



Optimizing CO₂ Emission Forecasting from Vehicles Using Deep Learning and the Football Optimization Algorithm

Omnia M. Osama^{1,*}

El-Sayed M. El-Rabaie²

Marwa M. Eid^{3,4}

¹ Department of Communications and Electronics, Delta Higher Institute of Engineering and 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: Omnia.osama@dhiet.edu.eg · srabie1@yahoo.com · mmm@ieee.org

Received: April 10, 2025 Revised: July 1, 2025 Accepted: September 8, 2025 ★ Corresponding author

ABSTRACT

Accurate prediction of CO₂ emissions from vehicles is essential for environmental regulation and sustainable transport design. Existing models often suffer from limited accuracy because of suboptimal hyperparameter configurations. This study enhances CO₂ emission forecasting by combining deep learning with metaheuristic optimization. An attention-based Encoder LSTM (EALSTM) model is trained on Canadian vehicle-emissions data, with hyperparameters tuned using the 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 FbOA–EALSTM model yields an RMSE of 0.00349, MAE of 0.00010, and R^2 of 0.984, outperforming the alternatives reported in the study. These results demonstrate the potential of domain-inspired metaheuristics to improve prediction accuracy. The proposed framework offers a scalable solution for emissions modeling and supports data-driven environmental policy and intelligent vehicle technologies.

Keywords: CO₂ emissions ▪ Metaheuristic optimization ▪ Football Optimization Algorithm (FbOA) ▪ Encoder LSTM (EALSTM) ▪ Transportation sustainability

1. INTRODUCTION

Global warming and climate change remain among the most pressing environmental challenges, driven primarily by rising atmospheric concentrations of greenhouse gases (GHGs) [1, 2]. Carbon dioxide (CO₂) is a central concern because of its large emissions volume, long atmospheric lifetime, and dominant contribution to human-induced warming. Road transportation is a major anthropogenic source of CO₂ and contributes substantially to national and global GHG inventories. Increasing vehicle ownership, rapid urbanization, expanding road infrastructure, and growing reliance on logistics and mobility have intensified the environmental burden of vehicle emissions. This trend is evident in emerging

economies, while developed nations also face continuing challenges in curbing emissions from personal and commercial fleets.

The urgency of climate action has increased interest in data-driven approaches for quantifying and mitigating vehicle emissions. Recent Canadian vehicle-level analyses use machine learning to estimate emissions [3], while eco-driving control methods can reduce emissions in connected-vehicle traffic [4]. Accurate, scalable emission models inform environmental regulations, carbon-target compliance, and vehicle-level emission standards. Manufacturers can use them to evaluate emissions during vehicle design, while consumers and urban planners can use estimates to support eco-labeling,

congestion pricing, low-emission zones, and emissions-based tolling [5]. Forecasting is also important for assessing alternative fuels, hybrid and electric drivetrains, and changing mobility patterns. Lifecycle assessments and scenario analyses allow stakeholders to examine decarbonization pathways, energy-policy effects, and predictive emission control in intelligent transportation systems [6].

Machine-learning (ML) methods are increasingly used to model the nonlinear relationships found in large, heterogeneous datasets. Unlike conventional regression approaches that rely on linearity or handcrafted features, deep-learning models can discover latent patterns and dependencies. Recurrent neural networks (RNNs), including Long Short-Term Memory (LSTM) networks [7, 8], Gated Recurrent Units (GRUs) [9], and bidirectional RNNs (BiRNNs), are widely used for time-dependent relationships. Attention mechanisms extend these architectures by weighting input sequences according to relevance, improving predictive fidelity in Encoder-Attention LSTM (EALSTM) models.

Applying deep learning to vehicle emissions nevertheless presents challenges. Vehicle datasets combine categorical and numerical attributes describing powertrain configurations, transmissions, drivetrains, fuel, and manufacturers. These features may be interrelated, introducing redundancy or noise and weakening generalization when handled poorly [10]. Deep-learning performance is also sensitive to hyperparameters such as learning rate, hidden-layer size, batch size, dropout, and activation functions. Inadequate tuning can cause underfitting, overfitting, or poor convergence, particularly across different vehicle classes or time periods. At the same time, evolving fleets—with electric vehicles, hybrid systems, and alternative fuels—require models that generalize across changing input spaces [11].

Research gap. Most existing vehicle-emission studies do not systematically optimize hyperparameters. Manual tuning and basic grid searches can be inefficient and may settle at local optima, especially for architectures with high-dimensional search spaces. Few studies also assess models on nationally representative, multi-year datasets spanning diverse vehicle specifications, fuels, and drivetrains. Scalable and generalizable CO₂ forecasting frameworks therefore remain an important research need.

Objectives and contributions. This study develops a generalizable deep-learning framework for vehicle-level CO₂ forecasting. It compares LSTM, GRU, BiRNN, RNN, ANN, and EALSTM architectures; applies multiple metaheuristic optimizers for automated hyperparameter tuning, including the Football Optimization Algorithm (FbOA); and evaluates the framework on a multi-year Canadian government dataset. Its contributions are a comparative benchmark of deep-learning architectures, a unified architecture-agnostic optimization framework, use of a national-scale vehicle dataset, introduction of a football-inspired optimizer to improve convergence and stability, and an analysis of the interaction between model architecture and hyperparameter sensitivity.

Organization. Section 2 reviews related studies. Section 3 describes the dataset, preprocessing, models, optimizers, and evaluation metrics. Section 4 presents baseline and optimized results, Section 5 discusses their implications and limitations, and Section 6 concludes the paper and outlines future work.

2. LITERATURE REVIEW

Predicting transportation-sector CO₂ emissions is important for climate policy, regulation, and sustainable development. Researchers have proposed computational and hybrid approaches to improve the precision, adaptability, and interpretability of forecasts across national and regional contexts. Table 1 provides a comparative overview of representative methods and findings across the reviewed studies.

In Cameroon, Sapnken et al. [12] developed an Optimized Wavelet Transform Hausdorff Grey Model, OWTGGM(1,N), to improve the stability of classical GM(1,N) models and reduce sensitivity to input noise. Wavelet filtering, fractional Hausdorff derivatives, and Rao-3 parameter optimization improved road-transport emission forecasts. In Thailand, Sujjaviriyasup and Pitiruek [13] combined Maximal Overlap Discrete Wavelet Transform (MODWT), Support Vector Regression (SVR), and Differential Evolution (DE). Their hybrid outperformed ARIMA and Holt benchmarks, illustrating the value of wavelets and metaheuristics for short- and long-term emission dynamics.

