts generally exhibit lower median emissions and narrower variability. In contrast, larger classes including standard SUVs, cargo vans, and especially passenger vans show significantly higher emission levels, with passenger vans reaching a median of nearly 400 g/km. These findings highlight the disproportionate contribution of heavy-duty and utility vehicles to total emissions, underscoring the importance of class-specific regulatory frameworks and the promotion of low-emission alternatives in high-impact categories.

DOI: https://doi.org/10.54216/JAIM.100203

Received: April 10, 2025 Revised: July 01, 2025 Accepted: September 08, 2025

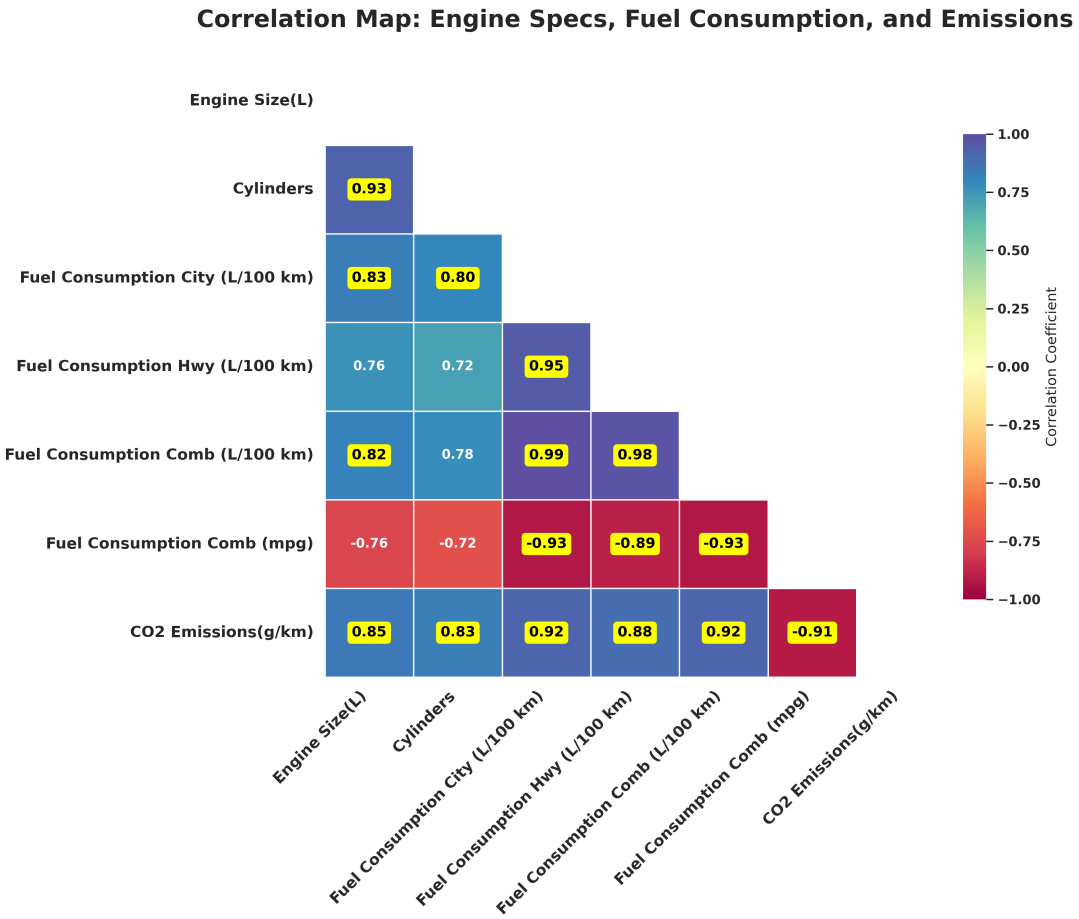

Figure 1: Correlation Map of Engine Specifications, Fuel Consumption, and CO_2 Emissions

3.3 Machine Learning Models

The core objective of this study is to develop and optimize a machine learning framework capable of accurately predicting CO_2 emissions from passenger vehicles based on real-world technical specifications. The modeling strategy is centered around a deep learning paradigm, with a particular emphasis on sequence-based architectures that are capable of modeling complex, nonlinear interactions among input features. The proposed model is the **Encoder Attention-based Long Short-Term Memory (EALSTM)** network, which extends the classical LSTM architecture by integrating a trainable attention mechanism into the encoder-decoder structure. The inclusion of attention allows the model to assign dynamic importance weights to each input timestep, thereby enhancing the learning of inter-feature relevance in structured vehicle data. The motivation for adopting EALSTM stems from the fact that while vehicle configuration attributes—such as transmission type, drivetrain, engine size, and fuel consumption—are not explicitly temporal, their inter-dependencies can be treated as pseudo-sequential for the purpose of modeling. For example, the influence

