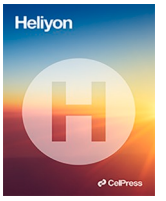
Heliyon 10 (2024) e31850

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon



Research article

Enhanced multi-layer perceptron for CO2 emission prediction with worst moth disrupted moth fly optimization (WMFO)

Oluwatayomi Rereloluwa Adegboye [a], Ezgi Deniz Ülker [b], Afi Kekeli Feda [c], Ephraim Bonah Agyekum [d], Wulfran Fendzi Mbasso [e,*], Salah Kamel [f]

[a] *Engineering Managment, University of Mediterranean Karpasia, Mersin-10, Turkey*
[b] *Computer Engineering, European University of Lefke, Mersin-10, Turkey*
[c] *Advanced Research Centre, European University of Lefke, Northern Cyprus, TR-10, Mersin, Turkey*
[d] *Department of Nuclear and Renewable Energy, Ural Federal University named after the first President of Russia Boris Yeltsin, 620002, 19 Mira Street, Ekaterinburg, Russia*
[e] *Technology and Applied Sciences Laboratory, UIT of Douala, P.O. Box 8689, Douala, University of Douala, Cameroon*
[f] *Electrical Engineering Department, Faculty of Engineering, Aswan University, 81542, Aswan, Egypt*

ARTICLE INFO

Keywords:
Moth fly optimization
Multi-layer perceptron
Machine learning
Optimization carbon emission

ABSTRACT

This study introduces the Worst Moth Disruption Strategy (WMFO) to enhance the Moth Fly Optimization (MFO) algorithm, specifically addressing challenges related to population stagnation and low diversity. The WMFO aims to prevent local trapping of moths, fostering improved global search capabilities. Demonstrating a remarkable efficiency of 66.6 %, WMFO outperforms the MFO on CEC15 benchmark test functions. The Friedman and Wilcoxon tests further confirm WMFO's superiority over state-of-the-art algorithms. Introducing a hybrid model, WMFO-MLP, combining WMFO with a Multi-Layer Perceptron (MLP), facilitates effective parameter tuning for carbon emission prediction, achieving an outstanding total accuracy of 97.8 %. Comparative analysis indicates that the MLP-WMFO model surpasses alternative techniques in precision, reliability, and efficiency. Feature importance analysis reveals that variables such as Oil Efficiency and Economic Growth significantly impact MLP-WMFO's predictive power, contributing up to 40 %. Additionally, Gas Efficiency, Renewable Energy, Financial Risk, and Political Risk explain 26.5 %, 13.6 %, 8 %, and 6.5 %, respectively. Finally, WMFO-MLP performance offers advancements in optimization and predictive modeling with practical applications in carbon emission prediction.

1. Introduction

A significant concern for many developing nations today is environmental degradation due to greenhouse gas (GHG) emissions [1]. The surge in industrial activities, driven by increased industrialization and urbanization, has led to higher GHG emissions, contributing to a notable increase in overall atmospheric concentrations [2]. This, in turn, has given rise to global warming and climate change, adversely impacting global agricultural productivity through factors like diminished rainfall, seasonal fluctuations, and temperature increases [3]. Consequently, the paramount focus for many countries is the mitigation of GHC emissions. The Paris Agreement,

* Corresponding author.
E-mail addresses: aoluwatayomi@gmail.com (O.R. Adegboye), ezgiulker@ecrowdsolutions.com (E.D. Ülker), kekelifeda@gmail.com (A.K. Feda), agyekumephraim@yahoo.com (E.B. Agyekum), fendzi.wulfran@yahoo.fr (W. Fendzi Mbasso), skamel@aswu.edu.eg (S. Kamel).

https://doi.org/10.1016/j.heliyon.2024.e31850
Received 26 January 2024; Received in revised form 1 May 2024; Accepted 22 May 2024
Available online 27 May 2024

established under the United Nations Framework Convention on Climate Change (UNFCCC), underscores the commitment to reducing GHG emissions. This agreement, signed in 2016 in Paris, France, signifies a collective effort to address the environmental challenges associated with escalating carbon emissions [4,5]. It is designed to restrain the rise in the global average temperature to below 2 °C, with a further aim of reducing it to 1.5 °C to mitigate the impacts of climate change [6]. Carbon dioxide stands out as a crucial greenhouse gas (GHG), primarily stemming from human activities, vehicular emissions, fossil fuel combustion (coal, oil, and natural gas), and various industrial sectors [7]. The production rate of CO2 from diverse human activities far surpasses its absorption rate, primarily attributed to diminishing forest cover [8,9]. Given the pivotal role of carbon dioxide in the surge of GHG emissions, there is a pressing need for policies that can effectively curb these CO2 emissions. Forecasting carbon dioxide production proves instrumental in identifying sectors requiring close monitoring and formulating new policies to mitigate GHG emissions.

The utilization of artificial intelligence techniques for environmental forecasts has experienced a notable surge in the last few years. These techniques are recognized for their stability, simplicity of use, and commendable effectiveness. One contributing factor to their increasing popularity is the success of black box methods in deciphering intricate patterns without requiring specific subject knowledge [10]. Among these techniques, categorized as mathematical techniques able to deduce patterns from data and apply them to other sets of data, the artificial neural network (ANN) stands out as the most widely utilized and renowned. ANNs possess a remarkable capability to adapt and model nonlinear relations between input and output variables thanks to their extensive interconnectivity [11]. Among ANN models, the multilayer perceptron (MLP) neural network stands out as particularly prevalent [9]. MLP is characterized by its composition of multiple layers, including an input layer, an output layer, and one or multiple hidden layers in between. Every layer, excluding the input layer, employs a non-linear activation function to establish a curvilinear correlation between the input and output. MLP facilitates learning by fine-tuning the weights of the neurons it self-assigns. The appeal of MLP lies in its capacity to tackle intricate non-linear challenges, excel with substantial input data, and deliver prompt predictions post-training. Moreover, it showcases the ability to maintain a consistent accuracy ratio across both limited and extensive datasets [12].

MLP was successfully applied and obtained noteworthy results in different domains, including cancer diagnosis [13], stochastic synthesis [14], wave forecasting [15], wind speed prediction [16], sediment load estimation [17], stock price index prediction [18], anomaly detection [19], landslide susceptibility mapping [20], speech emotions recognition, and intrusion detection [21]. Despite being regarded as a highly accurate approach, MLP comes with various challenges. The primary drawbacks include the need for fine-tuning configuration parameters, slow convergence speed, and a significant risk of encountering a local minimum point prematurely concluding the training phase without reaching the optimal solution [20,22]. In the realm of forecasting, some studies have utilized optimization techniques in an attempt to address the aforementioned limitations of MLP. Examples include the MLP-Intelligent Water Drop algorithm proposed by Pham et al., to improve the global optima search [23]. Two sets of historical river streamflow data were gathered to assess the MLP-IWD model's effectiveness. To evaluate the model input-output architecture's performance under various circumstances, four distinct scenarios were looked at. The suggested MLP-IWD model outperforms the traditional MLP-NN model, according to the outcomes, and greatly improves river streamflow forecasting precision. Dokur et al. suggested a novel approach to significant wave height prediction, which integrates swarm decomposition (SWD) and multi-strategy random weighted grey wolf optimizer (MsRwGWO) into a multi-layer perceptron (MLP) forecasting model [24]. The MsRwGWO effectively optimizes the MLP model's parameters, while the SWD technique creates more dependable, fixed, and regular variations of the original signal. As a result, the predicting precision is improved. To evaluate the effectiveness of the suggested model, real wave records from the North Atlantic Sea were utilized. A comparative analysis with cutting-edge prediction techniques based on deep learning was also carried out, and the results showed notable improvements in reliability. To attain the ideal MLP parameters, Ashraf et al. developed an improved MLP by incorporating a sophisticated Genetic Algorithm (GA), which mimics "seeded selection" in natural selection [25]. In addition to experimental results from the forecast of the productivity of Hemispherical Solar Stills (SSs), the proposed Enhanced Genetic Algorithm – Multi-Layer Perceptron (EGA-MLP) is contrasted with two different algorithms: the original MLP and Particle Swarm Optimization with MLP (PSO-MLP). Three distinct SS designs are tested for output under varying climatic conditions using the models. Among the models, EGA-MLP maintains the lowest RMSE and MAE values and the most outstanding R^2 values throughout every experiment. In comparison to the compared approaches and the experimental data, the suggested model offers the most reliable SS yield forecasts, suggesting potential advancements in SS system optimization and design. Ahmed aimed to develop an accurate model for predicting agricultural production, aiding farmers in effective planning [26]. Initially, the Crop Yield Prediction Dataset is normalized, followed by feature engineering to assess the significance of each feature in predicting crop yield. The prediction is conducted using the Multi-Layer Perceptron (MLP) model combined with the Spider Monkey Optimization (SMO) method. The MLP model effectively handles nonlinear relationships between features, while SMO optimizes feature weights. Data from the Food and Agriculture Organization and the World Data Bank are utilized to predict maize production in Saudi Arabia. The proposed MLP-SMO model's predictive performance was evaluated, and the results indicate superior prediction accuracy, outperforming other techniques considered in the statistical analysis. Other examples include the hybrid MLP Genetic Algorithm [27], the MLP based Grasshopper Optimization Algorithm [28], the hybrid MLP- Grey Wolf Optimization [29], the MLP-Particle Swarm Optimization [30], and the adaptive dynamic grey wolf-dipper throated optimization (ADGWDTO – MLP) [31].