Zadmirzaei et al. [14] combined a fuzzy undesirable slacks-based measure (FUNSBM) within Data Envelopment Analysis (DEA) with Artificial Neural Networks (ANN), Particle Swarm Optimization (PSO), and Artificial Immune Systems (AIS). Their FUNSBM-AIS configuration showed substantial CO₂ reduction potential. For the United Kingdom, a SHapley Additive exPlanations (SHAP)-enhanced ML framework identified road carbon intensity as a stronger predictor than population or GDP per capita, while improving interpretability [15]. At the vehicle level, an LSTM using On-Board Diagnostics (OBD-II) data enabled scalable emission monitoring without dedicated emission sensors [16].

In China, an Extreme Learning Machine optimized with Manta Ray Foraging Optimization (MRFO) was used to forecast transport emissions and examine peak-emission scenarios; the findings highlighted vehicle electrification and projected peaks between 2035 and 2043, depending on policy intensity [17]. A Canadian hybrid mathematical and ML framework forecast transportation energy demand and emissions, projecting a 50.02% increase by 2048 and identifying oil consumption as an influential factor [18]. For G8 countries, multilayer perceptrons (MLPs) optimized with Teaching-Learning-Based Optimization (TLBO) and Vortex Search (VS) achieved strong generalization, with TLBO providing the best training-testing trade-off [19]. In India, Marine Predators Algorithm (MPA) and Lightning Search Algorithm (LSA) were used to forecast GHG emissions; the study projected a 2.87-fold increase by 2050 and found MPA most accurate [20].

For Indonesia, system-dynamics modeling combined with the STIRPAT model evaluated carbon-tax scenarios and projected an earlier emissions peak under a higher-tax scenario [21]. For OECD Asia-Pacific countries, Moth Flame Optimization

Table 1. Summary of reviewed studies.

Ref.	Focus area	Methodology	Key findings and contributions
[12]	Road transport emissions in Cameroon	OWTHGM(1,N) with wavelet filtering and Hausdorff derivative	Improved prediction accuracy; supports policy planning in Cameroon.
[13]	Emission monitoring in Thailand	MODWT–SVR–DE hybrid model	Outperforms ARIMA, Holt, and SES; suitable for emission-control policies.
[14]	Environmental efficiency in Iran	FUNSBM with ANN, PSO, and AIS	FUNSBM–AIS reduces CO ₂ emissions by 53%; evaluates inefficiency in DMUs.
[15]	UK transportation sector	SHAP-enhanced ML model	Highlights road carbon intensity and improves interpretability.
[16]	Real-time vehicle emission estimation	LSTM with OBD-II data	Scalable monitoring without onboard emission sensors.
[17]	China transport emission peaks	MRFO–ELM hybrid model	Predicts emission peaks under different scenarios and stresses electrification.
[18]	Canada transport emission forecast	Mathematical and ML hybrid	Predicts a 50.02% emission rise by 2048; oil is an influential factor.
[19]	CO ₂ prediction in G8 nations	MLP with nature-inspired optimization	TLBO and VS show strong accuracy and generalization.
[20]	India's GHG emissions trajectory	MPA, LSA, SOS, and BSA	CO ₂ projected to rise 2.5–2.87 times by 2050; MPA is most accurate.
[21]	Indonesia carbon-tax scenarios	System dynamics with STIRPAT	Scenario analysis projects an earlier emissions peak under a carbon tax.
[22]	Emissions in OECD Asia-Pacific	MFO/SHO/GOA with MLP, GP, and RF	MFO–MLP achieves top accuracy and supports policy design.
[23]	Carbon emissions in Chinese cities	Interpretable machine-learning comparison	Extra-Trees performs best; energy use is a key driver.
[24]	Recycled-concrete sustainability	OMA–MBM for aggregate strength and mixture emissions	Predicts strength and identifies lower-emission mixes.
[25]	Green production-routing	MILP with PSO	Optimizes cost and emissions; PSO outperforms VNS.
[26]	Provincial CO ₂ emissions in China	Grey convolution model with UNAGO	Outperforms traditional models across multiple horizons.
[27]	Light-duty vehicle CO ₂ prediction	Regression models using UK VCA data	Decision-tree regression achieves the lowest MAE.
[28]	Urban CO ₂ in Central Europe	MLP with BBO, SCE, and FSA	SCE and BBO produce the best metrics ($R^2 = 0.9999$).
[29]	Tamil Nadu electricity planning	HSHP SO-based generation expansion	Balances modeled emissions, cost, and reliability.

(MFO) was coupled with MLP, Genetic Programming (GP), and Random Forest (RF), producing reported R^2 values above 0.999 [22]. In China, an interpretable machine-learning comparison across 254 cities found Extra-Trees to be the most accurate model and identified energy consumption as a major emissions driver [23].

Emission-aware AI has also been applied beyond transport. An OMA–MBM inference framework predicts recycled-concrete aggregate strength and identifies mixtures with lower CO₂ emissions [24]. For green production-routing, a mixed-integer linear programming model solved with PSO optimized both cost and emissions [25]. A grey convolution model with the UNAGO operator improved multi-step provincial CO₂ forecasts in China, reporting MAPE below 10% [26]. For UK light-duty vehicles, six regression models were evaluated using Vehicle Certification Agency data; decision-tree regression achieved the lowest MAE [27]. In Central Europe, MLP models tuned using Biogeography-Based Optimiza-

tion (BBO), Shuffled Complex Evolution (SCE), and Firefly Swarm Algorithm (FSA) achieved strong results, with SCE and BBO reporting $R^2 = 0.9999$ [28]. Finally, a hybrid spotted hyena–particle swarm optimization method was applied to generation expansion planning for older power plants in Tamil Nadu, balancing emissions, cost, and reliability [29].

3. MATERIALS AND METHODS

The proposed predictive framework integrates real-world data acquisition, preprocessing, model development, and hyperparameter optimization. Machine-learning models capture nonlinear relationships between vehicle specifications and CO₂ output, while metaheuristic algorithms tune model parameters to improve stability and generalization. The following subsections describe the dataset and experimental methodology.

3.1 Dataset Description

The dataset comes from the Government of Canada's open-data repository and covers CO₂ emissions from passenger vehicles over seven consecutive years. It contains 7,385 instances, each representing a vehicle-model configuration, and 12 attributes describing physical, mechanical, and fuel-related properties. These include drivetrain, transmission, fuel type, engine displacement, cylinder count, city and highway fuel consumption (L/100 km), and combined CO₂ emissions (g/km).