DOI: https://doi.org/10.54216/JAIM.100203

Received: April 10, 2025 Revised: July 01, 2025 Accepted: September 08, 2025

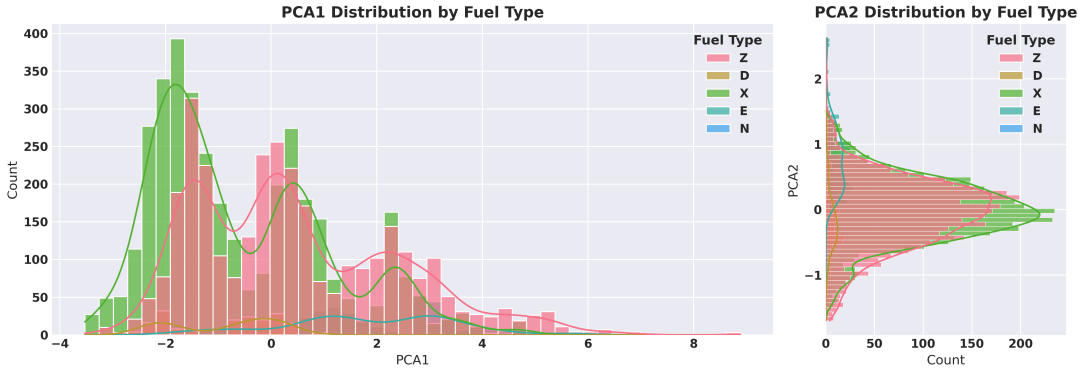

Figure 2: Distribution of PCA1 and PCA2 by Fuel Type

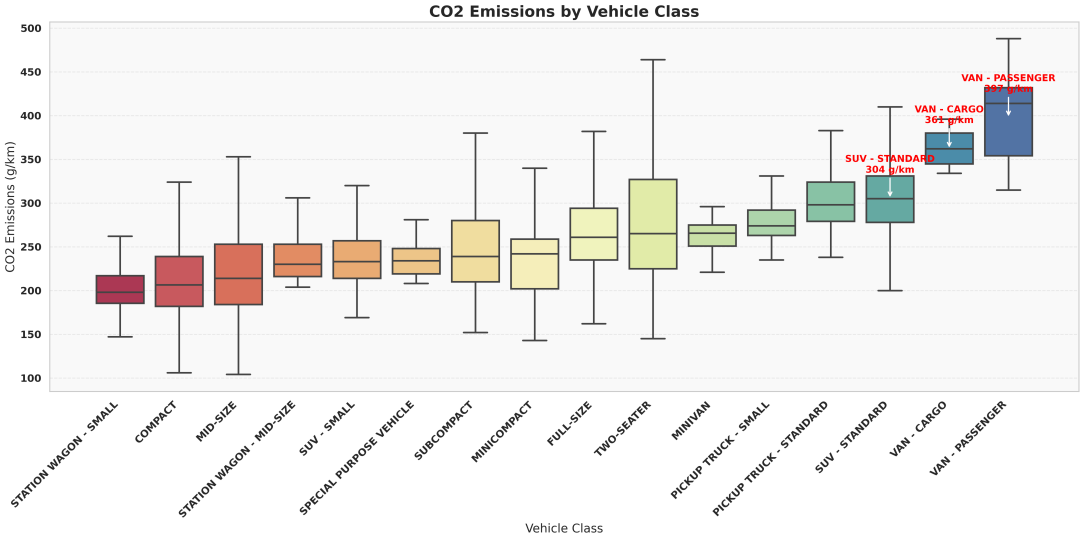

Figure 3: CO_2 Emissions Distribution by Vehicle Class

of fuel type on CO_2 emissions may be modulated by the presence of certain transmission systems or fuel efficiency ratings. Treating these interrelated attributes as a structured input sequence enables the model to learn latent interactions more effectively.

3.3.1 EALSTM Model Architecture

Let $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_T]$ represent the sequence of input vectors for a single vehicle, where T is the number of encoded attributes and each $\mathbf{x}_t \in \mathbb{R}^d$ denotes the t^{th} feature embedding. The encoder processes $\mathbf{X}$ through a series of LSTM cells:

DOI: https://doi.org/10.54216/JAIM.100203

Received: April 10, 2025 Revised: July 01, 2025 Accepted: September 08, 2025

$$\mathbf{h}_t, \mathbf{c}_t = \text{LSTM}(\mathbf{x}_t, \mathbf{h}_{t-1}, \mathbf{c}_{t-1}) \tag{1}$$

where $\mathbf{h}_t$ and $\mathbf{c}_t$ denote the hidden and cell states at timestep t, respectively.

The attention mechanism computes a context vector $\mathbf{c}$ as a weighted sum of encoder hidden states:

$$\alpha_t = \frac{\exp(e_t)}{\sum_{k=1}^{T} \exp(e_k)}, \quad \text{where} \quad e_t = \mathbf{v}^\top \tanh(\mathbf{W}_h \mathbf{h}_t + \mathbf{b}) \tag{2}$$

$$\mathbf{c} = \sum_{t=1}^{T} \alpha_t \mathbf{h}_t \tag{3}$$

Here, α_t represents the attention weight assigned to each timestep, and $\mathbf{c}$ is the attention-weighted context vector passed to the output layer. The final prediction of CO_2 emissions $\hat{y}$ is given by:

$$\hat{y} = \sigma(\mathbf{W}_o \mathbf{c} + \mathbf{b}_o) \tag{4}$$

where σ denotes the output activation function (typically linear for regression), and $\mathbf{W}_o$ and $\mathbf{b}_o$ are the output layer weights and bias. This attention mechanism allows the model to selectively emphasize influential input features without requiring explicit feature engineering or manual weighting. It also improves model interpretability by offering insight into which features the model considers most relevant during prediction[29].

3.3.2 Comparative Models

To establish a rigorous performance benchmark for the proposed EALSTM model, several well-established deep learning architectures are implemented and evaluated. These models, described below, represent a spectrum of recurrent and non-recurrent neural networks used in structured data modeling:

1. **LSTM (Long Short-Term Memory)**: A recurrent neural network architecture designed to overcome the vanishing gradient problem by incorporating memory cell units. LSTM captures long-term dependencies among features and has been widely used in sequence prediction tasks[30].
2. **GRU (Gated Recurrent Unit)**: A simplified variant of LSTM with fewer parameters, combining the forget and input gates into a single update gate. GRU models are often more computationally efficient while preserving comparable learning capacity[31].

DOI: https://doi.org/10.54216/JAIM.100203

Received: April 10, 2025 Revised: July 01, 2025 Accepted: September 08, 2025

3. **BiRNN (Bidirectional Recurrent Neural Network)**: Processes input sequences in both forward and backward directions, enabling the model to incorporate contextual information from the entire sequence. While less common in structured vehicle data applications, BiRNNs can offer enhanced interpretability by evaluating both prior and subsequent feature contributions[32].
4. **RNN (Standard Recurrent Neural Network)**: A fundamental sequence model that uses hidden states to capture temporal correlations. While its performance is limited by gradient instability in longer sequences, RNN serves as a baseline for understanding the incremental benefits of gated and bidirectional extensions[33].
5. **ANN (Artificial Neural Network)**: A classical feedforward neural network composed of multiple fully connected layers with nonlinear activation functions. Unlike recurrent models, ANNs do not account for sequential dependencies but are capable of modeling nonlinear relationships between input and output through depth and width in layer construction[34].

All models are implemented within a consistent training environment, with identical loss functions, optimizers, and regularization techniques. To ensure fair comparison, model hyperparameters are optimized independently using metaheuristic algorithms, which are discussed in Section **??**. The comparative evaluation aims to assess the trade-offs in model complexity, generalization capacity, and learning behavior when applied to structured vehicular emission prediction.

3.4 Metaheuristic Algorithms

To automate the hyperparameter optimization process of deep learning models and to enhance their predictive accuracy, this study integrates a class of population-based metaheuristic algorithms. Among these, the proposed Football Optimization Algorithm (FbOA) is the primary optimization mechanism, evaluated comparatively against a diverse group of well-established algorithms. Metaheuristics are particularly suited for hyperparameter tuning due to their capacity to navigate large, complex, and non-convex search spaces without relying on gradient information.

Football Optimization Algorithm (FbOA)

The Football Optimization Algorithm (FbOA) is a recent population-based metaheuristic inspired by the dynamic strategies and tactical formations observed in football (soccer). In this metaphor, each agent in the population represents a player, and the optimization process simulates team-based interactions such as passing, movement, and positioning. FbOA aims to strike an effective balance between exploration and exploitation by mimicking diverse passing strategies—short passes for local search and long passes for global search—alongside coordinated movement across the solution space[35].