The Moth–Flame Optimization (MFO) algorithm, a notable biomimetic metaheuristic, draws inspiration from moths' spiral motion around a source of light during the night [32]. Renowned for its ease of use and favorable time complexity, MFO has found utility in addressing diverse real-world issues, including engineering design problems [33], clustering [34], feature selection [35], constrained engineering optimization problems [36], permutation-based problems [37], scheduling [38], optimal power flow problem [39], and image processing [40]. Despite its applicability, an inherent limitation of MFO is observed, namely, a tendency towards inadequate exploration and reduction of population diversity, resulting in early convergence and local optima confinement before reaching global optimum. In response to these limitations, different variants of the MFO were recently proposed. Shehab et al. introduced

modifications to the Moth Flame Optimization (MFO) algorithm through a two-step process [41]. Firstly, they hybridized MFO with the local-based algorithm, hill climbing (HC), creating MFOHC. The effectiveness of the suggested models was contrasted against well-established meta-heuristic algorithms. Nadimi-Shahraki et al. suggested a migration-based moth–flame optimization (M-MFO) that places a primary emphasis on enhancing the positioning of less fortunate moths through stochastic migration in the initial iterations, facilitated by a random migration agent [42]. The novel M-MFO was subjected to assessment to validate M-MFO's effectiveness. Ma et al. also proposed enhancements to the MFO to address issues related to premature convergence and getting stuck in local optimum [43]. To tackle the diversity aspect, a diversity feedback control mechanism, in the form of an inertia weight, is incorporated into the Moth-Flame Optimization algorithm. Additionally, a small probability mutation is introduced after the location update phase. Comparative analyses with other improved algorithms from the literature demonstrate that the proposed method outperforms its counterparts. Some other research focused on merging MFO with ANN techniques for addressing different problems. Singh et al. introduced a novel hybrid optimization approach for multi-objective optimization, combining a back-propagation artificial neural network (ANN) and MFO [44]. Results obtained from the hybrid ANN–MFO algorithm offer effective and accurate process parameters, contributing to improved surface finish and part quality. Fathy et al. introduced an innovative approach utilizing the Moth-Flame Optimizer (MFO) to optimize the output power of a Solid Oxide Fuel Cell (SOFC) by finding optimal features for its model [45]. The SOFC model is constructed using an ANN trained on experimental datasets. Results demonstrate that the introduced ANN-MFO enhances SOFC power by 18.92 % and 5.56 % compared to ANN-GA and ANN-RMO. Bui et al., suggested a robust model for the prediction of Peak Particle Velocity (PPV) in open pit mines [46]. The primary model utilizes an Adaptive Fuzzy Inference Neural Network (ANFIS). To enhance its accuracy, MFO is incorporated into ANFIS, resulting in the MFO–ANFIS. The results demonstrate that the MFO–ANFIS model achieves the highest precision.

While the literature highlights the recognized benefits of MLP and its variants, it is important to note that MLP still exhibits certain drawbacks. Specifically, the performance of MLP is notably contingent on the initial weights and biases, which may lead to sub-optimal convergence and the risk of falling into local optima traps. To tackle these problems, our research introduces an innovative solution by training MLPs with the Worst Moth Disrupted Moth Fly Optimization, which incorporates a strategy called the Worst Moth Disruption Strategy (WMD). The research approach differs from the MFO-MLP introduced by Yamany et al. [47], where the traditional MFO was applied to optimize the weight and bias of the MLP. The introduction of the WMD strategy aims to bolster the diversity of the population and mitigate the risk of falling into strong exploitation traps. The WMD strategy employs a novel position update equation that incorporates the global worst moth, providing a means to escape the confines of the strong exploitation traps. Consequently, WMFO is paired with MLP as a hybridized approach and employed to predict the CO2 emission. This amalgamation facilitates the correct tuning and strategic choice of the most efficient parameters, specifically weight and bias for the network. The rationale behind selecting an MLP for carbon emission prediction within this research lies in its ability to capture non-linear relationships, flexibility in handling

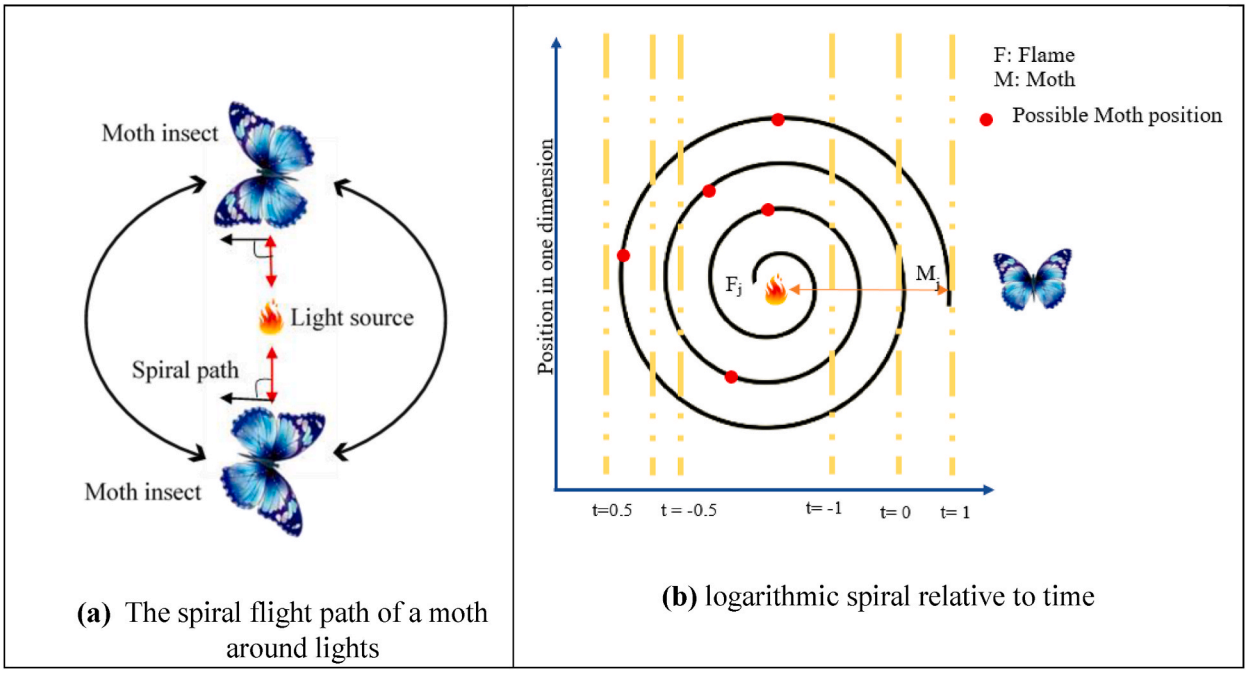


Fig. 1. The spiral flight path and logarithmic spiral movement of moths [32].

$$M_i = S(M_i \cdot F_j) \tag{1}$$

diverse predictors, feature learning capability, scalability, generalization ability, and synergistic integration with optimization algorithms. These characteristics make MLPs well-suited for modeling the complex and dynamic nature of carbon emission data, leading to more accurate predictions and insights for environmental sustainability efforts.

The remainder of the paper is thus organized. Section 2 briefly presents the Materials and Methodology, while section 3 focuses on the Proposed Method. In section 4 MLP- WMFO is examined through statistical evaluation, and the results are discussed. The conclusions drawn from these assessments are summarized in Section 5.

2. Materials and Methodology

2.1. Background of MFO

An innovative technique based on population optimization called MFO emulates the moth's nocturnal movement toward light sources [32]. Moths are known to keep a constant angle with the moon as they fly great distances using the "transverse orientation" technique. The distance separating moths and the source of light has a significant impact on the effectiveness of the technique. The theoretical framework for MFO primarily revolves around moths and their attraction to flame. The population functions as agents of exploration, and the flames indicate the best moth locations inside the search space. The MFO's primary operation S drives the moths to move throughout the problem space, as expressed in Eq (1). Each moth searches close to the correlated flame, as seen in Fig. 1a and adjusts its location to get a better result using the logarithmic spiral technique, as illustrated in Fig. 1b. A moth is never without its most effective solution in this scenario.Where F_j is the j^{th} flame and M_i is the i^{th} moth in Eq (1). The matrix of values in Eq. (2) represents the group of moths used in the algorithm.

$$M = \begin{bmatrix} m_{1,1} & m_{1,2} & \cdots & m_{1,d} \\ m_{2,1} & \ddots & \ddots & m_{2,d} \\ \vdots & \ddots & \ddots & \vdots \\ m_{n,1} & m_{n,2} & \cdots & m_{n,d} \end{bmatrix} \tag{2}$$

where n is the total number of moths present in the d problem space. As described in Eq (3), the fitness of every moth is calculated and stored in an array A.

$$A = Z(M) = [A_1, A_2 \ldots A_n]^T \tag{3}$$

Eq (4) presents the flames' matrix similar to the moth's matrix M. Similarly, Eq (5) provides the associated fitness value stored in an array E.

$$F = \begin{bmatrix} f_{1,1} & f_{1,2} & \cdots & f_{1,d} \\ f_{2,1} & \ddots & \ddots & f_{2,d} \\ \vdots & \ddots & \ddots & \vdots \\ f_{n,1} & f_{n,2} & \cdots & f_{n,d} \end{bmatrix} \tag{4}$$

$$E = Z(F) = [E_1, E_2 \ldots E_n]^T \tag{5}$$

Here Z represents the optimization problem's fitness function. The spiral equation, expressed in Eq. (1), which determines the path of every moth is detailed in Eq (6).

$$S(M_i \cdot F_j) = D_i \cdot e^{bt} \cdot \cos(2\pi t) + F_j \tag{6}$$

$$D_i = |F_j - M_i| \tag{7}$$

In Equations (6) and (7) F_j represents the j^{th} flame, D_i represents the distance separating a moth from a flame and M_i represents the i^{th} moth. The form of the "logarithmic spiral function" is determined by b. The moth moves toward the flame in steps, which are determined by the t parameter. To put it another way, t indicates the distance the moth to the flame at its next position. The moth is farthest from the flame when $t = 1$, while $t = -1$ denotes the moth's nearest location to the flame. The integer t is obtained as depicted in Eq (8) and Eq (9).

$$t = (a - 1) \times \text{rand} + 1 \tag{8}$$

$$a = -1 + \text{iter} \times \left(\frac{-1}{\text{Maxiter}}\right) \tag{9}$$

$$B = \text{round}\left(N - \text{iter} * \frac{N - 1}{\text{Maxiter}}\right) \tag{10}$$

In addition, as the search progresses, the quantity of flames gradually drops to equalize diversification and intensification. Because of this, moths only move during iterations according to the most suitable alternative. A proposed adaptive process for determining the

amount of flames is depicted in Eq (10): where N is the total number of flames, *iter* is the current iteration number, and *Maxiter* is the maximum iteration. The Flow chart of MFO is shown in Fig. 2

2.2. MLP training and hyperparameter optimization

The MLP comprises numerous layers, beginning with the input layer and concluding with the output layer, with intermediate layers termed as hidden layers [48]. The typical structure of an MLP is depicted in Fig. 3, which consists of just one hidden layer. Neurons in different levels are connected, and each connection has a weight.