Drivetrain abbreviations include 4WD/4X4 (four-wheel drive), AWD (all-wheel drive), SWB (short wheelbase), LWB (long wheelbase), and EWB (extended wheelbase). Transmission codes include A (automatic), M (manual), AM (automated manual), AS (automatic with select shift), and AV (continuously variable), commonly followed by a gear count from 3 to 10. Fuel codes are X for regular gasoline, Z for premium gasoline, D for diesel, E for ethanol blends such as E85, and N for natural gas. Categorical variables are transformed during preprocessing. Combined fuel consumption is supplied in L/100 km and miles per gallon (mpg); L/100 km is used for training to align with the target variable.

The target is the tailpipe CO₂ emission rate under combined city and highway driving, weighted 55% city and 45% highway, and measured in g/km. The government-sourced records are treated as reliable, but preprocessing addresses incomplete records and numerical consistency. Data are randomly divided into training and test subsets using an 80/20 split. No external dataset is included. The dataset was selected for its official status, breadth, and structured encoding of vehicle attributes, making it suitable for evaluating model architectures and optimization on real-world emissions data.

3.2 Data Preprocessing

Preprocessing was applied consistently to the training and testing data to improve quality and avoid leakage. Missing values were checked across columns; incomplete records were removed, and no imputation was needed because missing data were sparse. Drivetrain, fuel, and transmission variables were encoded using label encoding where ordinal relationships or embeddings were appropriate, and one-hot encoding for nominal features. Numerical attributes—including engine displacement, cylinder count, and fuel consumption—were scaled to the range [0, 1] using Min-Max scaling. The target was kept in its original units for interpretability, although it could be rescaled internally during loss computation. The data split used a fixed random seed; no stratification was applied because the target is continuous. The test set was reserved for final evaluation.

Correlation analysis (Figure 1) showed strong associations between engine size, cylinder count, and CO₂ emissions. City, highway, and combined fuel-consumption measures were positively correlated with emissions, whereas fuel efficiency in mpg was negatively correlated. These patterns indicate that larger engines and higher fuel consumption are associated with greater emissions.

Principal Component Analysis (PCA) was used to examine latent structure in the vehicle attributes. The first two components show clustering across fuel types (Figure 2), with

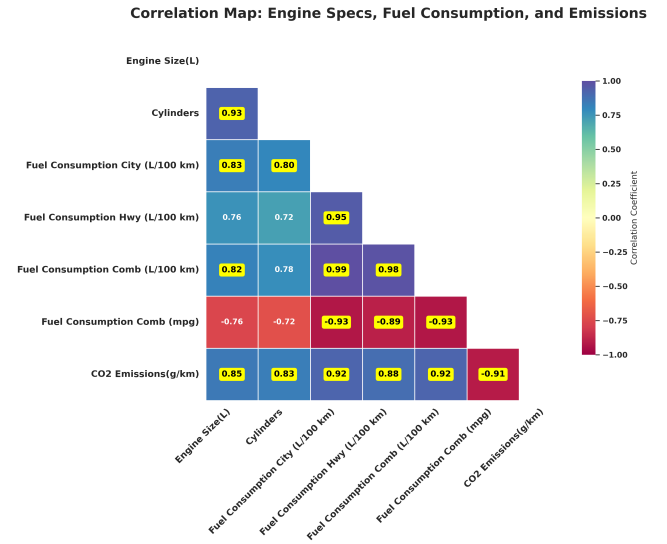


Figure 1. Correlation map of engine specifications, fuel consumption, and CO₂ emissions.

differences between gasoline (Z), diesel (D), ethanol (E), and natural gas (N). Figure 3 compares emissions across vehicle classes: smaller classes generally have lower medians and narrower distributions, while standard SUVs and vans show higher values; passenger vans approach a median of 400 g/km. These patterns motivate vehicle-class-sensitive emission policies.

3.3 Machine-Learning Models

The modeling strategy uses sequence-based deep-learning architectures to represent nonlinear interactions among vehicle attributes. The proposed Encoder Attention-based Long Short-Term Memory (EALSTM) network adds a trainable attention mechanism to an encoder-decoder structure. Attention assigns dynamic importance weights to input steps, allowing the model to learn inter-feature relevance. Although vehicle attributes such as transmission, drivetrain, engine size, and fuel consumption are not inherently temporal, they are treated as a structured pseudo-sequence so the model can learn latent interactions—for example, how the effect of fuel type may depend on transmission or fuel efficiency.

3.3.1 EALSTM Model Architecture

Let $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_T]$ be the encoded input sequence for one vehicle, where T is the number of attributes and $\mathbf{x}_t \in \mathbb{R}^d$ is the embedding at step t . The encoder updates its hidden and cell states as

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

The attention weights and context vector are

$$\alpha_t = \frac{\exp(e_t)}{\sum_{k=1}^T \exp(e_k)}, \quad (2)$$

$$e_t = \mathbf{v}^\top \tanh(\mathbf{W}_h \mathbf{h}_t + \mathbf{b}).$$

$$\mathbf{c} = \sum_{t=1}^T \alpha_t \mathbf{h}_t. \quad (3)$$

Here, α_t is the attention assigned to step t and $\mathbf{c}$ is the weighted context passed to the output layer. The predicted

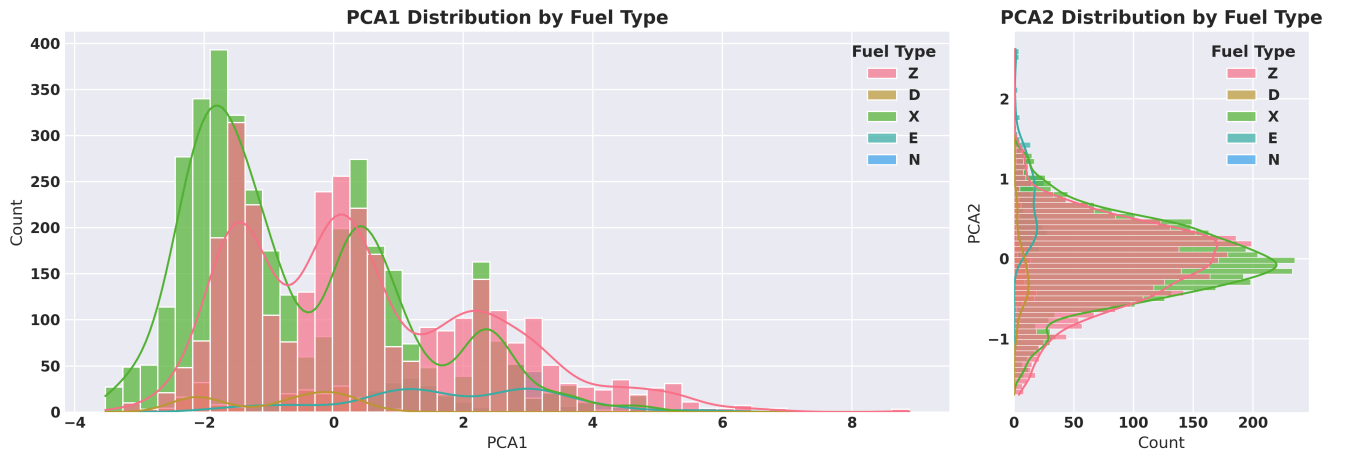


Figure 2. Distribution of PCA1 and PCA2 by fuel type.