DOI: https://doi.org/10.54216/JAIM.100203

Received: April 10, 2025 Revised: July 01, 2025 Accepted: September 08, 2025

Exploration Phase In the exploration stage, agents are encouraged to search diverse areas of the solution space. The velocity of each agent (player) is modeled as:

$$V_n = F_{\max}\left(b_x \cdot a_i(F_{\text{ext}} - F_{\min}) + r \cdot b_y \cdot a_j(F_{\text{best}} - F_{\min}) \cdot \cos\left(\frac{\pi}{\text{Iteration}}\right)\right) \tag{5}$$

where $F_{\max}$ is the maximum force, $F_{\min}$ the minimum force, F_{ext} the external influence, and F_{best} the best-known force or solution so far. The coefficients a_i, a_j, b_x, b_y control movement dynamics, and the cosine modulation ensures smooth convergence behavior across iterations.

Exploitation Phase To focus search around high-quality solutions, FbOA employs a local refinement mechanism expressed as:

$$Fb(S_{t+1}) = F_i + z_3 \cdot Fb(S_t) + K \cdot \sin\left(\frac{\pi}{\text{Iteration}}\right) \tag{6}$$

Here, F_i is the current position, $Fb(S_t)$ is the agent's state at iteration t, z_3 is a balance control parameter, and K modulates the shift toward exploitation as iterations progress. The sinusoidal term accelerates convergence in the final stages of the optimization process.

Mutation Strategy To avoid entrapment in local minima, a controlled mutation operator is applied:

$$S_t = K \cdot a_q\left(\frac{2n+1}{x}\right) + K \cdot \cos\left(\frac{\pi}{\text{Iteration}}\right) \tag{7}$$

where a_q is a scaling factor and K governs mutation intensity. This mechanism introduces diversity by shifting agents to unexplored regions of the search space.

Algorithm 1 Football Optimization Algorithm (FbOA)

```
Initialize population of agents with random positions
Evaluate initial fitness of each agent
while termination condition not met do
    for each agent do
        Compute velocity using Eq. (1)
        Update position using Eq. (2) for exploitation or exploration
        Apply mutation via Eq. (3)
        Evaluate new fitness
        if fitness improved then
            Update best-known solution
        end if
    end for
end while
Return best solution found
```

DOI: https://doi.org/10.54216/JAIM.100203

Received: April 10, 2025 Revised: July 01, 2025 Accepted: September 08, 2025

In this study, FbOA was used to optimize the key hyperparameters of the EALSTM model, including learning rate, batch size, number of hidden units, dropout ratio, and number of training epochs. Its design enables fast convergence, robustness, and strong generalization capabilities in deep learning applications.

Comparative Metaheuristic Algorithms

To benchmark the effectiveness of FbOA, eight additional metaheuristic algorithms were employed. Each was applied independently under the same experimental settings to ensure consistent and fair evaluation. These algorithms are summarized below:

1. **Differential Evolution (DE):** A robust population-based algorithm that perturbs individuals by vector differences from other randomly selected candidates. DE is widely used for its global search efficiency in continuous spaces[36].
2. **Genetic Algorithm (GA):** Modeled after natural evolution, GA operates using crossover, selection, and mutation. It encourages diversity and convergence by simulating biological inheritance over generations[37].
3. **Stochastic Fractal Search (SFS):** Based on fractal diffusion processes, SFS uses Gaussian walks to balance local and global exploration dynamically, making it effective in rugged optimization landscapes[38].
4. **JAYA Algorithm:** A parameter-less optimizer that iteratively shifts each candidate toward the best solution and away from the worst. Its minimal configuration overhead makes it suitable for black-box optimization tasks[39].
5. **Quantum-Inspired Optimizer (QIO):** Leverages principles of quantum computation, such as superposition and rotation gates, to enhance diversity and maintain exploration during stagnation[40].
6. **Bat Algorithm (BA):** Mimics the echolocation behavior of bats to switch between local and global search, using adaptive loudness and pulse rates to fine-tune search behavior[41].
7. **Biogeography-Based Optimizer (BBO):** Inspired by ecological migration, BBO evolves solutions by sharing features across populations (habitats), promoting diversity and exploitation[42].
8. **Multiverse Optimization (MVO):** Simulates search as a universe of candidate solutions, where agents evolve via white holes (exploration), black holes (exploitation), and wormholes (long jumps)[43].

Each optimizer was used to tune the same set of hyperparameters for all deep learning models included in this study. Their performance is comparatively analyzed in Result Section, where FbOA consistently demonstrates superior convergence speed and predictive robustness.

DOI: https://doi.org/10.54216/JAIM.100203

Received: April 10, 2025 Revised: July 01, 2025 Accepted: September 08, 2025

3.5 Evaluation Metrics

To assess the predictive performance of the machine learning models used for CO_2 emission estimation, this study employs a comprehensive suite of statistical evaluation metrics. These metrics capture different aspects of prediction accuracy, correlation, bias, and model generalization. The use of multiple metrics ensures a robust and multi-dimensional evaluation framework suitable for regression-based environmental modeling. Table 2 lists the definitions and mathematical expressions for the nine evaluation metrics used in this study.

Table 2: Regression evaluation 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)	$\text{MBE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)$
Correlation Coefficient (r)	$r = \frac{\sum(y_i - \bar{y})(\hat{y}_i - \bar{\hat{y}})}{\sqrt{\sum(y_i - \bar{y})^2}\sqrt{\sum(\hat{y}_i - \bar{\hat{y}})^2}}$
Coefficient of Determination (R^2)	$R^2 = 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y})^2}$
Relative Root Mean Squared Error (RRMSE)	$\text{RRMSE} = \frac{\text{RMSE}}{\bar{y}} \times 100$
Nash–Sutcliffe Efficiency (NSE)	$\text{NSE} = 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y})^2}$
Willmott's Index of Agreement (WI)	$\text{WI} = 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum\left(\lvert\hat{y}_i - \bar{y}\rvert + \lvert y_i - \bar{y}\rvert\right)^2}$

DOI: https://doi.org/10.54216/JAIM.100203
Received: April 10, 2025 Revised: July 01, 2025 Accepted: September 08, 2025

4 Results

This section presents a comprehensive and in-depth evaluation of the experimental outcomes obtained from the CO_2 emission prediction models proposed and described in the preceding sections. Each model was evaluated using the multi-metric framework outlined in Table 2, which includes both error-based and correlation-based indicators to ensure a holistic performance assessment. The experiments were conducted under rigorously controlled conditions, including standardized preprocessing pipelines, consistent train–test splits, and identical validation procedures.

The results are organized to first present the performance of the deep learning models under default or manually configured hyperparameter settings, which serves as the baseline scenario. This is followed by an analysis of the impact of metaheuristic-based optimization on these models, with a particular emphasis on the proposed Football Optimization Algorithm (FbOA). The goal of this structure is to provide a clear comparative landscape that highlights the performance improvements made possible by integrating intelligent, domain-inspired optimization strategies.