Each node within the hidden layer is capable of performing two fundamental processes: activation and summation [48]. The summing operation makes use of the multiplication of bias, input, and weights, as given in Eq (11).

$$S_j = \sum_{i=1}^{n} \omega_{ij} I_i + \beta_j \tag{11}$$

Here, n signifies the number of inputs, I_i represents the i-th input value, ω_{ij} stands for the connection weight, and finally, β_j denotes the bias. The activation process is applied to the outcome of Equation (11). Numerous activation functions are available, with a commonly employed approach being the utilization of the Rectified Linear Unit Activation Function (ReLU), as defined in Eq (12):

$$max(0, S_j) \tag{12}$$

The evaluation of the Artificial Neural Network's performance is established through the employment of a loss function. A prevalent method involves using the mean squared error (MSE) as the designated loss function. The MSE quantifies the cumulative sum of squared variances between the observed and forecasted values, as articulated in Eq (13):

$$MSE = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \tag{13}$$

Ultimately, if the observed data incorporates, for instance, three features aligned with three nodes in the input layer, and the hidden layer encompasses three hidden units, such an MLP can be expressed mathematically as shown in Eq (14):

$$S_1 = \begin{bmatrix} \omega_{11} & \omega_{21} \\ \omega_{12} & \omega_{22} \\ \omega_{13} & \omega_{23} \end{bmatrix} \times \begin{bmatrix} I_1 \\ I_2 \end{bmatrix} + \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} \tag{14}$$

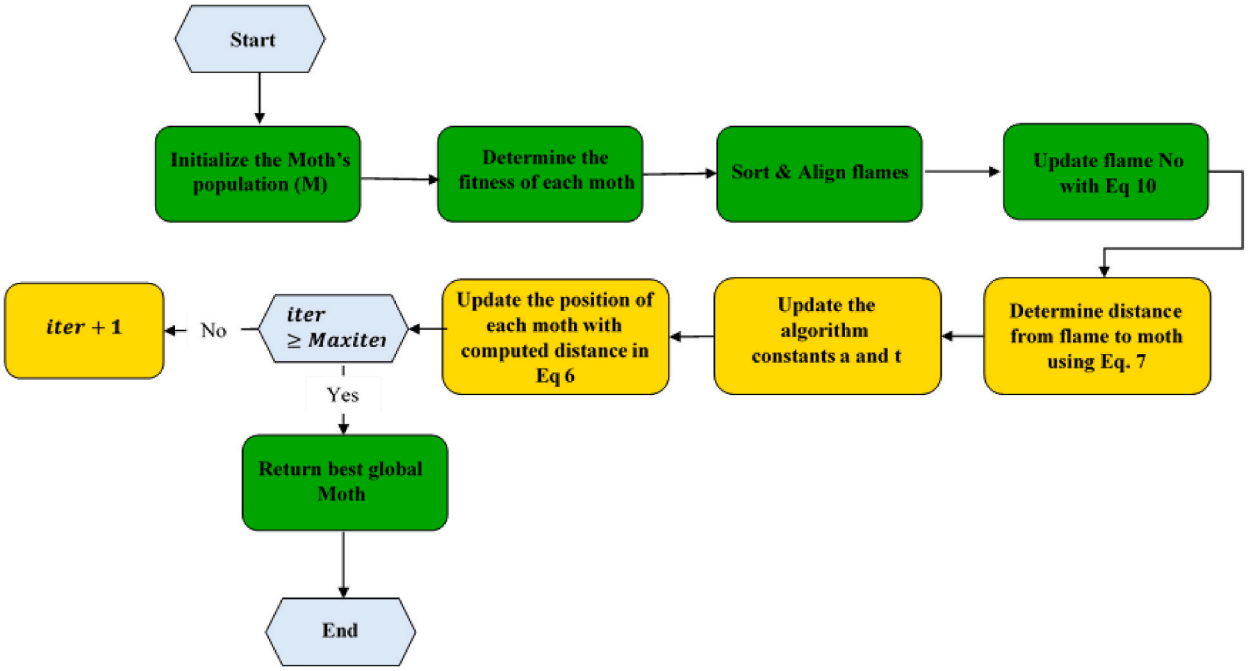


Fig. 2. MFO Flow chart.

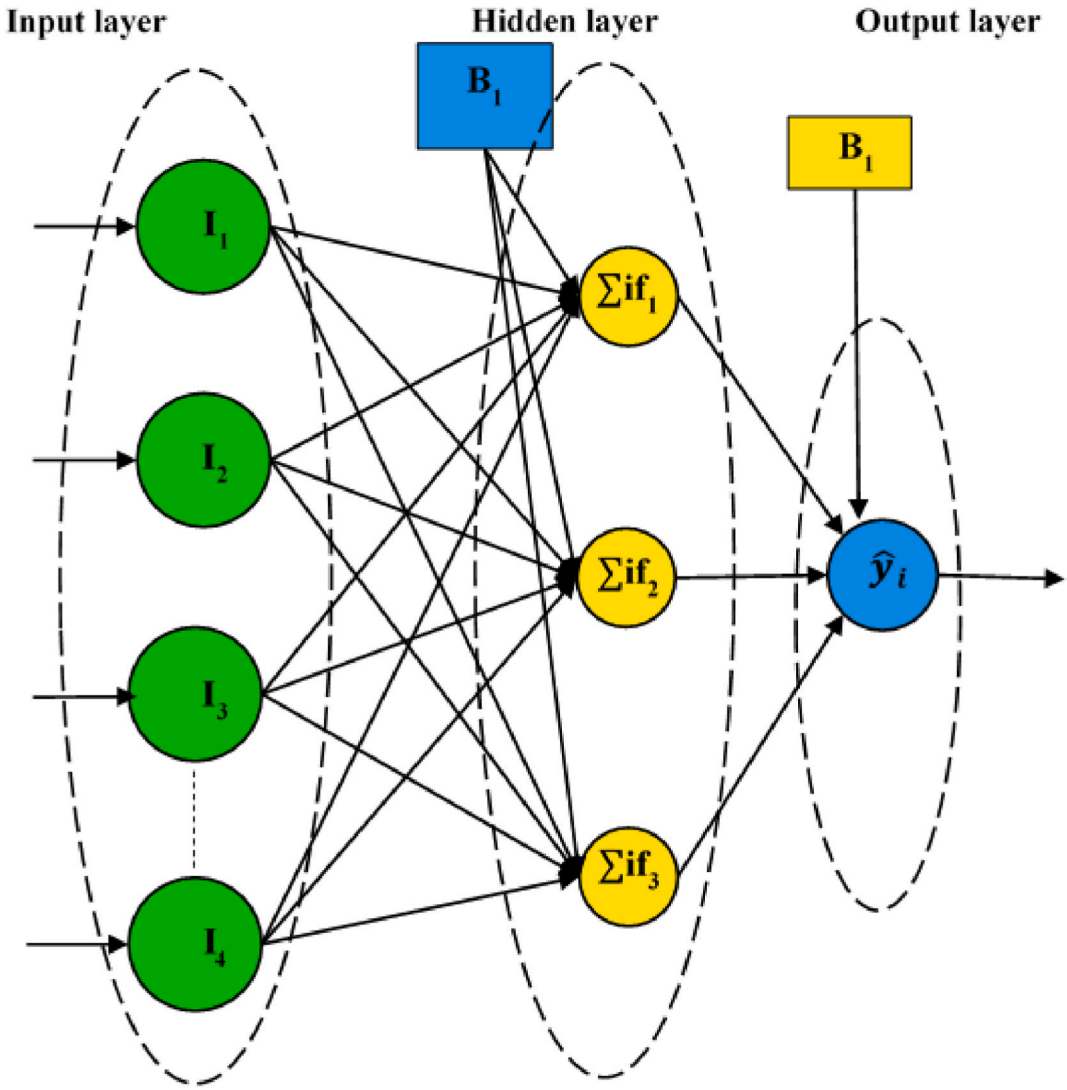


Fig. 3. Mlp structure.

3. Proposed method

3.1. Proposed MFO variant

3.1.1. Worst Moth Disruption Strategy (WMD)

In the iterative process of the MFO algorithm, moths adhere to a mathematical model that guides them toward the optimal flame. As the number of flames decreases, the exploitation process is intensified. While this can yield impressive results, it may also induce inefficiencies in the global search, with moths converging towards the diminishing number of remaining optimal flames. Consequently, some moths may become locally trapped, causing the population to stagnate and reducing their ability to escape local optimal spaces. To address this challenge, the WMD strategy is introduced to bolster the diversity of the population and mitigate the risk of falling into strong exploitation traps. The WMD strategy employs a novel position update equation that incorporates the global worst moth, providing a means to escape the confines of the strong exploitation trap [49]. This strategy is designed to enhance the exploration capabilities of the algorithm, fostering a more robust global search process, as shown in Eq (15).

$$X_i^{t+1} = r_4 * X_{\text{best}}^t - A * \left| C * X_{\text{best}}^t - X_i^t \right| + (1 - r_4) * X_{\text{worst}}^t \tag{15}$$

Where r_4 is an arbitrary float in the interval [0,1], and X_{worst}^t denotes the location of the worst moth, X_{best}^t denotes the location of the best moth. The arbitrary parameters r_4 and $(1–r_4)$ are specified. Due to the unpredictability of r_4, which fluctuates arbitrarily in the range of zero to one, the exploration at this point will take into account both the combined impact of X_{best}^t and X_{worst}^t instead of the best

moth only. X_i^t denote the current position of a given moth. A is obtained as shown in Eq (16) and Eq (17) while C is obtained by Eq (18).

$$A = 2a_1 r_1 - a_1 \tag{16}$$

$$a_1 = 2 - \frac{2iter}{\text{Maxiter}} \tag{17}$$

$$C = 2r_2 \tag{18}$$

Whereas r_1 and r_2 in the aforementioned equations are arbitrary values between [0,1], the resultant value of a_1 falls linearly from 2 to 0.

