CO2 Emission Predication Based on an Improved Swarm Optimization Technique

CO_2 Emission Predication Based on an Improved Swarm Optimization Technique

Aziza I. Hussein
College of Engineering, Energy and Technology Research Center., Effat University
Jeddah, Saudi Arabia
azibrahim@effatuniversity.edu.sa

Eslam Hussein Sayed
- Information System Department, Faculty of Computer Science, Nahda University,
Beni Suef, Egypt.
eslaamhusseinn@gmail.com

Rabab Hamed M. Aly
- The Higher Institute for Management and Information Technology,
Minya, Egypt.
rabab.cse2010@gmail.com

2025 1st Future International Conference on Artificial Intelligence and Cybersecurity (FICAC) | 979-8-3315-1383-2/25/$31.00 ©2025 IEEE | DOI: 10.1109/FICAC65757.2025.11341823

***Abstract*— Forecasting CO2 emissions is one of the most important measures governments must take to maintain human health in the current climate. CO_2 levels are a critical component of overall air quality and are closely linked to industrial activity. Industrial air pollution, particularly CO_2 emissions, remains one of the most significant challenges faced today. The use of optimization algorithms can support accurate predictions of CO2 emissions. This paper proposes an optimization strategy based on the proposed Bee- Particle Swarm optimization (BPSO) method to estimate CO_2 emissions. The proposed method integrates Bee Optimization and enhances it using particle swarm technique to achieve higher predictive accuracy for CO_2 emission forecasting. The model presented in this paper consists of two stages: the first involves data preprocessing using the Discrete Wavelet Transform (DWT), and the second involves feature selection and CO_2 emission prediction using the proposed BPSO algorithm. The results show that an increased number of features correlates with a heightened potential for improved accuracy. The outcomes demonstrate the effectiveness of the proposed methodology.**

Keywords*— *Carbon emission, Particle Swarm Optimization, Bee colony optimization, Discrete Wavelet Transform (DWT), Climate.

I. INTRODUCTION

The construction industry is a major contributor to world energy consumption and greenhouse gas emissions, attracting considerable scrutiny for its substantial carbon output. Carbon emission prediction is the process of predicting the future levels of carbon dioxide (CO_2) that will be emitted into the atmosphere because of a variety of human and natural activities. It is instrumental in the mitigation of climate change and the management of the environment by allowing industries to predict emission trends and devise effective strategies for reduction. This prediction typically depends on historical data, environmental variables, and sophisticated computational methods, including machine learning, optimization algorithms, and time-series forecasting. Governments can establish informed climate in industries can strive for sustainability, and scientists can gain a more comprehensive understanding of the potential long-term effects on global temperatures and ecosystems by accurately forecasting CO_2 emissions.

Artificial intelligence (AI) technologies are essential for enhancing low-carbon energy plans through precise and swift forecast, analysis, and optimisation of energy production and consumption. Simultaneously, big data technologies facilitate the collecting, integration, and analysis of extensive energy-related datasets, providing profound insights into market dynamics and industry trends. AI and big data collectively establish a robust framework for informed decision-making and the advancement of sustainable, low-carbon energy systems [1-4]. Furthermore, the Hybrid Carbon-Hydrogen Metallurgy Manufacturing Process (HCHMP) signifies a viable approach for the sustainable evolution of the iron and steel sector. In contrast to the conventional Pure Carbon Metallurgy Process (PCMP), which depends exclusively on carbon-based inputs and has been thoroughly examined for material, energy, and emission optimisations, HCHMP incorporates hydrogen as a reducing agent to diminish carbon emissions. Nevertheless, despite its promise, a significant deficiency exists in research regarding low-carbon optimisation models and methodologies specifically designed for HCHMP. Creating such models is crucial for enhancing efficiency, reducing environmental impact, and steering the sector towards more sustainable production methods [5,6]. In contrast, most metaheuristic solutions mimic physical, biological, or natural principles that exhibit intrinsic predictability and optimisation traits, rendering them exceptionally successful for CO_2 emission forecasting. These algorithms can traverse intricate, multidimensional data environments to discern patterns and correlations between industrial processes and emission outputs. Metaheuristics can improve the precision and efficacy of CO_2 emissions forecasting across diverse sectors, especially in energy-intensive businesses, by utilising their adaptive search capabilities.