4.1 Baseline Deep Learning Model Performance

Prior to any hyperparameter optimization, a suite of deep learning architectures was trained and evaluated on the same dataset using consistent configurations. The selected models include Encoder Attention-based LSTM (EALSTM), standard LSTM, Gated Recurrent Unit (GRU), Bidirectional Recurrent Neural Network (BiRNN), basic Recurrent Neural Network (RNN), and Artificial Neural Network (ANN). All models were implemented using an 80% training and 20% testing split, ensuring a fair and reproducible evaluation protocol across architectures.

Performance was quantified using a battery of metrics, including Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Bias Error (MBE), correlation coefficient (r), Coefficient of Determination (R^2), Relative RMSE (RRMSE), Nash–Sutcliffe Efficiency (NSE), and Willmott's Index of Agreement (WI). These metrics provide complementary perspectives: while MSE and RMSE penalize larger deviations, MAE offers insight into average prediction accuracy, and R^2, r, NSE, and WI reflect goodness-of-fit and model agreement.

The results, shown in Table 3, reveal that EALSTM significantly outperforms all other models across every metric. Specifically, it achieves the lowest RMSE (0.0862), MAE (0.0094), and MSE (0.00743), while also securing the highest R^2 (0.9059), NSE (0.8897), and WI (0.8869). These findings highlight the efficacy of the attention mechanism in enhancing temporal sequence learning by assigning contextually relevant weights to input features over time. The standard LSTM and GRU models follow closely, demonstrating robust temporal modeling capabilities, though they fall short in both precision and bias minimization compared to EALSTM.

Conversely, traditional RNN and ANN models exhibited weaker performance. RNN's instability in handling long sequences likely contributed to its higher RMSE (0.2813) and MAE (0.0509), while ANN, lacking sequence awareness, showed the poorest results overall with an RMSE of 0.3447 and an MAE of

DOI: https://doi.org/10.54216/JAIM.100203

Received: April 10, 2025 Revised: July 01, 2025 Accepted: September 08, 2025

Table 3: Performance of baseline deep learning models prior to optimization.

Model	MSE	RMSE	MAE	MBE	r	R^2	RRMSE	NSE	WI
EALSTM	0.00743	0.08621	0.00944	0.00920	0.89329	0.90589	0.73250	0.88972	0.88690
LSTM	0.02034	0.14260	0.02797	0.02501	0.87707	0.88967	0.92297	0.86741	0.86759
GRU	0.03048	0.17459	0.03254	0.02810	0.85363	0.86623	1.31403	0.84353	0.85894
BiRNN	0.05409	0.23257	0.03712	0.04723	0.85199	0.86459	0.41188	0.83303	0.82661
RNN	0.07913	0.28130	0.05085	0.05321	0.85313	0.85858	3.00056	0.82775	0.82176
ANN	0.11879	0.34466	0.07499	0.10136	0.83342	0.83249	3.57880	0.81693	0.81266

0.0750. These results underscore the importance of memory-based architectures in time-series emissions prediction.

Visual analysis further supports these findings. Figure 4 presents swarm-overlaid boxplots for each metric across all models. EALSTM's metrics are tightly clustered with minimal spread, reflecting its high consistency and stability. In contrast, RNN and ANN show broader distributions and several outliers, indicating volatility in predictions and susceptibility to overfitting or underfitting.

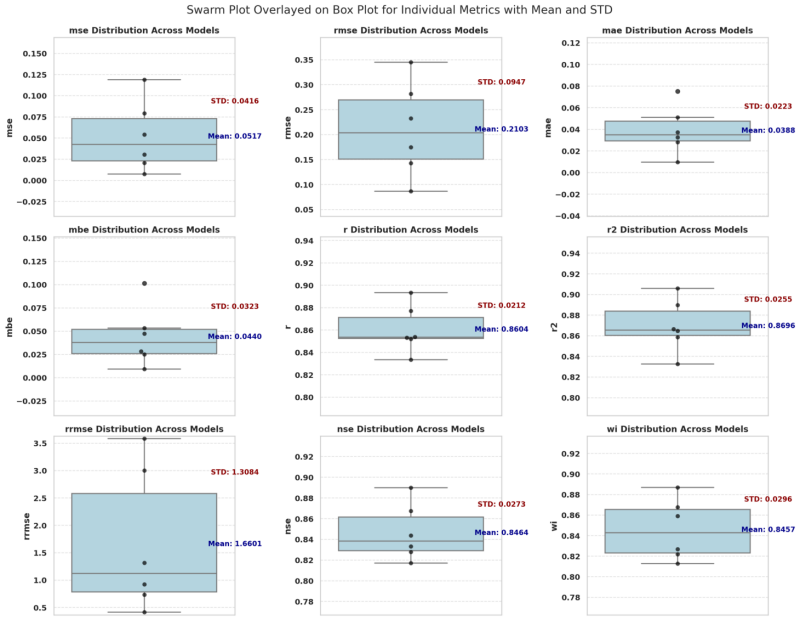

Figure 4: Swarm and boxplot visualization of model performance metrics with mean and standard deviation.

Figure 5 provides a consolidated performance heatmap across all models and metrics. Here, EALSTM emerges with a clear advantage across most indicators, shaded in deep green to denote superior values. The ANN and RNN models, by contrast, are heavily represented in the high-error spectrum, confirming their comparative underperformance.

Further analysis of metric relationships is shown in Figure 6, which demonstrates the linear correlation

DOI: https://doi.org/10.54216/JAIM.100203

Received: April 10, 2025 Revised: July 01, 2025 Accepted: September 08, 2025

Figure 5: Heatmap of comparative model metrics before optimization.

between MAE and RMSE. This supports the assumption that models with lower average absolute error also exhibit lower root-mean-square error, although the latter is more sensitive to extreme deviations.

4.2 Optimized Model Analysis

Having established the baseline, this section examines the impact of hyperparameter optimization on the predictive accuracy of the EALSTM model. Nine metaheuristic algorithms—including DE, GA, SFS, JAYA, QIO, BA, BBO, MVO, and the proposed FbOA—were employed to fine-tune critical model parameters. All tuning was performed using the same training and testing sets to ensure comparability.

Table 4 presents the performance metrics post-optimization. The FbOA-EALSTM combination outperformed all other configurations, achieving an exceptionally low RMSE of 0.00349, a negligible MAE of 0.00010, and an outstanding R^2 value of 0.984. This reflects not only precise predictions but also a high degree of consistency between observed and predicted outputs.