3.1.2. Proposed WMFO

In the proposed WMFO, the Worst Moth Disruption Strategy (WMD) is specifically designed to address two key challenges encountered in the Moth Flame Optimization (MFO) algorithm: population stagnation and low diversity. In terms of population stagnation, in traditional optimization algorithms like MFO, as the optimization process progresses, there's a risk of the population becoming stagnant, with moths converging towards local optimal solutions. Population stagnation occurs when moths get trapped in local optima, preventing them from exploring the search space further and potentially missing out on better solutions due to a lack of better information by the best flame. WMD disrupts this stagnation by introducing a mechanism to escape local optima. It does this by considering not only the best moth's position (which might be trapped in a local optimum) but also the worst moth's position in the population and random coefficient to adjust the current moth position to encourage it to break free of stagnation. Also, low diversity within the population can hinder the algorithm's ability to explore the search space effectively, leading to suboptimal solutions. WMD aims to increase diversity by leveraging information from the worst-performing moths in the population. By considering the position of the worst moth, WMD encourages exploration of regions of the search space that may have been overlooked by moths focused on local optima. This diversity-promoting aspect of WMD helps prevent premature convergence and encourages moths to explore a broader range of solutions. WMD prevents local trapping. By incorporating information from the worst moth's position, WMD introduces additional diversity into the population, reducing the likelihood of moths getting stuck in local optima. The WMD strategy disrupts the trapping effect by guiding moths away from regions where other moths are concentrated (locally trapped), thus promoting the exploration of different areas of the search space. This action prevents premature convergence and ensures that the algorithm continues to search for better solutions across the entire search space. Finally, the WMD strategy enhances the global search capabilities of the MFO algorithm by promoting the exploration of diverse regions of the search space. Therefore, in a few words, the Worst Moth Disruption Strategy (WMD) effectively addresses population stagnation and low diversity in the Moth Flame Optimization (MFO) algorithm by leveraging information from both the best and worst moths in the population. By disrupting local trapping, increasing diversity, and fostering improved global search capabilities, WMD enhances the overall performance of the optimization algorithm. Furthermore, this section details the implementation of WMFO; the steps below expand each phase of WMFO as given in Fig. 4.

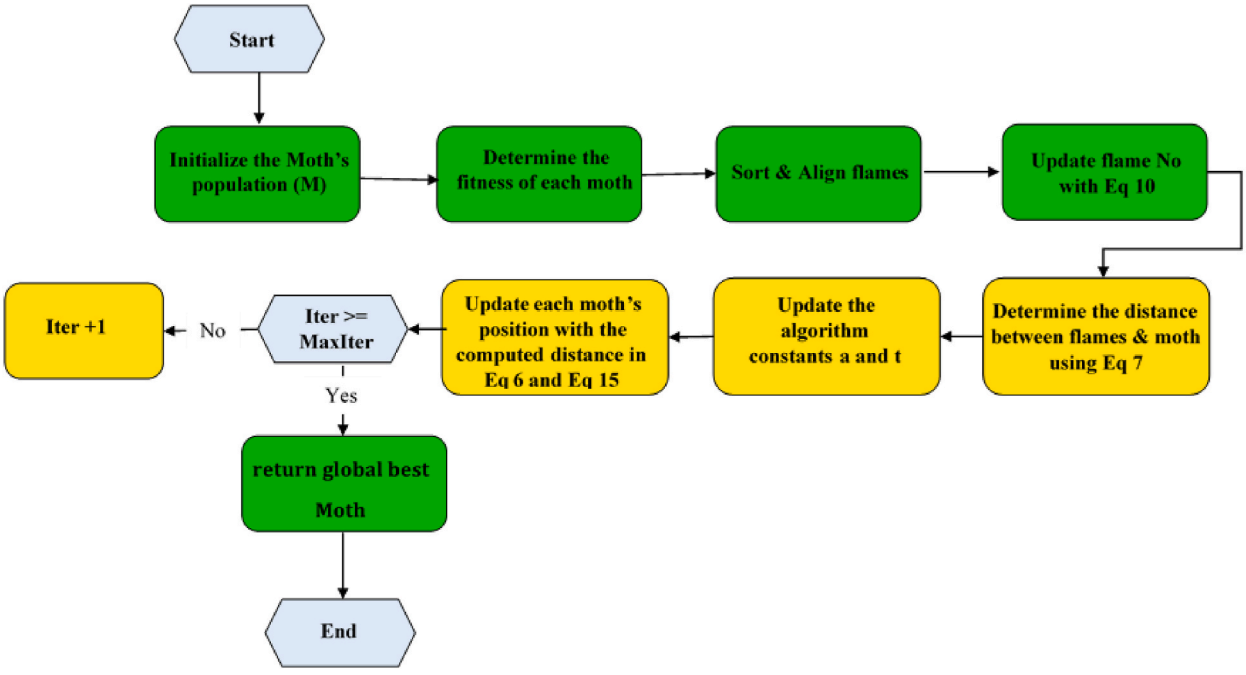


Fig. 4. WMFO Flow chart.

1. Initialization of Moths and Flames: at the initial stage, a population of moths and corresponding flames is randomly initialized. Each moth represents a potential solution to the optimization problem, and its position is determined by a set of decision variables defined by a vector. The flames represent the quality of the solutions and are associated with fitness values, indicating how well they perform in the optimization process.
2. Fitness Evaluation: The fitness of each moth is evaluated by utilizing its respective position to evaluate an objective function (e.g. to train and assess the Multi-Layer Perceptron (MLP) model. This involves using the weights and biases of each moth to adjust the weight and bias of the MLP and compute a performance metric such as Mean Squared Error (MSE)). The results of the objective function serve as a measure of how well each moth contributes to solving the optimization problem.
3. Selection and Ranking of Flames: Once the fitness of each moth is evaluated, the flames are ranked based on their respective fitness values. This ranking helps identify the best-performing solutions within the population. The selection process may involve selecting a subset of flames based on specific criteria, such as the top-performing solutions or a random selection.
4 Update the number of Flames: At the beginning of the iteration, all moths have corresponding flames, so the initial number of flames is typically equal to the total number of moths in the population. As the algorithm progresses, the number of flames is adjusted dynamically to balance exploration (searching new areas) and exploitation (focusing on promising areas).
5 Determine the distance between Flames and Moth: A distance function is used to measure the difference between the positions of a moth (represented by its solution vector) and its corresponding flame. This distance represents how far the moth's current solution is from the current best solution found so far.
6. Update Strategy for Moths: After ranking the flames, the moths' positions are updated based on a strategy that incorporates both exploitation and exploration aspects, as expressed in Equ 6. Furthermore, this work introduced the Worst Moth Disruption Strategy (WMD), which plays a crucial role in updating moth positions. It introduces a novel position update equation that considers the global worst moth's location, aiming to prevent moths from becoming trapped in local optima. The update equation involves adjusting each moth's position based on its current position, the best moth's position, and the worst moth's position, along with random coefficients to introduce variability. The solution of the WMD is compared to the solution of the traditional MFO update strategy, the best solution is adopted by the moth.
7. Termination Criteria: The optimization process iterates until a termination criterion is met. This criterion could be a maximum number of iterations, reaching a certain level of convergence, or other predefined conditions. Upon termination, the best-performing moth (flame) is identified, and its associated parameters (e.g weights and biases of MLP) are extracted for further analysis and application in the MLP model. The parameterization of WMFO involves setting various parameters that govern the algorithm's behavior, such as the population size, maximum number of iterations, coefficients used in the update equation, and termination criteria. These parameters are expressed in the subsequent section.

In conclusion, with the MFO algorithm, moths update their positions based on a mathematical model that guides them towards

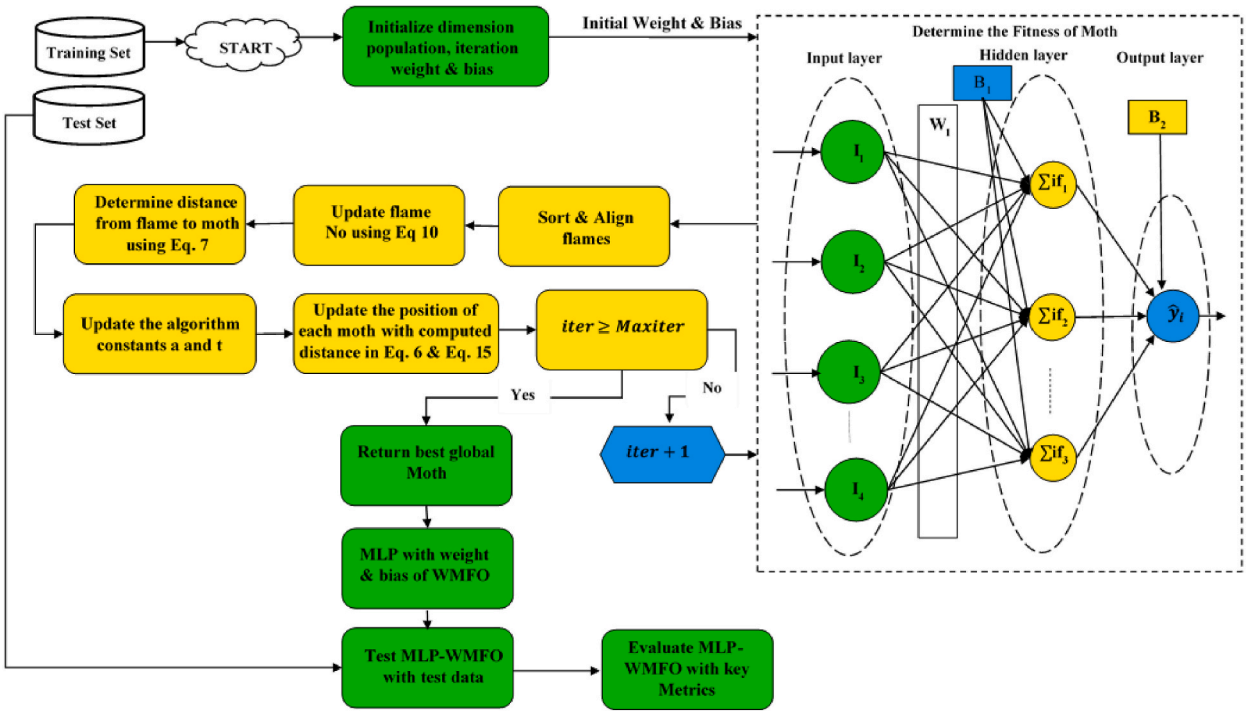


Fig. 5. MLP-WMFO architecture.

optimal solutions (flames). Traditionally, this update equation considers only the position of the best moth (the global optimum) when updating the positions of other moths. WMFO defers from the baseline MFO by utilizing the WMD strategy; this modifies the position update equation by incorporating information from both the best and worst moths in the population. Specifically, it introduces a novel update equation that takes into account the position of the worst moth as well. By considering the position of the worst moth, WMD aims to disrupt local trapping and increase diversity within the population, thus promoting the exploration of different regions of the search space.