The prediction of CO2 emissions in this paper is based on Bee Optimization and is enhanced with the Particle Swarm Optimization (PSO) technique to achieve greater predictive accuracy for CO2 emission forecasting.

There are two stages in this model: the first involves preprocessing data using the Discrete Wavelet Transform (DWT), and the second involves feature selection and prediction of CO2 emissions using the proposed Bee-Particle Swarm optimization (BPSO). Furthermore, BPSO algorithm performed better than other algorithms in terms of prediction effectiveness and accuracy rate. Section II provides an overview of related works in this paper. The proposed model is illustrated in Section III. The results of the research are presented in Section IV. The conclusion is provided in Section V.

II. RELATED WORKS

Presently various researchers have introduced a variety of optimisation techniques applicable to carbon emission analysis and prediction. This section focuses on previous research industries aimed at improving the accuracy and efficiency of CO_2 emission forecasting models, particularly using advanced optimisation strategies (see TABLE I.).

Study	Year	Technology	Strengths	Limitations
[7]	2025	To identify key drivers of urban expansion and predict future growth patterns, the authors used advanced Geographic Information Systems (GIS) tools and machine learning techniques. Using this approach, they were able to shed light on how urbanization interacts with cultural and historical landscapes.	Focus on Heritage Preservation: By concentrating on a heritage-rich area like Gharb Sohail, the study emphasizes the importance of balancing development with cultural preservation.	Data Limitations: The accuracy of machine learning models heavily depends on the quality.
[8]	2024	Least Squares Support Vector Machine (LSSVM) optimised using Whale Optimisation Algorithm (WOA) for forecasting emissions in power systems incorporating renewable energy sources.	High prediction accuracy with Mean Square Error (MSE); effectively handles non-linear relationships.	Some difficulties in adapt to developing conditions in power systems.
[9]	2024	The authors propose a deep learning-based approach combining the Long Short-Term Memory (LSTM) model and scenario simulation techniques to forecast carbon emissions in China's logistics industry.	- High forecasting accuracy using LSTM models. - Provides peak carbon emission pathways, aiding long-term strategic planning	- Data limitations may affect model robustness in real-world applications
[10]	2024	The CNN-BRNN model is proposed as a hybrid deep learning model.	The hybrid architecture effectively captures both spatial and temporal dependencies.	Does not include CO2 emissions directly as an output variable.
[11]	2023	Enhances prediction accuracy by combining Oscillatory Particle Swarm Optimization (OPSO) with Long Short-Term Memory (LSTM) networks.	Significant improvement over standard LSTM models; effectively mitigates overfitting.	The integration of OPSO increases the computational complexity.
[12]	2023	The paper compared SVM, Random Forest, and k-NN models for predicting PM2.5 and PM10 concentrations. The paper showed the impact of training data characteristics and evaluates model performance on a variety of international datasets.	The paper provided practical insights into how different machine learning algorithms, including data preprocessing and hybrid models, handle air pollution prediction. Additionally, the impact of the dataset on the accuracy of the prediction is highlighted	Focused on particulate matter (PM) rather than directly on CO_2 emissions
[13]	2023	The paper proposed a hybrid model combining Ensemble Empirical Mode Decomposition (EEMD) and Particle Swarm optimisation-based backpropagation Neural Networks (PSOBP).	This is the first time EEMD has been applied to the prediction of carbon emissions, effectively handling non-stationary, nonlinear patterns of data.	The performance of EEMD is heavily dependent on accurate decomposition; errors may cascade into predictions if EEMD is not accurate.
[14]	2022	Several machine learning (ML) approaches were used to investigate transportation-related CO2 emissions. To analyze and predict emission patterns in urban areas, the study combined spatial and environmental data. To identify dominant influencing factors and spatial distribution patterns, ML algorithms such as Random Forest (RF), Gradient Boosting, and others are used.	Utilizes high-performance models such as RF and GBM to capture non-linear relationships and feature importance.	The results of tree-based ML models such as RF/GBM are often difficult to understand by non-technical stakeholders due to their lack of interpretability.
[15]	2021	In the proposed system, real-time in-vehicle sensor data is used to estimate the CO2 emissions of the vehicle using a Recurrent Neural Network (RNN) -based Long Short-Term Memory (LSTM) model.	Based on public OBD-II data, the proposed model provides a scalable and efficient method of monitoring automobile emissions at the vehicle level and has been evaluated	Collecting vehicle-level emission data could raise data privacy or cybersecurity challenges in real-world applications.
[16]	2021	They used deep learning models (LSTM and GRU) to forecast carbon emissions.	According to the authors, deep learning can capture complex nonlinear trends in CO_2 emissions data.	The Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) models were implemented.