A balanced assessment of predictive performance requires not only examining the mean values of evaluation metrics but also understanding their variability across models. Line plots provide a concise way to compare central tendencies with corresponding deviations, thereby offering a dual perspective on accuracy and stability. As depicted in Figure 7, the mean values of correlation-based indicators (r, R^2, NSE,

DOI: https://doi.org/10.54216/JAIM.100203

Received: April 10, 2025 Revised: July 01, 2025 Accepted: September 08, 2025

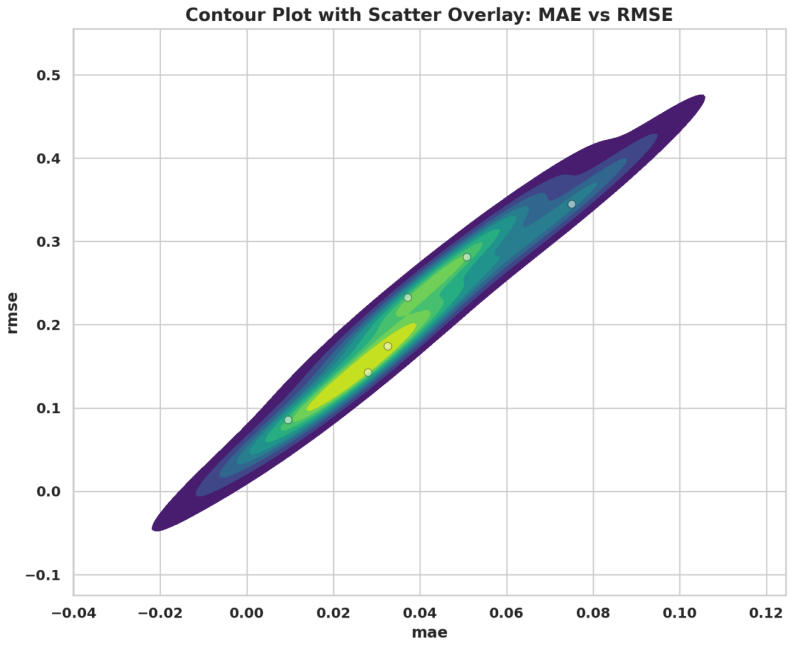

Figure 6: Contour plot with scatter overlay: relationship between MAE and RMSE.

Table 4: Performance of EALSTM model after hyperparameter optimization using different metaheuristic algorithms.

Optimizer + Model	**MSE**	**RMSE**	**MAE**	**MBE**	**r**	**R^2**	**RRMSE**	**NSE**	**WI**
FbOA + EALSTM	1.22E-05	0.00349	0.00010	0.00012	0.979	0.984	0.0474	0.977	0.980
DE + EALSTM	1.52E-04	0.01232	0.00207	0.00078	0.951	0.961	0.0664	0.965	0.968
GA + EALSTM	3.70E-04	0.01924	0.00208	0.00091	0.950	0.955	0.0754	0.961	0.960

WI) remain high, reflecting strong predictive capacity, while error-based metrics such as MSE, RMSE, and MAE exhibit relatively low averages. However, the relative root mean squared error (RRMSE) shows both the highest mean and standard deviation, indicating greater instability and sensitivity across models. By contrast, metrics such as MAE and MBE display consistently low variability, suggesting robustness. This visualization highlights the importance of complementing mean-based evaluation with dispersion measures to avoid misleading conclusions regarding model performance.

The comparative advantage of FbOA lies in its hybrid exploration–exploitation mechanism, inspired by football team dynamics, which allows it to avoid local minima while converging rapidly on high-fidelity hyperparameter sets. In contrast, other optimizers such as MVO and BBO exhibited less stable results, with higher RMSE and lower correlation metrics, suggesting less effective navigation of the hyperparameter space.

Visualizations such as Figure 7, Figure 8, and Figure 9 further confirm the statistical robustness of the

DOI: https://doi.org/10.54216/JAIM.100203

Received: April 10, 2025 Revised: July 01, 2025 Accepted: September 08, 2025

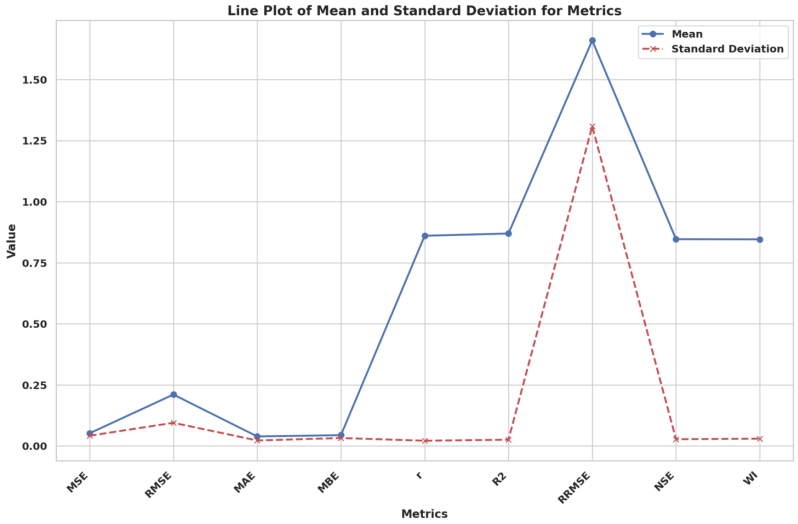

Figure 7: Line Plot of Mean and Standard Deviation for Model Evaluation Metrics

optimization framework. In particular, the tight clustering and near-normal distributions of metrics under FbOA indicate high reliability and generalization capability.

Assessing whether model evaluation metrics follow an approximately normal distribution is an important step in validating statistical assumptions and ensuring the robustness of comparative analyses. Quantile–Quantile (Q–Q) plots provide a graphical means to compare the empirical distribution of observed values with the theoretical quantiles of a reference distribution, typically the normal distribution. As shown in Figure 8, the alignment of points along the diagonal line indicates that most metrics, including MSE, RMSE, MAE, MBE, r, R^2, RRMSE, NSE, and WI, closely follow normality assumptions with minor deviations at the tails. Such consistency suggests that the metrics are not heavily skewed and that variance-based statistical tests can be reliably applied. This verification step strengthens the credibility of the performance evaluation framework, as it confirms that observed deviations across models are statistically interpretable within a standard probabilistic setting.

Understanding the distributional properties of performance metrics provides deeper insights into the consistency and reliability of predictive models. Density plots combined with kernel density estimation (KDE) enable the visualization of both empirical distributions and their smoothed approximations, thereby highlighting underlying statistical patterns. As illustrated in Figure 9, the distributions of metrics such as MSE, RMSE, MAE, and MBE are generally unimodal and symmetric, indicating stability across models. Correlation-based metrics (r, R^2, NSE, and WI) also exhibit well-defined peaks, suggesting strong agreement in predictive performance with relatively low variability. Meanwhile, the relative RMSE (RRMSE) shows a broader distribution, reflecting greater sensitivity to fluctuations compared to other error-based measures. These density–KDE overlays provide a valuable perspective on the statistical behavior of evaluation metrics, ensuring that model comparison is grounded in both central tendency and distributional consistency.