3.2. Proposed MLP-WMFO

The depicted flowchart in Fig. 5 illustrates the iterative process of WMFO applied to MLP's performance for CO2 emission prediction. The initial phase involves randomly initializing the weights and biases of each moth within the WMFO, assigning numerical values within the range [−1, 1]. These weights signify inter-neuron connection strength, while biases denote activation function offsets. Subsequently, the fitness of each moth is computed by utilizing its respective weights and biases to train and evaluate the MLP model, employing the Mean Squared Error (MSE) metric for model assessment; the train and test data utilized the 80:20 split ratio. The flames are then ranked based on their fitness, and the number of flames is updated according to equation (10). A crucial step follows, involving the computation of distances between moths and flames, facilitating the update of MFO parameters. The individual moths' positions are adjusted based on the calculated distances, iterating until the maximum specified iteration is reached, if not algorithm's procedure will revert the step of calculating the fitness of each and making adjustments to the decision variable of each moth which comprises of weight and bias. Upon completion, the MLP model is established using the optimal weights and biases derived from the best-ranking moth. The subsequent phase evaluates the MLP-WMFO performance on test data, incorporating key metrics to understand the performance of the model. The process is reiterated until the specified maximum iteration is achieved, ensuring thorough optimization of the MLP for enhanced predictive accuracy. The source of MLP-WMFO is available at https://github.com/MetaHeuLab/MLP-WMFO.

4. Experiments and discussion

4.1. Comparison of WMFO with well-known optimizers on CEC 2015

The WMFO algorithm's performance was assessed through experimentation on fifteen test functions featured in IEEE CEC15. Designated as F1–F15, the Names and Optimal solutions of these functions are expressed in Table 1 [50], and the search range of all the functions is between [−100, 100]. This test suite encompasses a variety of optimization problems, including Unimodal instances (F1–F2), Straightforward Multimodal instances (F3–F5), Hybrid Function instances (F6–F8), and Composite Function instances (F9–F15). Each test case represents a distinct optimization problem with its own characteristics and challenges. The study employed thirty-dimension in the evaluation of these functions, and the outcomes were derived from thirty reiterations. The assessment was conducted under the constraint of a maximum of one thousand function evaluations. The finely tuned parameter configurations for both the proposed WMFO algorithm and other optimizers can be found in Table 2. Peer algorithms in this evaluation are Arithmetic Optimization algorithm (AO) [51], Gaussian Mutation Specular Reflection Learning with Local Escaping Operator Artificial Electric Field Algorithm (GRLEO-AEFA) [52], Moth Flame Optimization (MFO) [32], Sine Cosine Algorithm [53] and Transient Search Optimization(TSO) [54].

The work utilized several performance metrics to assert the performance of WMFO. The low average value of thirty runs

Table 1
Detailed characteristics of CEC15 functions.

$$OE = \left(\frac{N-L}{N}\right) \times 100 \tag{19}$$

Function Number	Function Name	Optimal Solution
F_1	Rotated High Conditioned Elliptic	100
F_2	Rotated Cigar	200
F_3	Shifted and Rotated Ackley's	300
F_4	Shifted and Rotated Ackley's Function	400
F_5	Shifted and Rotated Schwefel's	500
F_6	Hybrid Function 1 (N = 3)	600
F_7	Hybrid Function 2 (N = 4)	700
F_8	Hybrid Function 3(N = 5)	800
F_9	Composition Function 1 (N = 3)	900
F_{10}	Composition Function 2 (N = 3)	1000
F_{11}	Composition Function 3 (N = 5)	1100
F_{12}	Composition Function 4 (N = 5)	1200
F_{13}	Composition Function 5 (N = 5)	1300
F_{14}	Composition Function 6 (N = 7)	1400
F_{15}	Composition Function 7 (N = 10)	1500

demonstrates the consistency of WMFO in achieving optimal solutions. Standard Deviation (STD) provides information about the dispersion or variability of the optimization algorithm's performance across multiple runs. A lower STD suggests more consistent performance. Overall Effectiveness (OE) is a metric used to assess the overall performance of an optimization algorithm across all test functions. It takes into account the number of functions where the algorithm outperforms others, as well as the losses incurred in cases where it performs worse. Wilcoxon Signed-Rank Test is a statistical test used to compare the performance of WMFO against other optimization algorithms in a pairwise manner. It determines whether the differences in performance are statistically significant. The Friedman test is employed to compare the average rankings of different optimization algorithms across multiple test functions. It helps identify whether there are substantial differences in performance among the algorithms.

Table 3 displays the results of 15 benchmark challenges. There are two rows designated for each function to showcase the Average (AVG) and Standard Deviation (STD) values for every algorithm used to solve the corresponding function. Bold values denote superiority over non-bold values. Every algorithm's ideal parameters are selected by analyzing their pertinent literature. According to the experimental results, WMFO finds the best solution more often than its peer algorithms in most function assessments, especially F1, F2, F3, F7, F9, F10, F11, F12, F13, and F14. This implies that WMFO has a clear benefit when handling challenging issues. Furthermore, the STD variation in WMFO is generally minimal, highlighting the algorithm's strong stability. The findings tabulated in Table 3 are further examined to assess the "Overall Effectiveness (OE)" of WMFO in comparison to the compared optimizer. Equation (19) illustrates that the total number of test functions (N) and losses (L) incurred by each algorithm serve as pivotal metrics for determining the OE of the compared algorithms. "Here, N represents the overall number of functions, while L denotes the losses sustained by a given algorithm. In the presented tables, W and T are employed to denote the count of wins and ties" [55,56], respectively. WMFO achieved an OE of 66.6 % over MFO.

The Friedman test is a non-parametric test used to determine if there are statistically significant differences in the performance of multiple algorithms across multiple datasets or test functions [57]. It ranks the algorithms based on their performance and compares the average ranks to assess if there are significant differences. The methodology of the Friedman test is as follows. Let k be the number of algorithms being compared, and let N be the number of test functions or datasets used for evaluation. Rank the algorithms based on their performance on each test function. Ties are assigned the average rank. Calculate the average rank (R_j) for each algorithm across all test functions. Calculate the Friedman statistic (Q) using Eq (20):

$$Q = \frac{12}{Nk(k+1)} \sum_{j=1}^{k} R_j^2 - 3N(k+1) \tag{20}$$

Determine the critical value of (Q) from the chi-square distribution with $k-1$ degrees of freedom at a chosen significance level (e.g., ($\alpha = 0.05$)). If the calculated (Q) statistic is greater than the critical value, reject the null hypothesis, indicating that there are significant differences in performance among the algorithms.

The Wilcoxon signed-rank test is a non-parametric test used to compare the performance of two algorithms paired across multiple datasets or test functions [58]. It evaluates whether the median difference in performance between the paired algorithms is significantly different from zero. The Methodology of Wilcoxon is as follows. Let D_i represent the performance differences between WMFO and another algorithm for each dataset or test function. Calculate the absolute performance differences. $(|D_i|)$ between WMFO and the other algorithm for each function. Rank the absolute performance differences. Calculate the sum of the ranks for positive (T_+) and negative (T_-) performance differences. Calculate the test statistic W using Eq (21):

$$W = min(T_+, T_-) \tag{xx}$$

Determine the critical value of W from the Wilcoxon signed-rank at a chosen significance level (e.g., ($\alpha = 0.05$)). If the calculated W statistic is less than the critical value, reject the null hypothesis, indicating that there is a significant difference in performance between WMFO and the other algorithm. These statistical tests provide a robust methodology for assessing the significance of performance differences between WMFO and other algorithms in the evaluation process. In order to provide a more thorough verification of the experimental results of WMFO, we examined and validated WMFO's performance utilizing the Wilcoxon Signed-Rank Test. Table 4 tabulates the results. It can be seen from the table that WMFO Superior results as "+", which is the total number of functions where WMFO performed better recorded higher values, where "-" denotes the opposite of the former, while " = " denotes similar performance between the compared algorithm and WMFO. Furthermore, the average ranking of WMFO was evaluated using the Friedman test; the results are graphically displayed in Fig. 6 to facilitate a clearer understanding of the comparison results. The results of the Wilcoxon signed-rank test show that, when compared to other algorithms, all the P-values are less than 0.05, this shows that WMFO achieves a significant improvement compared to other algorithms. Friedman test results show a slight variation in WMFO's average ranking;

Table 2
Parameter settings of optimizers.

Algorithms	Parameter setting
AO	$\mu = 0.499$
GRLEO-AEFA	$k_0 = 500$, $\alpha = 30$
MFO	b = 1, a = [-2, −1]
SCA	$a = 2$
TSO	$k = 2$, $z \in [0,2]$ $k = 2$
WMFO	b = 1, a = [-2, −1]

Table 3
Results of AO, GRLEO-AEFA, MFO, SCA, TSO, and WMFO on CEC15.