TABLE I. LIST OF RELATD WORKS

Based on previous studies, we propose an innovative optimization-constructed approach. By using real-time recorded data, this method uses the Bee- Particle Swarm optimization (BPSO) to select features, followed by predicting CO2 emissions. This approach will be discussed in more detail in the following section.

III. Methodology

The main objective behind this work is to use a proposed technique to select features of CO_2 and classify with predicting CO_2 emission. Fig.1 shows the main blocks of this work.

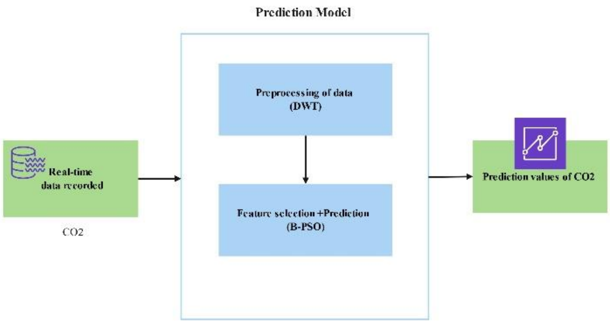


Fig. 1. Proposed system.

A. Real time data recorded

The first step in the system model involves data collection. The data collection of this system is database CSV files from an online dataset for providing an interactive data Explorer for analyzing historical greenhouse gas (GHG) emissions across global regions [17]. All greenhouse gases in these databases including CO_2, CH_4, N_2O, and F-gases. Based on historical emissions data obtained from the Climate Watch platform, this paper uses Egypt as the target country for CO2 emission prediction. Real-world, time-series records of greenhouse gas emissions are included in this dataset [17], which serves as the basis for training and evaluating the proposed prediction model.

B. Preprocessing stage

Preprocessing procedure is applied by using Discrete Wavelet Transform (DWT) method based on the criteria of the technique on python [18]. This part focuses on the extraction of features from datasets based on the DWT. The DWT is an effective method for decomposing nonstationary data into different scales for the purpose of analyzing it. The process involves breaking down the CO2 CSV file into a set of multi-resolution coefficients to reveal various facets of the data, such as trends, discontinuities, and recurring patterns.

Based on a scaling function and a wavelet function, DWT extracts the characteristics of CO2 data. During this decomposition process, two distinct components are produced: the approximation component and the detail component. Both components contain important information about the CO2 data [19].

C. Feature selection and prediction based on BPSO

The selection of features plays an important role in determining the accuracy of prediction models. In the case of a dataset with many features, the resulting high-dimensional space may become cluttered and noisy, negatively affecting the accuracy of the prediction. The PSO method was found to perform higher than average accuracy values when applied to feature selection techniques [20,21]. Using it to select the features and improving it with Bee Colony Optimization (BCO) to help in the prediction values [22], we applied it in this paper. According to PSO, the main equation for feature selection is as follows:
Every particle in PSO flies in the search space at a velocity established by its own memory of flying and that of its companion. . Based on the assumption that a minimization problem is being addressed, the following equations are given:

$$\boldsymbol{p^t_{best_i}} = x_i^* | f(x_i^*) = min_{k=1,2,..,t}(\{f(x_i^k)\}), \quad (1)$$

$$where\ i \in \{1,2, \dots. N\}$$

$$\boldsymbol{g^t_{best_i}} = x_*^t | f(x_*^t) = min(\{f(x_i^k)\}), \quad i = 1,2, \dots. N, \quad k = 1,2, \dots.., t \quad (2)$$