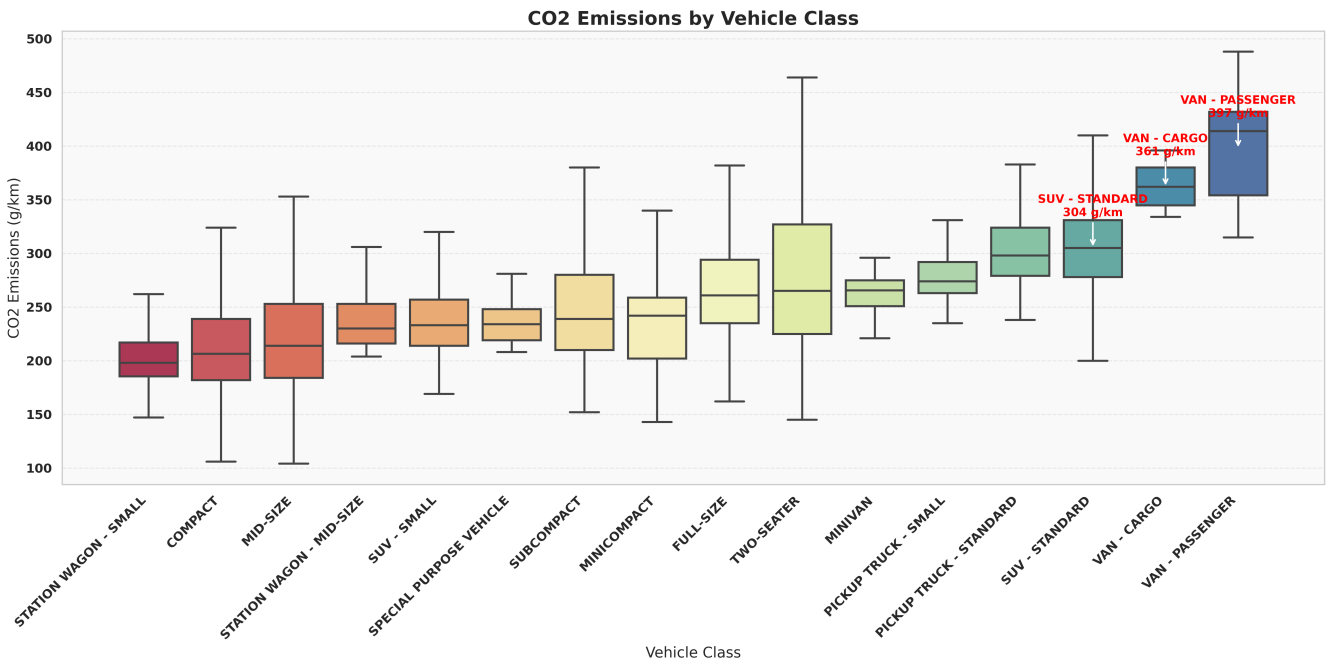


Figure 3. CO₂ emissions distribution by vehicle class.

emissions are

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

where σ is typically linear for regression, and $\mathbf{W}_o$ and $\mathbf{b}_o$ are output weights and bias. Attention allows the model to emphasize influential features and provides an indication of their relative relevance [30].

3.3.2 Comparative Models

The benchmark architectures are:

1. **LSTM**: a recurrent architecture with memory cells designed to capture long-term dependencies and mitigate vanishing gradients [7].
2. **GRU**: a lower-parameter alternative to LSTM that combines gating operations and can be more computationally efficient [31].
3. **BiRNN**: processes the input in forward and backward directions to incorporate context from the full sequence [32].
4. **RNN**: a basic sequence model that uses hidden states to represent dependencies; it serves as a baseline despite gradient-instability limitations on longer sequences [33].

5. **ANN**: a feedforward network with fully connected layers and nonlinear activations. It models nonlinear input–output relationships but does not represent sequence dependencies [34].

Models use consistent training conditions, loss functions, optimizers, and regularization. Hyperparameters are optimized independently for fair comparison. Performance is assessed using the regression metrics summarized in Table 2.

3.4 Metaheuristic Algorithms

Population-based metaheuristics navigate complex, non-convex hyperparameter spaces without requiring gradient information. FbOA is the primary optimizer, and its performance is compared with eight established algorithms.

3.4.1 Football Optimization Algorithm

The Football Optimization Algorithm (FbOA) is inspired by football tactics. Each population member represents a player, and interactions such as passing, movement, and positioning model search-space exploration. Short passes represent local search and long passes represent global search, balancing exploration and exploitation [35].

Table 2. Regression evaluation metrics used in this study.

Metric	Mathematical definition
Mean Squared Error (MSE)	$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$
Root Mean Squared Error (RMSE)	$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$
Mean Absolute Error (MAE)	$MAE = \frac{1}{n} \sum_{i=1}^n y_i - \hat{y}_i $
Mean Bias Error (MBE)	$MBE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)$
Correlation coefficient (r)	$r = \frac{\sum_i (y_i - \bar{y})(\hat{y}_i - \bar{\hat{y}})}{\sqrt{\sum_i (y_i - \bar{y})^2} \sqrt{\sum_i (\hat{y}_i - \bar{\hat{y}})^2}}$
Coefficient of determination (R^2)	$R^2 = 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \bar{y})^2}$
Relative RMSE (RRMSE)	$RRMSE = \frac{RMSE}{\bar{y}} \times 100$
Nash–Sutcliffe Efficiency (NSE)	$NSE = 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \bar{y})^2}$
Willmott's Index of Agreement (WI)	$WI = 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (\hat{y}_i - \bar{y} + y_i - \bar{y})^2}$

Exploration. The velocity of agent n is modeled as

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

where $F_{\max}$ and $F_{\min}$ are maximum and minimum forces, F_{ext} is external influence, and F_{best} is the best-known force or solution. Coefficients a_i , a_j , b_x , and b_y control movement; the cosine term modulates movement over iterations.