		AO	GRLEO-AEFA	MFO	SCA	TSO	WMFO
F1	AVG	3.6456E+10	6.9271E+5	7.0909E+9	1.2186E+10	5.8588E+10	**6.2619E+3**
	STD	6.5392E+9	5.8381E+5	3.9516E+9	3.0826E+9	8.1183E+9	**6.1193E+3**
F2	AVG	5.8758E+4	2.7098E+4	1.1631E+5	3.7535E+4	6.1292E+4	**1.1208E+4**
	STD	1.9514E+3	7.7737E+3	3.4627E+4	5.2381E+3	**6.9676E+1**	3.8887E+3
F3	AVG	3.4085E+2	3.2349E+2	3.2702E+2	3.3564E+2	3.4426E+2	**3.1256E+2**
	STD	2.6121	3.9757	4.2014	2.6576	**2.3572**	3.2310
F4	AVG	7.0070E+3	**3.3037E+3**	5.3180E+3	7.7578E+3	9.2309E+3	4.0930E+3
	STD	5.3736E+2	8.0855E+2	8.5195E+2	**3.1988E+2**	4.3118E+2	5.8342E+2
F5	AVG	5.0232E+2	**5.0071E+2**	5.0118E+2	5.0268E+2	5.0511E+2	5.0159E+2
	STD	4.2609E-1	4.7218E-1	6.4460E-1	**2.9684E-1**	8.2698E-1	9.5097E-1
F6	AVG	6.0431E+2	**6.0030E+2**	6.0114E+2	6.0181E+2	6.0538E+2	6.0045E+2
	STD	2.1407E-1	**7.3050E-2**	6.2851E-1	6.3506E-1	3.1832E-1	1.1517E-1
F7	AVG	7.5819E+2	7.0052E+2	7.1743E+2	7.2548E+2	8.0073E+2	**7.0050E+2**
	STD	9.2893	**1.2389E-1**	1.5909E+1	3.3349	1.1998E+1	2.4315E-1
F8	AVG	1.7254E+6	**8.0890E+2**	2.1143E+5	6.8903E+4	1.4591E+7	8.1674E+2
	STD	9.0363E+5	**4.1096**	6.4876E+4	5.4944E+4	9.6767E+6	7.9787
F9	AVG	9.1341E+2	9.1247E+2	9.1337E+2	9.1328E+2	9.1380E+2	**9.1228E+2**
	STD	3.1122E-1	3.4637E-1	3.3708E-1	2.6311E-1	**1.0932E-1**	4.6530E-1
F10	AVG	3.4185E+7	4.4330E+5	1.1923E+6	1.1327E+7	1.0832E+8	**1.5386E+5**
	STD	1.3118E+7	3.1623E+5	9.0587E+5	4.6574E+6	3.8081E+7	**8.0172E+4**
F11	AVG	2.8951E+7	3.4984E+3	4.8496E+3	7.6145E+6	2.3299E+8	**2.1769E+3**
	STD	1.7465E+7	3.2993E+3	4.1318E+3	6.3879E+6	1.0948E+8	**1.0048E+3**
F12	AVG	1.9892E+11	3.5193E+3	4.4556E+3	1.0040E+9	5.4860E+12	**1.7294E+3**
	STD	1.8944E+10	8.5541E+2	1.4501E+3	3.6250E+8	5.0903E+12	**4.0368E+2**
F13	AVG	1.9708E+3	1.5674E+3	1.6400E+3	1.5901E+3	2.8935E+3	**1.5575E+3**
	STD	1.2260E+2	1.5783E+1	3.6200E+1	8.0397	4.6041E+2	**0**
F14	AVG	2.7245E+3	2.3671E+3	2.0240E+3	3.1776E+3	9.4118E+3	**1.6094E+3**
	STD	4.0759E+2	2.5516E+2	6.6618E+1	2.3092E+2	3.0387E+3	**5.9790**
F15	AVG	2.9461E+3	**2.3645E+3**	2.5901E+3	2.9034E+3	3.0529E+3	2.8024E+3
	STD	7.7479E+1	2.8152E+2	1.1835E+2	1.0478E+2	**3.7531E+1**	1.6011E+2
	(L-T-W)	15-0-0	10-0-5	15-0-0	15-0-0	15-0-0	**5-0-10**
	OE	0 %	33.3 %	0 %	0 %	0 %	**66.6 %**

Table 4
Wilcoxon rank sum test for AO, GRLEO-AEFA, MFO, SCA, TSO, and WMFO.

WMFO vs	+/ = /-	R-	R+	P-Value
AO	15/0/0	0	8	0.000655
GRLEO-AEFA	10/0/5	5.80	9.10	0.007829
MFO	13/0/2	4	8.62	0.003143
SCA	15/0/0	0	8	0.000655
TSO	15/0/0	0	8	0.000655

nonetheless, it regularly outperforms GRLEO-AEFA, taking the lead in total performance. In summary, the comparison experiment findings validate that WMFO performs better than its peer algorithms.

To illustrate the superiorities of WMFO in comparison to alternative algorithms, this study documented the optimization search procedures of each algorithm and represented them as an iterative curve, depicted in Fig. 7. The horizontal axis signifies the count of evaluations, while the vertical axis corresponds to the fitness value. Initially, WMFO exhibits commendable convergence accuracy in the unimodal and straightforward multimodal function classifications of F1, F2, and F3, displaying a quicker search pace when juxtaposed with peered algorithms.

Furthermore, a notable observation emerges in the context of hybrid and combinatorial functions, specifically F7, F9, F10, F11, F12, F13, and F14, wherein WMFO demonstrates exceptional outcomes in tackling intricate optimization challenges. Expanding on the graphical representation, WMFO exhibits a distinct edge both during the global search phase of the search process and the local search phase of the iteration, swiftly pinpointing the prevailing global answer. Moreover, an unmistakable inflection point in the form of a noticeable decrease is evident amid the algorithmic iteration concerning the function evaluations of F3 and F4 within WMFO. In contrast, MFO encounters limitations in sustaining progression, signaling the robust capability of WMFO to escape local optima in comparison to MFO. Ultimately, the graphical representations in the figure affirm that WMFO surpasses other algorithms in both search and exploitation capacities, solidifying its status as a high-performance optimization algorithm.

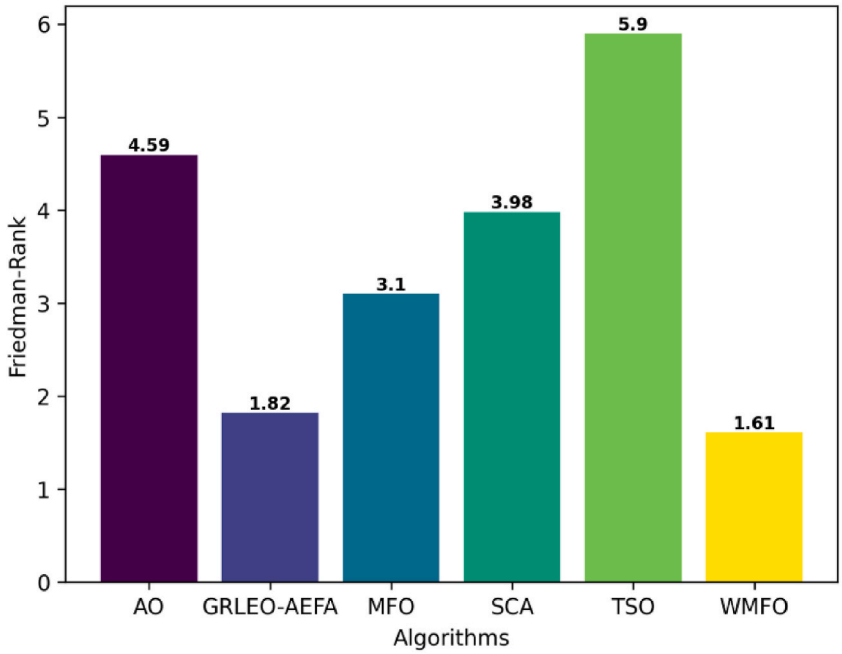


Fig. 6. Friedman rank of AO, GRLEO-AEFA, MFO, SCA, TSO and WMFO

4.2. Carbon prediction analysis with proposed MLP-WMFO

4.2.1. Carbon emission dataset

In this investigation, we explore the factors influencing ecological degradation in Finland across a spectrum of contributors. The dataset for this study was obtained from the World Bank and Our World In Data (OWD) [59,60]. The dependent variable is represented by CO2 levels, while Gas Efficiency, Financial Risk, Oil Efficiency, Economic Growth, Renewable Energy, and Political Risk stand as independent variables. A comprehensive breakdown of these variables is available in Table 5. For methodological consistency, the hybrid model was trained and tested using three-monthly data covering from 1990 to 2021. Correlation Heat map between all features and the relationship between each independent variable and dependent variable are illustrated in Figs. 8 and 9 respectively.

4.2.2. Prediction performance analysis metrics

It is imperative to employ statistical analyses to understand the optimality of WMFO in optimizing MLP, facilitating insight into the effectiveness of MLP-WMFO compared to other hybrid MLP optimizers for Carbon Emission prediction. As a result, a wide range of indicators of effectiveness will be employed to assess every optimizer, providing insightful information on the precision and dependability of the models utilized for prediction [61]. These measures are computed using accepted methods and recommendations found in Ref. [62]. By carefully examining and contrasting the results obtained from these measures, the optimizer with the best accuracy may be identified. Table 6 provides a summary of how these performance measures were calculated and determined.

Here, N represents the data count, σ stands for the standard deviation of observations, Y_i^{Exp} denotes the ith experimental value, Y_i^{MLP} signifies the ith approximated value using MLP, and $\overline{Y}$ represents the mean of the data.