Where i denotes the particle's index, t denotes the current iteration number, f denotes the objective function to be optimized (minimized), x is the position vector (or a potential solution), and N denotes the number of particles in the swarm [23].
According to the following equations, the velocity v and position x of each particle i are updated at the current iteration t + 1 as follows:

$$v_i^{t+1} = \omega \boldsymbol{v_i^t} + c_1 \boldsymbol{r_1}(\boldsymbol{p^t_{best_i}} - \boldsymbol{x_i^t}) + c_2 \boldsymbol{r_2}(\boldsymbol{g^t_{best_i}} - \boldsymbol{x_i^t}) \quad (3)$$

$$\boldsymbol{x_i^{t+1}} = \boldsymbol{x_i^t} + v_i^{t+1}, \quad (4)$$

The velocity vector v represents the velocity vector, the inertia weight ω represents the inertia weight employed to balance the local exploitation r1 and the global exploitation r2, and the random vectors are uniformly distributed within the interval [0, 1]. Additionally, we developed an enhanced velocity update equation (3) based on BCO [24,25]to improve selected PSO features by (5)

$$v_i^{t+1} = \omega v_i^t + c_1 r_1 (p^t_{best} - x_i^t) + c_2 r_2 (g^t_{best} - x_i^t) + \phi_1 . \psi_{ij}(t) \quad (5)$$

The equation is updated by adding two values:
- ϕ_1 is the Bee-inspired assuming local search coefficient based on (6).
- $\psi_{ij}(t)$Bee-guided perturbation based on (7)

$$\phi_1 = 0.1. (1- {}^{t}/_{t_{max}}) \quad (6)$$

$$\psi_{ij}(t) = \frac{f(x_{ij})}{\sum f} \cdot \eta_{ij,where\ \eta_{ij \sim N(0,1)}} \quad (7)$$

D. *The algorithm of BPSO*

General algorithm of BPSO to select the features and Predict the CO_2 emissions

Start

Step 1:

- The data collection for this system consists of CSV files sourced from an online dataset that provides an interactive data explorer for analyzing historical greenhouse gas (GHG) emissions. In this context, the focus is on Egypt's emissions data, enabling a detailed analysis of the country's GHG trends over time within the global framework.
- Preprocessing for the datasets based on DWT.

Step 2:

- Applying the PSO using the subsequent procedures by using equations (1) to (4).
- The initialization of PSO within the interval values [0, 1].

Step 3:

- The solution (position) of velocity is updated by BCO equation(5)

Step 4:

- Update ϕ_1, $\psi_{ij}(t)$ based on equation (6), and (7) for BCO.

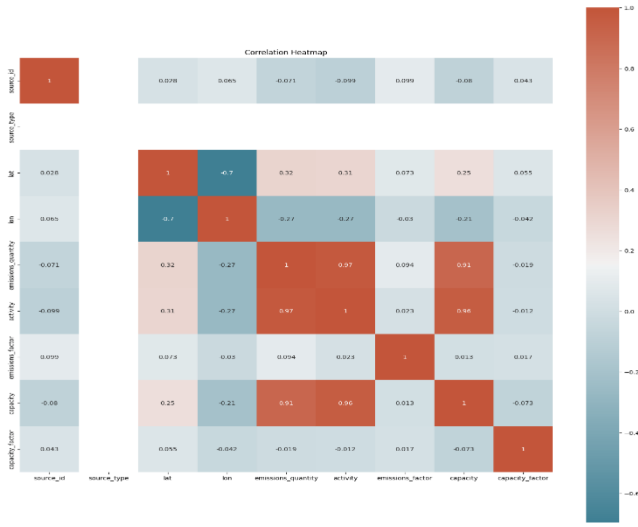
Step 5:

- Evaluation of the features and predict the CO_2 emissions based on the optimal extracted features

End

IV. Results and Discussion

This paper focuses on Co2 emission prediction and how to choose optimal features for detecting optimal results. The model was simulated using Python 3.7 tools on a PC with an i7-11800H, GeForce RTX 4GB, 16GB RAM, and 1TB SSD. Furthermore, the proposed model for predicting and detecting features of CO_2 is based on real-time recording, as well as recorded data.
The data were gathered from historical climate data for Egypt between June 2021 and January 2025[17]. Using the Python library, the data of all transportation areas is visualised using the data analysis function and the data is visualized using a heatmap of the correlation between data points (see Fig.2)
By applying the BPSO, the model Performance illustrates a powerful case for feature selection, establishing that reducing the number of features simplifies the model and improves its predictive ability (see Fig.3).



a) Accuracy with all 8 features: The model achieves 97.63% accuracy when using the complete set of features.
b) Accuracy with 2 selected features: After feature selection, the model's performance improves to 98.16% accuracy despite using only 2 features.