DOI: https://doi.org/10.54216/JAIM.100203

Received: April 10, 2025 Revised: July 01, 2025 Accepted: September 08, 2025

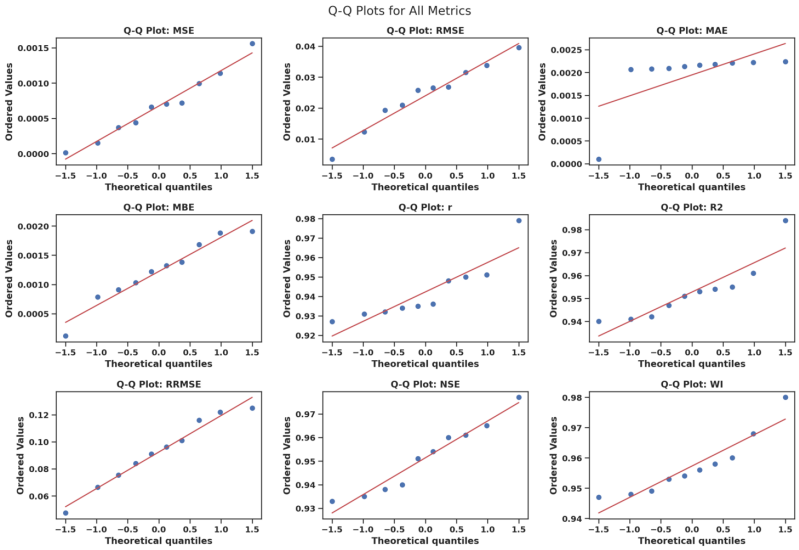

Figure 8: Q–Q Plots for Model Evaluation Metrics

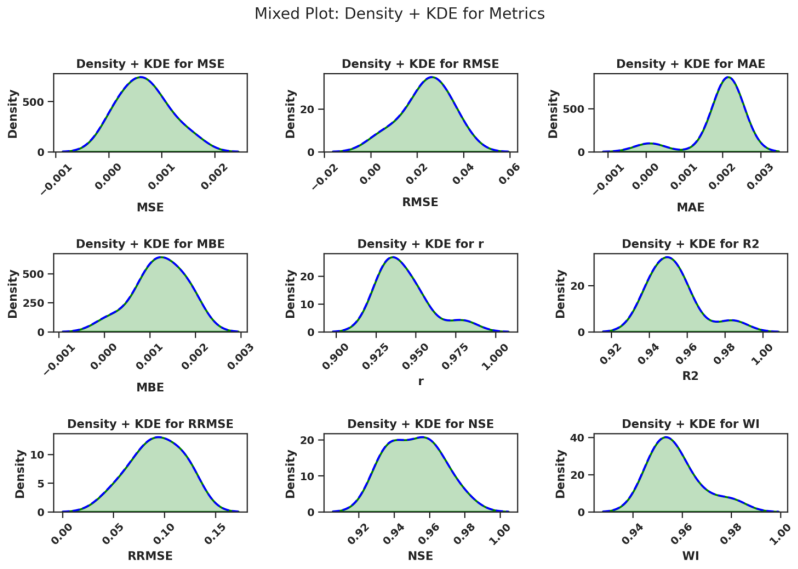

Figure 9: Density and KDE Plots for Model Evaluation Metrics

5 Discussion

The empirical results presented in Section 3 offer valuable insights into the effectiveness of both the machine learning architectures and the metaheuristic algorithms employed for CO_2 emission prediction. This section

DOI: https://doi.org/10.54216/JAIM.100203

Received: April 10, 2025 Revised: July 01, 2025 Accepted: September 08, 2025

discusses the key findings with an emphasis on model behavior, optimization efficiency, generalization performance, and methodological implications.

The baseline evaluation established a clear distinction in the inherent predictive capabilities of the models. Deep learning architectures—particularly those incorporating recurrent structures—consistently outperformed traditional regressors. Among these, the EALSTM model emerged as the most promising due to its attention mechanism, which enabled it to weigh input features according to their temporal and contextual significance. However, despite this structural advantage, even EALSTM exhibited limitations under manually tuned or default hyperparameter settings. This observation underscores the importance of advanced hyperparameter optimization strategies, especially for high-capacity models.

Upon introducing metaheuristic-based optimization, substantial improvements were observed across all evaluated models. The optimized configurations resulted in consistently lower prediction errors and higher correlation-based metrics, reflecting enhanced learning dynamics and better generalization to unseen data. The improvements were particularly significant in the EALSTM model, indicating that its complex architecture benefits more from intelligent parameter tuning than simpler models.

Among the nine optimization algorithms tested, the proposed Football Optimization Algorithm (FbOA) delivered the most favorable outcomes across all key performance indicators. The algorithm demonstrated a superior ability to balance exploration and exploitation, owing to its football-inspired tactical mechanisms such as adaptive velocity modulation, directional passing strategies, and mutation-based diversification. These mechanisms collectively contributed to faster convergence, lower error variance, and improved robustness. Notably, FbOA also exhibited low sensitivity to initial parameter configurations, which further supports its applicability in real-world scenarios where search space boundaries may not be well-defined in advance.

Other optimization algorithms, including DE, GA, and SFS, also showed competitive results but generally required more iterations to converge and were more sensitive to parameter initialization. Algorithms such as QIO and MVO, although conceptually innovative, displayed higher variance and less consistency in converging toward global optima across multiple runs. The performance differences between these methods reinforce the necessity of domain-specific tuning strategies—FbOA's design, grounded in dynamic agent coordination and spatial movement, aligns well with the multi-dimensional nature of the CO_2 emission modeling problem.

An important aspect of this discussion is the practical significance of the observed improvements. Reductions in RMSE and MAE metrics translate directly into more accurate emission estimates, which can influence fuel policy simulations, regulatory assessments, and energy-efficient vehicle design. The gains in NSE and WI further indicate that the optimized models are not only accurate in absolute terms but also reliable in reproducing the variability and trends in real-world data.

Despite the positive findings, some limitations are worth noting. The optimization process, while automated, remains computationally expensive due to the repeated training required across multiple parameter configurations. Future studies may investigate surrogate-assisted optimization or dynamic learning rate adjustment to reduce computational cost. Additionally, while FbOA demonstrated strong generalizability in this dataset, its performance should be evaluated further across different domains and problem settings to confirm its scalability and robustness.

DOI: https://doi.org/10.54216/JAIM.100203

Received: April 10, 2025 Revised: July 01, 2025 Accepted: September 08, 2025

Overall, the discussion highlights the complementary strengths of deep learning architectures and metaheuristic optimization strategies in environmental modeling. The results strongly support the use of tailored, domain-inspired optimizers such as FbOA to maximize the predictive potential of advanced neural models for CO_2 emission forecasting.