4.3. Carbon prediction experiment, results and discussion

This study uses the MLP to estimate carbon emissions. The performance of a number of optimizers, including the recently proposed WMFO, is assessed to improve the network's process of training. The primary findings from the evaluation are demonstrated and discussed in this section. Each model's effectiveness in enhancing the MLP network's training procedure is carefully analyzed and contrasted. An 80:20 ratio is used to generate training and testing datasets using a random selection procedure from the data. Fig. 10 shows the convergence trajectories for each hybridized MLP model. With 50 iterations taken into account, the convergence is assessed using the MSE metric. Fig. 10 shows that, after 4 iterations, the proposed WMFO optimizer has the best level of convergence, indicated by my MSE of 0.00186. Furthermore, its MSE value maintained the lowest convergence rate at 50 iterations. The MLP-WMFO model's fast convergence rate and low MSE are due to its hybrid optimization strategy, which combines the strengths of both stochastic optimizers and the newly introduced WMD strategy. This allows it to effectively balance exploration and exploitation, leading to faster convergence to the global minimum. The MLP-GRLEO-AEFA model has a moderate convergence rate, and its MSE is slightly lower than that of the MLP-AO model, but both fall behind MLP-WMFO. The MLP-MFO, MLP-SCA, and MLP-TSO models all have slow convergence rates, and their MSEs are significantly higher than those of the other models. This suggests that these optimizers are not effective

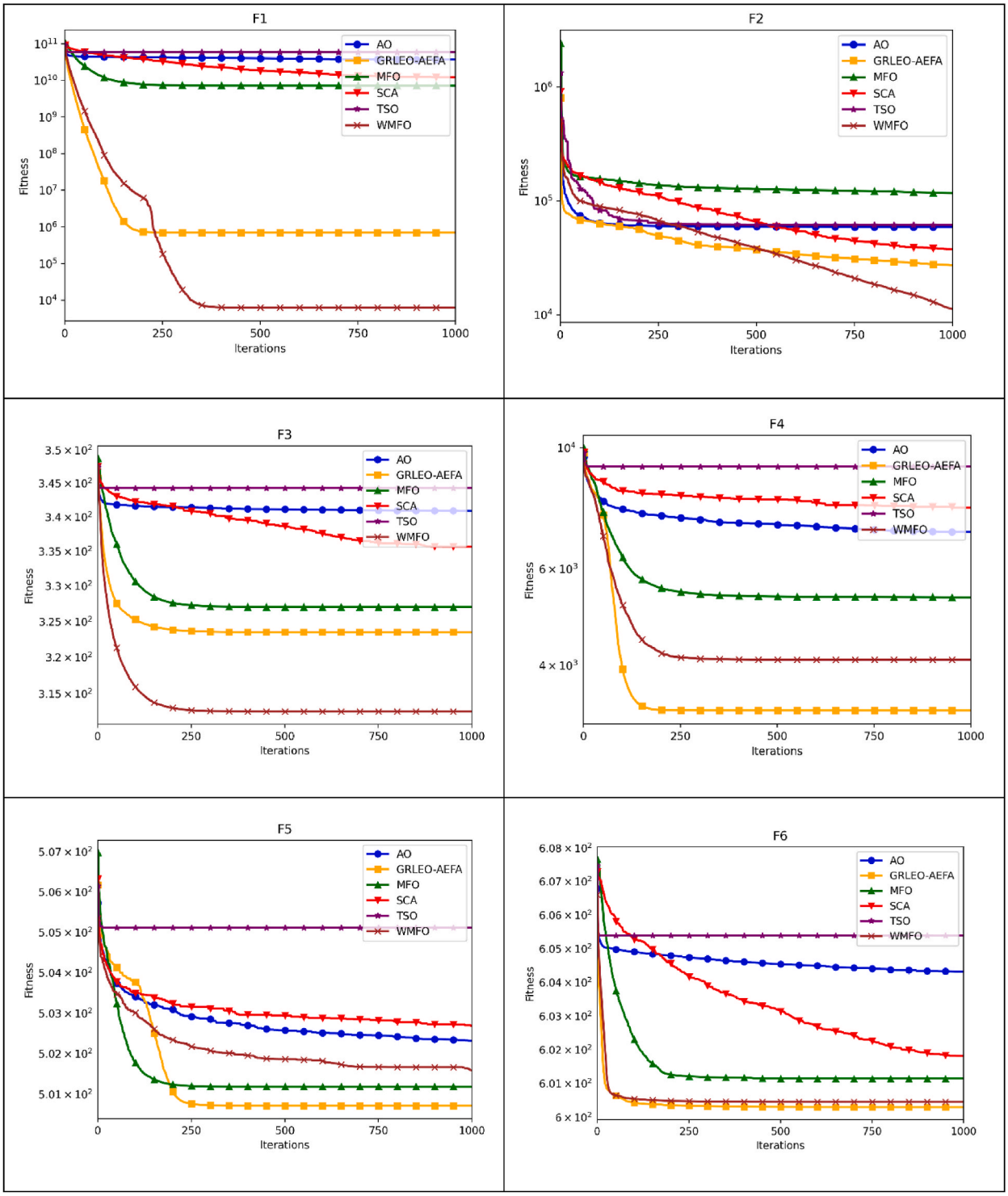


Fig. 7. Convergence plots of AO, GRLEO-AEFA, MFO, SCA, TSO and WMFO on F1–F15.

at either exploration or exploitation, and they may be prone to getting stuck in local minima.

Fig. 11 presents a comparative analysis of actual CO2 values and estimated CO2 values produced by various MLP optimizer-enhanced models during both training and testing phases. The pink lines represent the absolute error rates of each algorithm. As can be observed, the predicted CO2 values closely align with the observed values for most models. Notably, the MLP-WMFO model

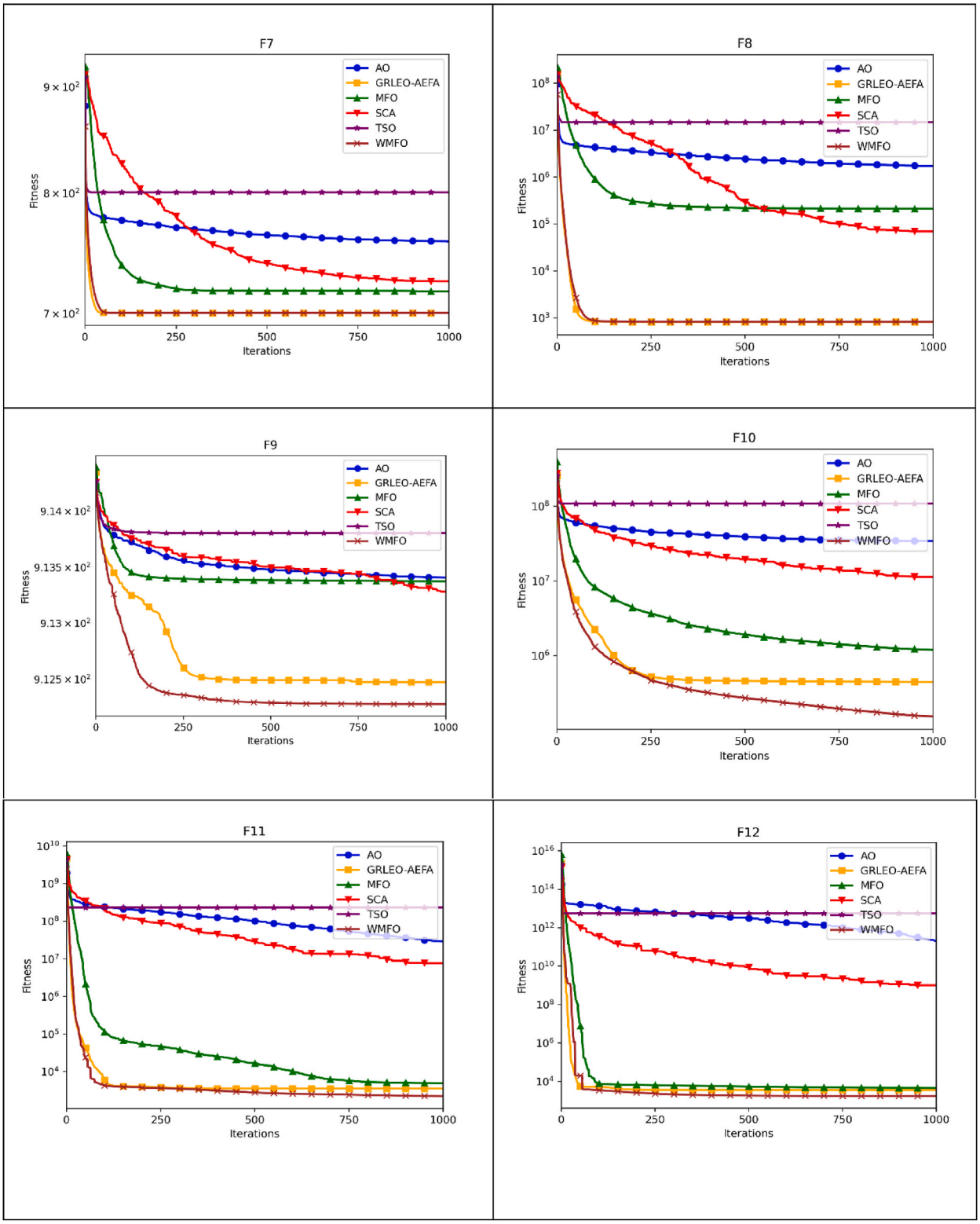


Fig. 7. (*continued*).