Exploitation. Local refinement is expressed as

$$F_b(S_{t+1}) = F_i + z_3 F_b(S_t) + K \sin \left(\frac{\pi}{\text{Iteration}} \right), \quad (6)$$

where F_i is the current position, $F_b(S_t)$ is the agent state at iteration t , z_3 controls balance, and K modulates the exploitation shift.

Mutation. To reduce premature convergence, a mutation operator is applied:

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

where a_q is a scaling factor and K controls mutation intensity. This term introduces diversity by moving agents toward unexplored regions.

Algorithm 1: Football Optimization Algorithm (FbOA)

- 1: Initialize a population of agents at random positions.
- 2: Evaluate the initial fitness of every agent.
- 3: While the termination condition is not met:
 - a) For each agent, compute velocity using Equation (5).
 - b) Update position using the exploration or exploitation rule.

- c) Apply mutation using Equation (7) and evaluate the new fitness.
- d) If fitness improves, update the best-known solution.

4: Return the best solution found.

FbOA tunes the EALSTM learning rate, batch size, number of hidden units, dropout ratio, and training epochs. Its design is intended to provide fast convergence and robust generalization.

3.4.2 Comparative Metaheuristic Algorithms

Eight additional optimizers were run independently under the same experimental settings [36, 37, 38, 39, 40, 41, 42, 43]:

1. **Differential Evolution (DE):** perturbs candidates using scaled differences among population members and supports global search in continuous spaces [36].
2. **Genetic Algorithm (GA):** simulates selection, crossover, and mutation to evolve candidate solutions [37].
3. **Stochastic Fractal Search (SFS):** uses fractal diffusion and Gaussian walks to balance global exploration and local search [38].
4. **JAYA:** moves candidates toward the best solution and away from the worst without algorithm-specific control parameters [39].
5. **Quantum-Inspired Optimizer (QIO):** uses quantum-inspired superposition and rotation gates to maintain diversity and exploration [40].
6. **Bat Algorithm (BA):** mimics echolocation, using adaptive loudness and pulse rates to switch between local and global search [41].
7. **Biogeography-Based Optimizer (BBO):** shares solution features among habitats to promote migration, diversity, and exploitation [42].

8. **Multiverse Optimization (MVO)**: models candidate solutions as universes connected through white holes, black holes, and wormholes [43].

Each optimizer tunes the same hyperparameter set. Their predictive performance is compared in Section 4.

3.5 Evaluation Metrics

Nine metrics quantify prediction error, bias, correlation, and agreement: 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). Their definitions are shown in Table 2, where y_i is the observation, $\hat{y}_i$ the prediction, and $\bar{y}$ and $\bar{\hat{y}}$ their respective means.

4. RESULTS

All models were evaluated using the metrics in Table 2, under consistent preprocessing, data splits, and validation procedures. Results first compare deep-learning architectures with baseline hyperparameters and then examine metaheuristic optimization, with particular focus on FbOA.

4.1 Baseline Deep-Learning Model Performance

EALSTM, LSTM, GRU, BiRNN, RNN, and ANN were trained and evaluated using an 80/20 training–testing split. EALSTM achieved the best baseline results, with RMSE 0.08621, MAE 0.00944, MSE 0.00743, R^2 0.90589, NSE 0.88972, and WI 0.88690. LSTM and GRU followed. RNN and ANN performed less well; the RNN had RMSE 0.28130 and MAE 0.05085, while ANN had RMSE 0.34466 and MAE 0.07499. The baseline values are reported in Table 3.

The swarm-overlaid boxplots in Figure 4 show tightly clustered EALSTM metrics and broader spreads and outliers for RNN and ANN. The metric heatmap in Figure 5 provides a consolidated comparison and highlights EALSTM’s baseline advantage. Figure 6 shows the relationship between MAE and RMSE: models with lower average absolute error also tend to have lower RMSE, although RMSE is more sensitive to extreme deviations.

4.2 Optimized Model Analysis

The EALSTM hyperparameters were tuned using nine optimizers: DE, GA, SFS, JAYA, QIO, BA, BBO, MVO, and FbOA. Table 4 reports the values available in the source study. FbOA–EALSTM achieved RMSE 0.00349, MAE 0.00010, and $R^2 = 0.984$, outperforming the DE–EALSTM and GA–EALSTM rows shown in the table.

Figure 7 compares mean metric values with standard deviations. Correlation-based indicators remain high, while error measures are comparatively low; RRMSE has the greatest variability, whereas MAE and MBE show lower variability. The authors attribute FbOA’s comparative advantage to its exploration–exploitation mechanism, while MVO and BBO are described as less stable. Figures 8 and 9 show Q–Q plots and density/KDE distributions for the evaluation metrics. The study reports that most metrics are approximately aligned with normality, with minor tail deviations; RRMSE displays a broader distribution.

5. DISCUSSION

The results distinguish the predictive capabilities of the evaluated architectures. Recurrent models generally outperform the feedforward ANN, and EALSTM is the strongest baseline, consistent with its attention mechanism’s ability to weight input features by contextual relevance. Its baseline performance nevertheless leaves room for improvement, emphasizing the importance of hyperparameter optimization for high-capacity models.

Metaheuristic tuning is reported to lower prediction errors and increase correlation-based metrics, with particularly strong improvement for EALSTM. Among the nine optimizers, FbOA achieved the best values in the reported comparison. Its adaptive velocity, directional passing, and mutation mechanisms are proposed to balance exploration and exploitation, converge efficiently, and reduce sensitivity to initial configurations. DE, GA, and SFS were competitive but reportedly required more iterations or showed greater sensitivity to initialization; QIO and MVO were described as more variable.

The practical value of lower RMSE and MAE lies in more accurate emission estimates for policy simulations, regulatory assessment, and energy-efficient vehicle design. Higher NSE and WI indicate stronger agreement with observed variability and trends. The optimization process remains computationally expensive because it requires repeated training over parameter configurations. Surrogate-assisted optimization and dynamic learning-rate adjustment could reduce that cost. FbOA should also be tested on additional datasets and domains to establish scalability and robustness. Overall, the findings support combining deep-learning architectures with domain-inspired optimization for vehicle-emission forecasting.