6 Conclusion

This study proposed a robust and scalable framework for forecasting vehicle-level CO_2 emissions using deep learning models optimized via metaheuristic algorithms. Leveraging a comprehensive real-world dataset from the Canadian Government, the research systematically evaluated the performance of several neural network architectures—including EALSTM, LSTM, GRU, BiRNN, RNN, and ANN—under baseline and optimized configurations. A suite of nine metaheuristic algorithms was employed to perform automated hyperparameter tuning, with the proposed Football Optimization Algorithm (FbOA) serving as the primary optimization engine. The results demonstrated that hyperparameter optimization significantly enhances model performance across all evaluated metrics. The EALSTM model, in particular, benefited substantially from this optimization due to its architectural complexity and capacity for learning contextual relationships. Among the tested optimizers, FbOA consistently achieved the best performance, outperforming other established algorithms such as DE, GA, and MVO in terms of prediction accuracy, convergence speed, and robustness. The football-inspired strategy of FbOA—incorporating dynamic agent coordination, directional exploration, and adaptive mutation—proved highly effective for navigating high-dimensional search spaces. The improvements observed in key metrics such as RMSE, R^2, NSE, and WI suggest that the proposed optimization framework is not only statistically significant but also practically valuable. More accurate CO_2 emission predictions can directly support environmental policy-making, energy-efficient automotive design, and emissions labeling. Furthermore, the methodological approach is generalizable and could be extended to other environmental modeling tasks, including fuel consumption forecasting, emissions reduction simulations, and lifecycle assessment.

Future research may focus on several promising directions. One area involves the development of hybrid and adaptive optimization strategies that combine FbOA with other metaheuristics or reinforcement learning components to enable dynamic self-adjustment during the search process. Another opportunity lies in embedding the optimized models into real-time edge computing environments, allowing for intelligent CO_2 estimation directly within onboard vehicular systems or traffic management platforms. Additionally, validating the framework on international datasets and across different transport sectors—such as freight, maritime, or aviation—would help establish its cross-domain applicability. Enhancing model interpretability through tools like SHAP values, attention heatmaps, or causal feature attribution would also strengthen trust and transparency in emissions modeling. Lastly, integrating the optimized models into intelligent policy support systems would allow governments and regulatory agencies to simulate and evaluate long-term impacts of taxation, incentive programs, and emissions caps under diverse operational scenarios. In conclusion, this research illustrates the synergistic benefits of combining advanced deep learning architectures with domain-inspired metaheuristic optimization. The proposed FbOA-EALSTM framework represents a powerful and generalizable solution for accurate CO_2 emissions modeling and lays the foundation for future advancements in intelligent, data-driven environmental forecasting.

DOI: https://doi.org/10.54216/JAIM.100203

Received: April 10, 2025 Revised: July 01, 2025 Accepted: September 08, 2025

References

[1] Gazi Mahabubul Alam, Sharia Arfin Tanim, Sumit Kanti Sarker, Yutaka Watanobe, Md. Rashedul Islam, M. F. Mridha, and Kamruddin Nur. Deep learning model based prediction of vehicle co2 emissions with explainable ai integration for sustainable environment. *Scientific Reports*, 15, 01 2025.

[2] Yıldıray Yiğit and Murat Karabatak. A deep reinforcement learning-based speed optimization system to reduce fuel consumption and emissions for smart cities. *Applied Sciences*, 15:1545–1545, 02 2025.

[3] Esra'a Alrashydah, Thaar Alqahtani, and Abdulnaser M. Al-Sabaeei. Emissions of conventional and electric vehicles: A comparative sustainability assessment. *Sustainability*, 17:6839–6839, 07 2025.

[4] Vindula Jayawardana, Baptiste Freydt, Ao Qu, Cameron Hickert, Edgar N. Sánchez, C.Y. Tang, Mark A. Taylor, Blaine Leonard, and Cathy Wu. Mitigating metropolitan carbon emissions with dynamic eco-driving at scale. *Transportation Research Part C Emerging Technologies*, 179:105146–105146, 08 2025.

[5] Ishan Shivansh Bangroo. Ai-based predictive analytic approaches for safeguarding the future of electric/hybrid vehicles. *arXiv (Cornell University)*, 01 2023.

[6] Qi Wang. Towards zero-emission urban mobility: Leveraging ai and lca for targeted interventions. *Building Simulation*, 17:1653–1657, 10 2024.

[7] Juraj Kováč, Peter Malega, Erik Varjú, Jozef Svetlík, and Rudolf Stetulič. Reduction of carbon footprint in mechanical engineering production using a universal simulation model. *Applied Sciences*, 15:5358–5358, 05 2025.

[8] Nima Khodadadi, S. K. Towfek, Ahmed Mohamed Zaki, Amal H. Alharbi, Ehsan Khodadadi, Doaa Sami Khafaga, Laith Abualigah, Abdelhameed Ibrahim, Abdelaziz A. Abdelhamid, and Marwa M. Eid. Predicting normalized difference vegetation index using a deep attention network with bidirectional GRU: a hybrid parametric optimization approach. *International Journal of Data Science and Analytics*, October 2024.

[9] Matteo Böhm, Mirco Nanni, and Luca Pappalardo. Gross polluters and vehicle emissions reduction. *Nature Sustainability*, 5:699–707, 06 2022.

[10] Samuel Akintomide Ajayi, Charles A. Adams, Gift Dumedah, and Atinuke Adebanji. The impact of vehicle engine characteristics on vehicle exhaust emissions for transport modes in lagos city. *Urban Planning and Transport Research*, 12, 02 2024.

[11] F. E. Sapnken, K. R. Hong, H. Chopkap Noume, and J. G. Tamba. A grey prediction model optimized by meta-heuristic algorithms and its application in forecasting carbon emissions from road fuel combustion. *Energy*, 302:131922, 2024.

[12] T. Sujjaviriyasup and K. Pitiruek. A hybridization of modwt-svr-de model emphasizing on noise reduction and optimal parameter selection for prediction of co2 emission in thailand. *Cogent Engineering*, 11(1):2317540, 2024.

[13] M. Zadmirzaei, F. Hasanzadeh, A. Susaeta, and E. Gutiérrez. A novel integrated fuzzy dea–artificial intelligence approach for assessing environmental efficiency and predicting co2 emissions. *Soft Computing*, 28(1):565–591, 2024.

DOI: https://doi.org/10.54216/JAIM.100203

Received: April 10, 2025 Revised: July 01, 2025 Accepted: September 08, 2025

[14] Q. Qiao, H. Eskandari, H. Saadatmand, and M. A. Sahraei. An interpretable multi-stage forecasting framework for energy consumption and co2 emissions for the transportation sector. *Energy*, 286:129499, 2024.

[15] M. Singh and R. K. Dubey. Deep learning model based co2 emissions prediction using vehicle telematics sensors data. *IEEE Transactions on Intelligent Vehicles*, 8(1):768–777, 2023.

[16] W. Wang and J. Wang. Determinants investigation and peak prediction of co2 emissions in china's transport sector utilizing bio-inspired extreme learning machine. *Environmental Science and Pollution Research International*, 28(39):55535–55553, 2021.