Fig. 2. The correlation heatmap of the dataset

c) Feature reduction: The algorithm successfully reduced the feature set from 8 to 2, eliminating 75% of the original features while improving accuracy by 0.53 percentage points.

d) The model is more scalable due to a 75% reduction in features, which reduces computational costs and speeds up inference.
e) The higher accuracy suggests that the selected features.

Table II shows the comparing between the BPSO with the other algorithms from pervious works and Fig.4.

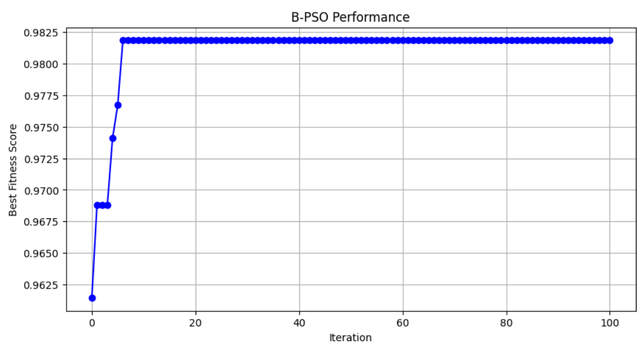


Fig. 3. The performance of the BPSO model.

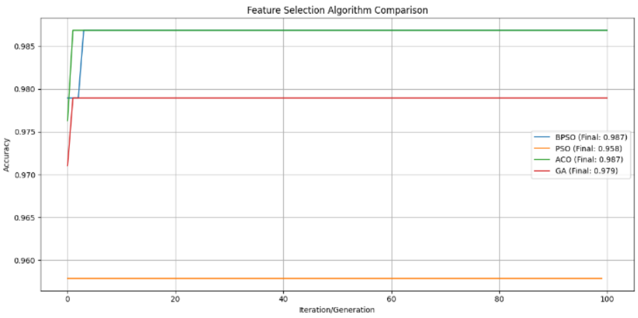


Fig. 4. The feature selection of the BPSO model vs the pervious works

TABLE II. COMPARE The BPSO MODEL WITH PERVIOUS WORK

Method	Iteration	Accuracy	Feature Reduction
PSO[23]	**100**	**95.7%**	**75%**
ACO[22]	**100**	**98%**	**75%**
GA[24]	**100**	**97.8%**	**12.5%**
B-PSO of this paper	**100**	**98.8%**	**75%**

As compared to previous work, PSO provides higher accuracy in prediction after being improved by BCO, as well as more practical values and optimal features.

V. CONCLUSION

Climate change refers to the long-term change in temperatures and weather patterns. Such shifts may be natural, because of changes in the sun's activity or large volcanic eruptions. The burning of fossil fuels like coal, oil, and gas has been the primary driver of climate change since the 1800s. The use of optimization algorithms can support accurate prediction of CO2 emissions. The purpose of this paper is to propose an optimization strategy based on the Bee-Particle Swarm Optimization (BPSO) method to estimate the emissions of carbon dioxide. By incorporating Bee Optimization and particle swarm technique, the proposed method achieves higher predictive accuracy for CO2 emission forecasting. It consists of two stages: the first stage involves data preprocessing using the Discrete Wavelet Transform (DWT), and the second stage involves feature selection and CO2 emission prediction using the proposed BPSO algorithm. Results indicate that an increase in features correlates with greater potential for improving accuracy. This methodology has been shown to be effective based on the outcomes.

REFERENCES

[1] A. A. Ajala et al., “An examination of daily CO_2 emissions prediction through a comparative analysis of machine learning, deep learning, and statistical models,” *Environ. Sci. Pollut. Res.*, pp. 1–26, 2025.

[2] J. Zhao et al., “Carbon emission prediction model and analysis in the Yellow River basin based on a machine learning method,” *Sustainability*, vol. 14, no. 10, p. 6153, 2022.

[3] S. Khurana et al., “Predictive modeling of engine emissions using machine learning: A review,” *Mater. Today: Proc.*, vol. 38, pp. 280–284, 2021.

[4] V. G. Nguyen et al., “An extensive investigation on leveraging machine learning techniques for high-precision predictive modeling of CO_2 emission,” *Energy Sources, Part A: Recovery, Util., Environ. Eff.*, vol. 45, no. 3, pp. 9149–9177, 2023.