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

Swarm Plot Overlayed on Box Plot for Individual Metrics with Mean and STD

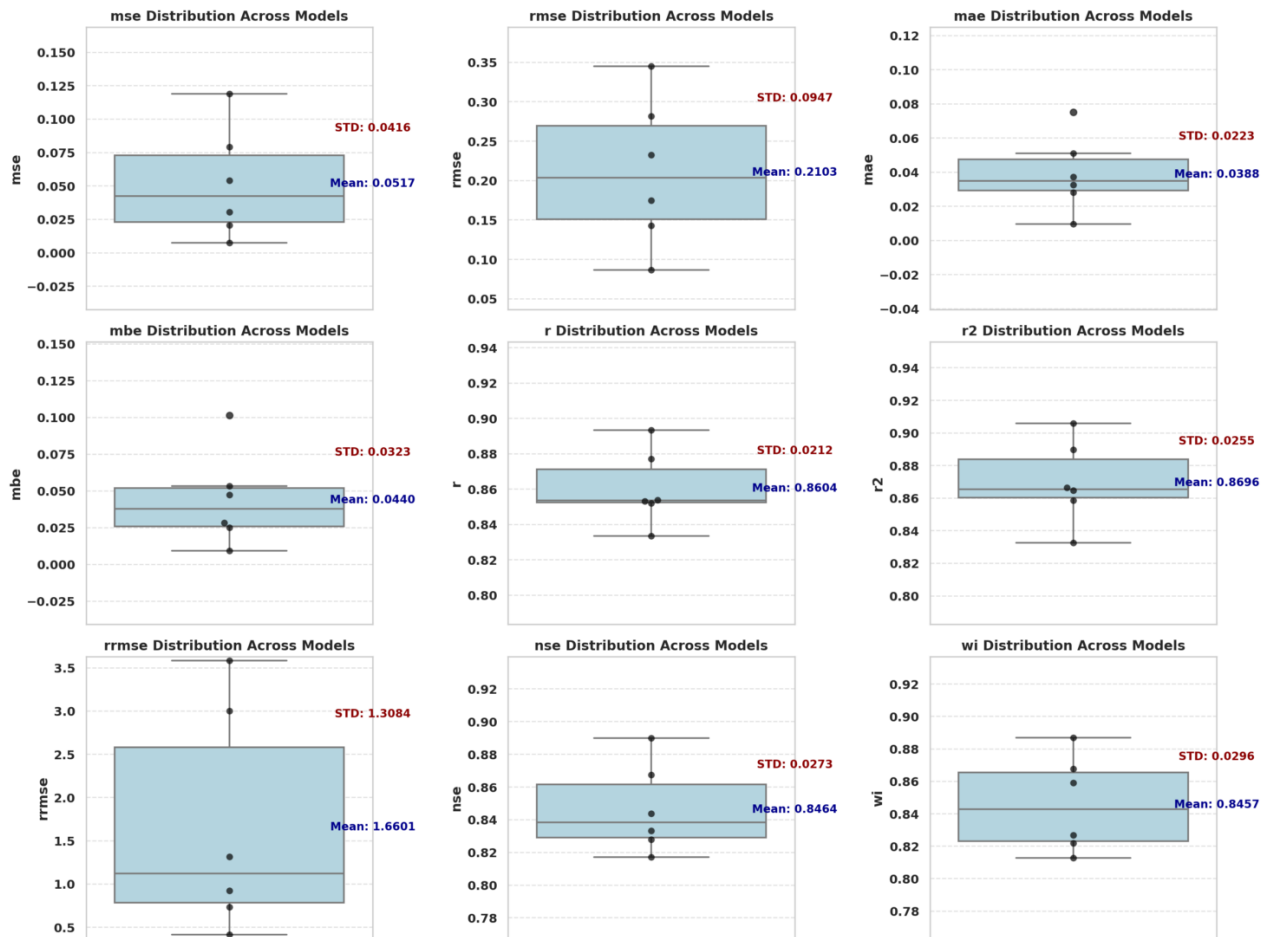


Figure 4. Swarm and boxplot visualization of model-performance metrics with mean and standard deviation.

Table 4. Performance of EALSTM after hyperparameter optimization.

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



Figure 5. Heatmap of comparative model metrics before optimization.

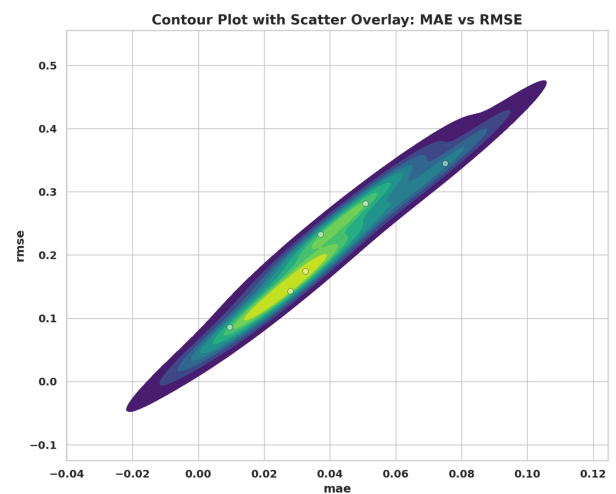


Figure 6. Contour plot with scatter overlay showing the relationship between MAE and RMSE.

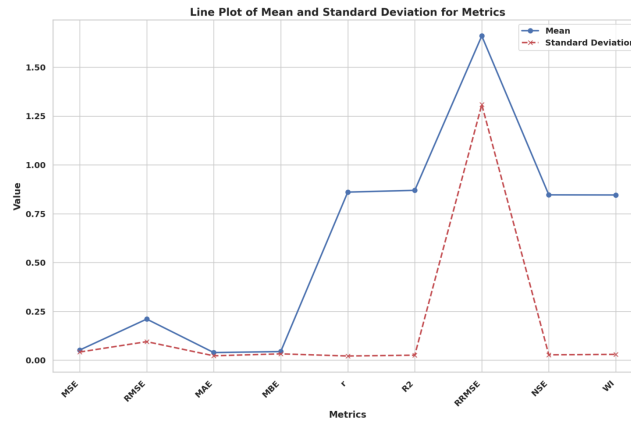


Figure 7. Line plot of mean and standard deviation for model-evaluation metrics.