[17] M. Emami Javanmard, Y. Tang, Z. Wang, and P. Tontiwachwuthikul. Forecast energy demand, co2 emissions and energy resource impacts for the transportation sector. *Applied Energy*, 338:120830, 2023.

[18] H. Moayedi, A. Mukhtar, N. B. Khedher, I. Elbadawi, M. B. Amara, Q. TT, and N. Khalilpoor. Forecasting of energy-related carbon dioxide emission using ann combined with hybrid metaheuristic optimization algorithms. *Engineering Applications of Computational Fluid Mechanics*, 18(1):2322509, 2024.

[19] H. Bakır, Ü. Ağbulut, A. E. Gürel, G. Yıldız, U. Güvenç, M. E. M. Soudagar, A. T. Hoang, B. Deepanraj, G. Saini, and A. Afzal. Forecasting of future greenhouse gas emission trajectory for india using energy and economic indexes with various metaheuristic algorithms. *Journal of Cleaner Production*, 360:131946, 2022.

[20] M. Almsallti, A. B. Alzubi, and O. R. Adegboye. Hybrid metaheuristic optimized extreme learning machine for sustainability focused co2 emission prediction using globalization-driven indicators. *Sustainability*, 17(15):6783, 2025.

[21] L. K. Foong, V. Blazek, L. Prokop, S. Misak, F. Atamurotov, and N. Khalilpoor. Improve carbon dioxide emission prediction in the asia and oceania (oecd): nature-inspired optimisation algorithms versus conventional machine learning. *Engineering Applications of Computational Fluid Mechanics*, 18(1):2391988, 2024.

[22] S. Ma, X. Wang, S. Cheng, Y. Liu, Y. Wang, and J. Wang. Improving carbon dioxide emission predictions through a hybrid model utilising an advanced sparrow search algorithm. *Environmental Technology*, 46(17):3422–3437, 2025.

[23] A. Bashir, S. I. Haruna, Y. E. Ibrhim, and S. I. Abba. Metaheuristic-enhanced ann framework for compressive strength prediction in concrete with supplementary cementitious materials. *Innovative Infrastructure Solutions*, 10(9):432, 2025.

[24] M. Salehi Sarbijan and J. Behnamian. Multi-product production routing problem by consideration of outsourcing and carbon emissions: particle swarm optimization. *Engineering Optimization*, 53(8):1298–1314, 2021.

[25] S. Ding, J. Ye, and Z. Cai. Multi-step carbon emissions forecasting using an interpretable framework of new data preprocessing techniques and improved grey multivariable convolution model. *Technological Forecasting and Social Change*, 208:123720, 2024.

DOI: https://doi.org/10.54216/JAIM.100203

Received: April 10, 2025 Revised: July 01, 2025 Accepted: September 08, 2025

[26] A. B. Ghorbal, A. Grine, I. Elbatal, E. M. Almetwally, M. M. Eid, and E.-S. M. El-Kenawy. Predicting carbon dioxide emissions using deep learning and ninja metaheuristic optimization algorithm. *Scientific Reports*, 15(1):4021, 2025.

[27] H. Moayedi, A. Mukhtar, S. Alshammari, M. Boujelbene, I. Elbadawi, Q. T. Thi, and M. Mirzaei. Prediction of co2 emission for the central european countries through five metaheuristic optimization techniques helping multilayer perceptron. *Engineering Applications of Computational Fluid Mechanics*, 18(1):2327437, 2024.

[28] J. Booma, P. Anitha, S. Amosedinakaran, and A. Bhuvanesh. Real-time electricity capacity expansion planning using chaotic ant lion optimization by minimizing carbon emission. *Journal of the Chinese Institute of Engineers*, 48(3):239–253, 2025.

[29] Y. Su and C.-C. J. Kuo. On extended long short-term memory and dependent bidirectional recurrent neural network. *Neurocomputing*, 356:151–161, 2019.

[30] Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural Computation*, 9(8):1735–1780, 1997.

[31] Kyunghyun Cho, Bart van Merrienboer, Caglar Gulcehre, Dzmitry Bahdanau, Fethi Bougares, Holger Schwenk, and Yoshua Bengio. Learning phrase representations using rnn encoder-decoder for statistical machine translation. *arXiv preprint arXiv:1406.1078*, 2014.

[32] Mike Schuster and Kuldip K. Paliwal. Bidirectional recurrent neural networks. *IEEE Transactions on Signal Processing*, 45(11):2673–2681, 1997.

[33] Shun-Ichi Amari. Learning patterns and pattern sequences by self-organizing nets of threshold elements. *IEEE Transactions on Computers*, C-21(11):1197–1206, 1972.

[34] Frank Rosenblatt. The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological Review*, 65(6):386–408, 1958.

[35] E.-S. M. El-Kenawy, F. H. Rizk, A. M. Zaki, M. E. Mohamed, A. Ibrahim, A. A. Abdelhamid, N. Khodadadi, E. M. Almetwally, and M. M. Eid. Football optimization algorithm (fboa): A novel metaheuristic inspired by team strategy dynamics. *Journal of Artificial Intelligence and Metaheuristics*, pages 21–38, 2024.

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

[37] Colin Reeves. Genetic algorithms. In *Handbook of Metaheuristics*, volume 146, pages 109–139. 2010.

[38] O. Bingöl, U. Guvenc, S. Duman, and S. Paçaci. Stochastic fractal search with chaos. In *2017 International Artificial Intelligence and Data Processing Symposium (IDAP)*, pages 1–6, 2017.

[39] R. Venkata Rao. Jaya: A simple and new optimization algorithm for solving constrained and unconstrained optimization problems. *International Journal of Industrial Engineering Computations*, 7:19–34, 2016.

[40] K.-H. Han and J.-H. Kim. Quantum-inspired evolutionary algorithm for a class of combinatorial optimization. *IEEE Transactions on Evolutionary Computation*, 6(6):580–593, 2002.

DOI: https://doi.org/10.54216/JAIM.100203

Received: April 10, 2025 Revised: July 01, 2025 Accepted: September 08, 2025

[41] X.-S. Yang. A new metaheuristic bat-inspired algorithm. In J. R. González, D. A. Pelta, C. Cruz, G. Terrazas, and N. Krasnogor, editors, *Nature Inspired Cooperative Strategies for Optimization (NICSO 2010)*, pages 65–74. Springer, 2010.

[42] Dan Simon. Biogeography-based optimization. *IEEE Transactions on Evolutionary Computation*, 12(6):702–713, 2008.

[43] S. Mirjalili, S. Mirjalili, and A. Hatamlou. Multi-verse optimizer: a nature-inspired algorithm for global optimization. *Neural Computing and Applications*, 27, 2015.

DOI: https://doi.org/10.54216/JAIM.100203

Received: April 10, 2025 Revised: July 01, 2025 Accepted: September 08, 2025