[5] X. Li and X. Zhang, “A comparative study of statistical and machine learning models on carbon dioxide emissions prediction of China,” *Environ. Sci. Pollut. Res.*, vol. 30, no. 55, pp. 117485–117502, 2023.

[6] W. Feng et al., “Application of neural networks on carbon emission prediction: A systematic review and comparison,” *Energies*, vol. 17, no. 7, p. 1628, 2024.

[7] H. Hasanoon et al., “Exploring urban growth drivers in heritage areas using machine learning: The case of Gharb Sohail, Aswan,” *J. Eng.*, 2025.

[8] Y. Luo et al., “An adaptive prediction method for carbon emissions of power systems containing new energy based on least-squares support vector machine,” *Sustain. Energy Res.*, vol. 11, no. 1, p. 36, 2024.

[9] T. Chen and M. Wang, “Deep learning-based carbon emission forecasting and peak carbon pathways in China’s logistics industry,” *Sustainability*, vol. 16, no. 5, p. 1826, 2024.

[10] F. Aljuaydi et al., “A deep learning prediction model to predict sustainable development in Saudi Arabia,” *Appl. Math. Inf. Sci.*, vol. 18, no. 6, pp. 1345–1366, 2024.

[11] Y. Chen et al., “Research of carbon emission prediction: An oscillatory particle swarm optimization for long short-term memory,” *Processes*, vol. 11, no. 10, p. 3011, 2023.

[12] A. Pandya et al., “A comparative and systematic study of machine learning (ML) approaches for particulate matter (PM) prediction,” *Arch. Comput. Methods Eng.*, vol. 31, no. 2, pp. 595–614, 2024.

[13] W. Sun and C. Ren, “Short-term prediction of carbon emissions based on the EEMD-PSOBP model,” *Environ. Sci.*, vol. 28, no. 40, pp. 56580–56594, 2021.

[14] X. Li, A. Ren, and Q. Li, “Exploring patterns of transportation-related CO_2 emissions using machine learning methods,” *Sustainability*, vol. 14, no. 8, p. 4588, 2022.

[15] J. Zhao et al., “Carbon emission prediction model and analysis in the Yellow River basin based on a machine learning method,” *Sustainability*, vol. 14, no. 10, p. 6153, 2022.

[16] H. Huang, X. Wu, and X. Cheng, “The prediction of carbon emission information in Yangtze River economic zone by deep learning,” *Land*, vol. 10, no. 12, p. 1380, 2021.

[17] https://www.climatewatchdata.org/data-explorer/.

[18] K. M. Aljebory, M. J. Yashar, and T. S. Mohammed, “Classification of EMG signals: Using DWT features and ANN classifier,” *IAENG Int. J. Comput. Sci.*, vol. 51, no. 1, 2024.

[19] Q. Sun et al., “Integrated prediction of carbon price in China based on heterogeneous structural information and wall-value constraints,” *Energy*, vol. 306, p. 132483, 2024.

[20] Z. Wang et al., “Intelligent construction of a supply chain finance decision support system and financial benefit analysis based on deep reinforcement learning and particle swarm optimization algorithm,” *Int. J. Manag.*, vol. 8, no. 3, pp. 28–41, 2025.

[21] A. K. Mahapatra et al., “A fast-flying particle swarm optimization for resolving constrained optimization and feature selection problems,” *Cluster Comput.*, vol. 28, no. 2, p. 91, 2025.

[22] R. H. M. Aly, K. H. Rahouma, and H. F. A. Hamed, “Brain tumors diagnosis and prediction based on applying the learning metaheuristic optimization techniques of particle swarm, ant colony and bee colony,” *Procedia Comput. Sci.*, vol. 163, pp. 165–179, 2019.

[23] A. G. Gad, “Particle swarm optimization algorithm and its applications: A systematic review,” *Arch. Comput. Methods Eng.*, vol. 29, no. 5, pp. 2531–2561, 2022.

[24] M. Jawad et al., “Energy optimization and plant comfort management in smart greenhouses using the artificial bee colony algorithm,” *J. Clean. Prod.*, vol. 15, no. 1, p. 1752, 2025.

[25] V. Alevizos et al., “Optimizing carbon footprint in ICT through swarm intelligence with algorithmic complexity,” *arXiv*, preprint arXiv:2501.17166, 2025.