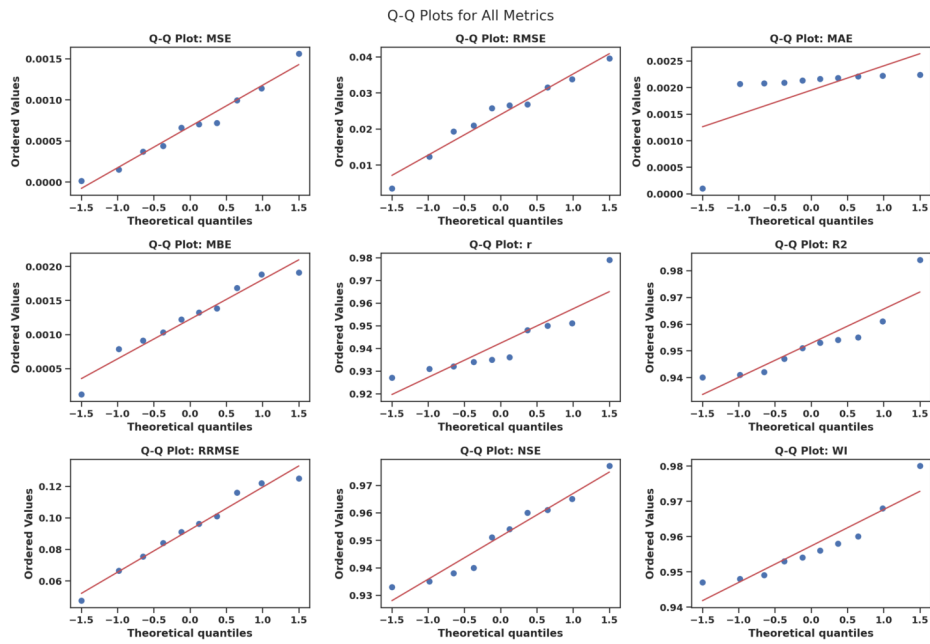


Figure 8. Q-Q plots for model-evaluation metrics.

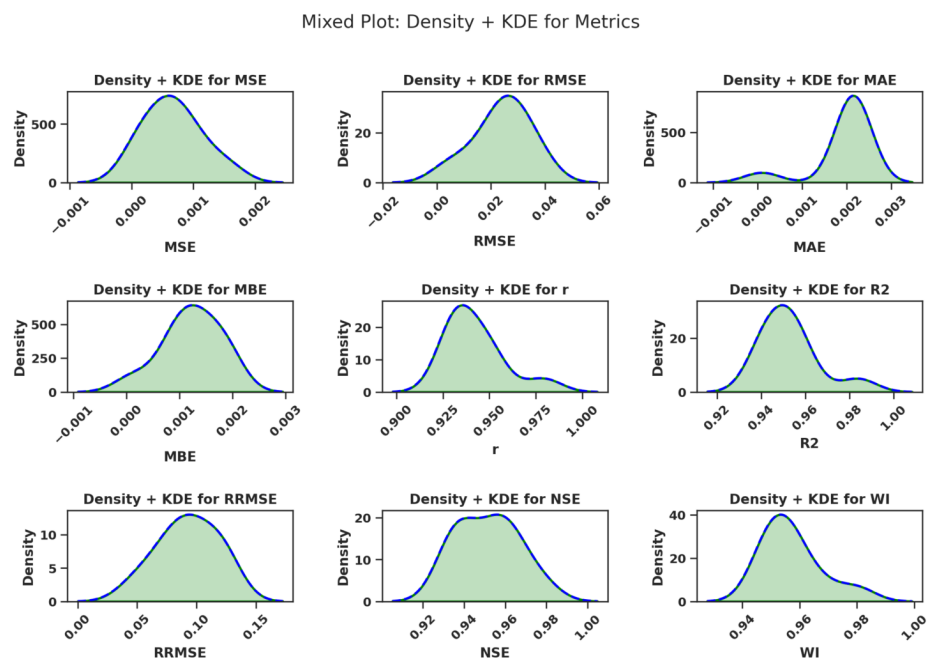


Figure 9. Density and kernel-density-estimation plots for model-evaluation metrics.

6. CONCLUSION

This study presents a framework for vehicle-level CO₂ emission forecasting using deep-learning models optimized by metaheuristic algorithms. A Canadian government dataset was used to compare EALSTM, LSTM, GRU, BiRNN, RNN, and ANN under baseline and optimized configurations. Nine metaheuristics were evaluated, with the Football Optimization Algorithm (FbOA) as the primary proposed optimizer.

Hyperparameter optimization improved the reported performance, particularly for EALSTM. FbOA achieved the best reported results, including RMSE 0.00349, $R^2 = 0.984$, NSE 0.977, and WI 0.980. Its football-inspired coordination, directional exploration, and adaptive mutation are presented as effective for navigating the hyperparameter search space. More accurate emission predictions can inform policy, vehicle design, and emission labeling, and the approach may extend to fuel-consumption forecasting and lifecycle assessment.

Future work may combine FbOA with other metaheuristics or reinforcement learning, deploy optimized models in edge-computing vehicle systems, and validate the framework on international datasets and other transport sectors such as freight, maritime, and aviation. Further directions include improving interpretability with SHAP, attention visualizations, or causal attribution, and integrating optimized models into policy-support systems. The FbOA–EALSTM framework offers a basis for further data-driven environmental forecasting.

REFERENCES

- [1] World Meteorological Organization, “State of the climate vehicle exhaust emissions for transport modes in lagos mate 2024 update for COP29,” World Meteorological Organization, Geneva, Switzerland, Tech. Rep., 2024.
- [2] United Nations Environment Programme, “Emissions gap report 2024: No more hot air ... please!” United Nations Environment Programme, Nairobi, Kenya, Tech. Rep., 2024.
- [3] D. Guo, F. Kou, B. He, and Y. Zhong, “Towards carbon neutrality: Machine learning analysis of vehicle emissions in Canada,” *Sustainability*, vol. 16, no. 23, p. 10526, 2024.
- [4] K. Lu, Y. Chen, Y. Tong, J. Zhang, Y. Luo, and J. Wang, “Eco-driving control for CAVs at signalized intersections: Adapting to traffic uncertainties,” *Transportation Research Part D: Transport and Environment*, vol. 132, p. 104270, 2024.
- [5] I. S. Bangroo, “AI-based predictive analytic approaches for safeguarding the future of electric/hybrid vehicles,” arXiv preprint, 2023.
- [6] Q. Wang, “Towards zero-emission urban mobility: Leveraging ai and lca for targeted interventions,” *Build-ing Simulation*, vol. 17, pp. 1653–1657, 2024.
- [7] S. Hochreiter and J. Schmidhuber, “Long short-term memory,” *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [8] S. Li, Z. Tong, and M. Haroon, “Estimation of transport CO₂ emissions using machine learning algorithm,” *Transportation Research Part D: Transport and Environment*, vol. 133, p. 104276, 2024.
- [9] N. Khodadadi, S. K. Towfek, A. M. Zaki, A. H. Alharbi, E. Khodadadi, D. S. Khafaga, L. Abualigah, A. Ibrahim, A. A. Abdelhamid, and M. M. Eid, “Predicting normal-ized difference vegetation index using a deep attention network with bidirectional gru: A hybrid parametric optimization approach,” *International Journal of Data Science and Analytics*, 2024.
- [10] M. Böhm, M. Nanni, and L. Pappalardo, “Gross pollutants and vehicle emissions reduction,” *Nature Sustainability*, vol. 5, pp. 699–707, 2022.
- [11] S. A. Ajayi, C. A. Adams, G. Dumedah, and A. Adebanji, “The impact of vehicle engine characteristics on city,” *Urban Planning and Transport Research*, vol. 12, 2024.
- [12] F. E. Sapnken, K. R. Hong, H. C. Noume, and J. G. Tamba, “A grey prediction model optimized by metaheuristic algorithms and its application in forecasting carbon emissions from road fuel combustion,” *Energy*, vol. 302, p. 131922, 2024.
- [13] 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*, vol. 11, no. 1, p. 2317540, 2024.
- [14] M. Zadmiraee, 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*, vol. 28, no. 1, pp. 565–591, 2024.
- [15] 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*, vol. 286, p. 129499, 2024.
- [16] M. Singh and R. K. Dubey, “Deep learning model based co₂ emissions prediction using vehicle telematics sensors data,” *IEEE Transactions on Intelligent Vehicles*, vol. 8, no. 1, pp. 768–777, 2023.
- [17] 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*, vol. 28, no. 39, pp. 55 535–55 553, 2021.
- [18] 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*, vol. 338, p. 120830, 2023.

- [19] H. Moayedi, A. Mukhtar, N. B. Khedher, I. Elbadawi, M. B. Amara, Q. T. Thi, and N. Khalilpoor, "Forecasting of energy-related carbon dioxide emission using ann combined with hybrid metaheuristic optimization algorithms," *Engineering Applications of Computational Fluid Mechanics*, vol. 18, no. 1, p. 2322509, 2024.
- [20] 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*, vol. 360, p. 131946, 2022.
- [21] K. Rokhmawati, V. Sarasi, and L. T. Berampu, "Scenario analysis of Indonesia carbon tax impact on carbon emissions using system dynamics modeling and STIRPAT model," *Geography and Sustainability*, vol. 5, pp. 577–587, 2024.
- [22] 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*, vol. 18, no. 1, p. 2391988, 2024.
- [23] W. Yu, L. Xia, and Q. Cao, "A machine learning algorithm to explore the drivers of carbon emissions in Chinese cities," *Scientific Reports*, vol. 14, p. 23609, 2024.
- [24] M.-Y. Cheng and R. R. Khasani, "An innovative moment balanced inference engine for predicting recycled concrete aggregate strength and minimizing mixture CO₂ emissions," *Construction and Building Materials*, vol. 441, p. 137482, 2024.
- [25] M. Salehi Sarbijan and J. Behnamian, "Multi-product production routing problem by consideration of outsourcing and carbon emissions: Particle swarm optimization," *Engineering Optimization*, vol. 53, no. 8, pp. 1298–1314, 2021.
- [26] S. Ding, J. Ye, and Z. Cai, "Multi-step carbon emissions forecasting using an interpretable framework of new data preprocessing techniques and improved grey multi-variable convolution model," *Technological Forecasting and Social Change*, vol. 208, p. 123720, 2024.
- [27] J. Udoh, J. Lu, and Q. Xu, "Application of machine learning to predict CO₂ emissions in light-duty vehicles," *Sensors*, vol. 24, no. 24, p. 8219, 2024.
- [28] 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*, vol. 18, no. 1, p. 2327437, 2024.
- [29] A. Arun Kumar, S. Suresh, A. Ramkumar, and A. Bhuvanesh, "A novel generation expansion planning in older power plants – hybrid spotted hyena–particle swarm optimization," *Opto-Electronics Review*, vol. 32, no. 3, p. e150612, 2024.
- [30] Y. Su and C.-C. J. Kuo, "On extended long short-term memory and dependent bidirectional recurrent neural network," *Neurocomputing*, vol. 356, pp. 151–161, 2019.
- [31] K. Cho, B. van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, "Learning phrase representations using rnn encoder-decoder for statistical machine translation," arXiv preprint, arXiv:1406.1078, 2014.
- [32] M. Schuster and K. K. Paliwal, "Bidirectional recurrent neural networks," *IEEE Transactions on Signal Processing*, vol. 45, no. 11, pp. 2673–2681, 1997.
- [33] S.-I. Amari, "Learning patterns and pattern sequences by self-organizing nets of threshold elements," *IEEE Transactions on Computers*, vol. C-21, no. 11, pp. 1197–1206, 1972.
- [34] F. Rosenblatt, "The perceptron: A probabilistic model for information storage and organization in the brain," *Psychological Review*, vol. 65, no. 6, pp. 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*, pp. 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*, vol. 11, pp. 341–359, 1997.
- [37] C. Reeves, "Genetic algorithms," Chapter in Handbook of Metaheuristics, volume 146, Springer, pp. 109–139, 2010.
- [38] O. Bingöl, U. Güvenc, S. Duman, and S. Paçacı, "Stochastic fractal search with chaos," In Proceedings of the 2017 International Artificial Intelligence and Data Processing Symposium (IDAP), pp. 1–6, 2017.
- [39] R. V. Rao, "Jaya: A simple and new optimization algorithm for solving constrained and unconstrained optimization problems," *International Journal of Industrial Engineering Computations*, vol. 7, pp. 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*, vol. 6, no. 6, pp. 580–593, 2002.
- [41] X.-S. Yang, "A new metaheuristic bat-inspired algorithm," Chapter in Nature Inspired Cooperative Strategies for Optimization (NISCO 2010), edited by J. R. González, D. A. Pelta, C. Cruz, G. Terrazas, and N. Krasnogor, Springer, pp. 65–74, 2010.

- [42] D. Simon, “Biogeography-based optimization,” *IEEE Transactions on Evolutionary Computation*, vol. 12, no. 6, pp. 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*, vol. 27, 2015.