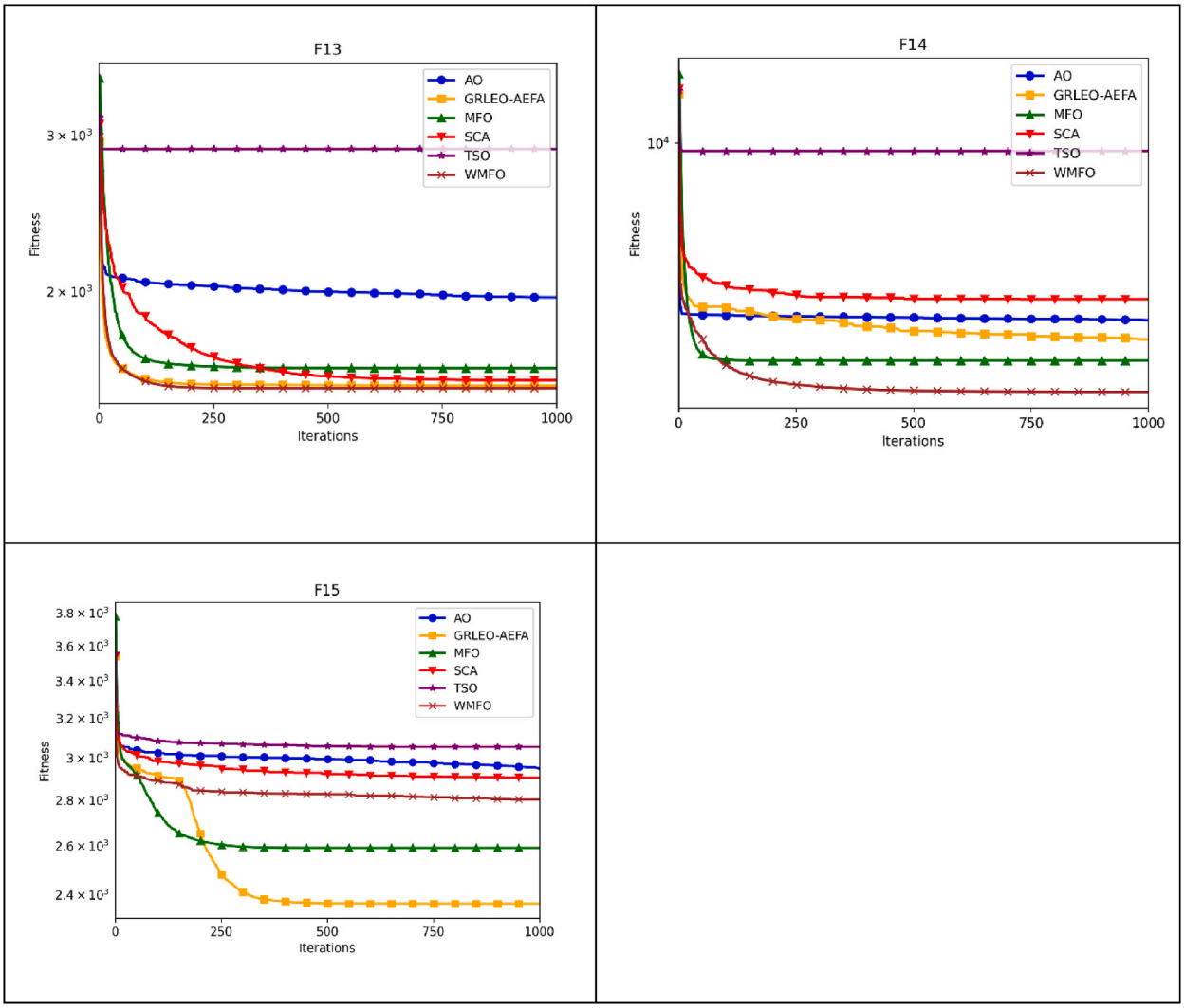


Fig. 7. (*continued*).

Table 5
Data sources and description.

Symbol	Factor	Quantity
GASF	Gas Efficiency	GDP, PPP (constant 2017 international $)/Gas Consumption
FR	Financial Risk	Index
CO_2	CO_2 Emissions	Metric Tonnes Per Capita
GDP	Economic Growth	GDP Per Capita US$ 2015
PR	Political Risk	Index
OIF	Oil Efficiency	GDP, PPP (constant 2017 international $)/Oil Consumption
REC	Renewable Energy	Per Capita

exhibits the most consistent proximity between these two lines, as evidenced by the relatively flat error line. This implies that the MLP-WMFO model generates more accurate CO2 predictions. To elaborate further, the MLP-WMFO model's exceptional accuracy stems from its hybrid optimization strategy, which effectively balances exploration and exploitation. This balance enables the model to effectively navigate the complex landscape of the CO2 prediction problem, leading to more precise estimations. In contrast, the other models exhibit varying degrees of error, with MLP-AO and MLP-GRLEO-AEFA demonstrating slightly higher error rates than MLP-WMFO. This suggests that these models may not be as effective at balancing exploration and exploitation, potentially leading to overfitting. MLP-MFO, MLP-SCA, and MLP-TSO exhibit the highest error rates, indicating that these models struggle to accurately predict CO2 levels. This may be attributed to their inability to effectively balance exploration and exploitation, leading to inefficient

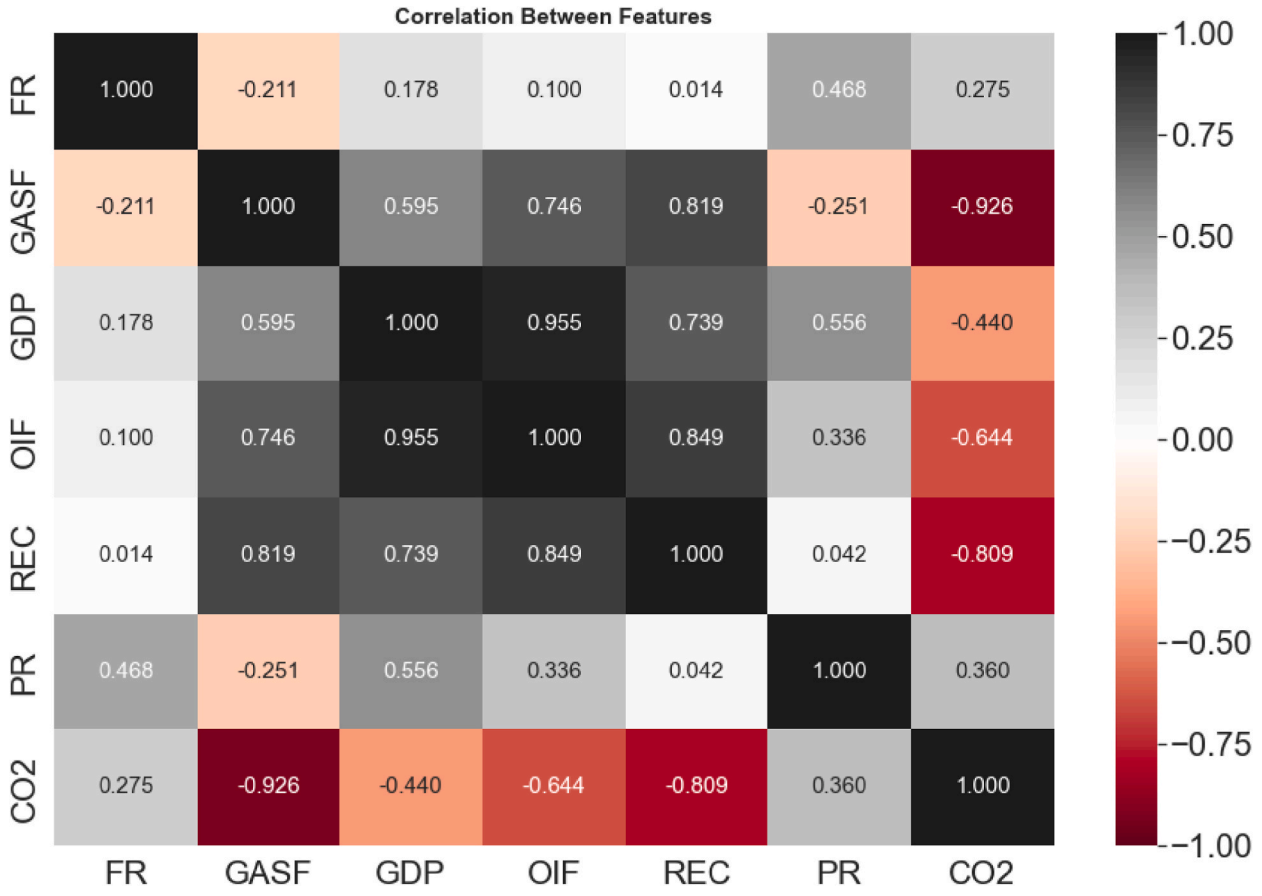


Fig. 8. Correlation heatmap of variables.

optimization and inaccurate estimations. In conclusion, the MLP-WMFO model stands out as the most accurate CO2 prediction model, as evidenced by its consistently low error rates and proximity between predicted and observed values. This superior performance can be attributed to its hybrid optimization strategy with the newly proposed strategy, which enables it to effectively navigate the complexity of the CO2 prediction problem.

Fig. 12 presents a comprehensive analysis of the correlation between actual CO2 values and predicted CO2 values generated by various MLP optimizer-enhanced models. The R^2 values for each algorithm are provided for both test and train datasets, alongside a plot of the predicted vs. actual values. The red dotted line represents the theoretical perfect correlation, where every predicted value exactly coincides with the corresponding actual value. As evident in the plot, the majority of MLP models exhibit a close alignment with the red dotted line, indicating satisfactory performance. However, the MLP-WMFO model stands out as the clear frontrunner, with its data points consistently falling close to the red line. This exceptional correlation is reflected in the highest R^2 value across all models, reaching an impressive 0.978 for the total dataset. The high R^2 value of the MLP-WMFO model translates to an exceptional 97.8 % accuracy between actual and predicted CO2 values. This suggests that the MLP-WMFO model effectively captures the underlying patterns in the data, leading to highly accurate CO2 predictions. In contrast, the other MLP models demonstrate varying degrees of correlation, with MLP-AO and MLP-GRLEO-AEFA exhibiting slightly lower R^2 values compared to MLP-WMFO. This suggests that these models may not be as effective at capturing the complex relationships within the data. MLP-MFO, MLP-SCA, and MLP-TSO exhibit the lowest R^2 values, indicating the most significant discrepancies between predicted and actual CO2 values. This may be attributed to their inability to accurately capture the nuances of the data, leading to less reliable predictions. Finally, the MLP-WMFO model emerges as the most accurate CO2 prediction model, demonstrating an unmatched ability to capture the underlying patterns in the data and generate highly precise estimations. Its exceptional performance is underscored by its highest R^2 value and its consistent alignment with the red dotted line, representing the theoretical ideal correlation.

The outcomes of R^2, NNSE, MSE, MAE, and MSLE for different MA models and MLP throughout training, testing, and the entire experiment are shown in Tables 7–9. Based on a thorough investigation, the MLP-WMFO model shows lower error rates when compared to the MSE, MSLE, and MAE measures of peer optimizer. The models MLP-AO and MLP-GRLEO-AEFA also show somewhat reduced error rates after MLP-WMFO. Furthermore, the NSE and R^2 prediction accuracy indicators demonstrate how well the MLP-WMFO model performs in comparison to the other hybrid models that were examined. The higher NNSE and R^2 values corresponding to MLP-WMFO indicate exceptional results in observing the pattern between the observed and projected values, hence improving

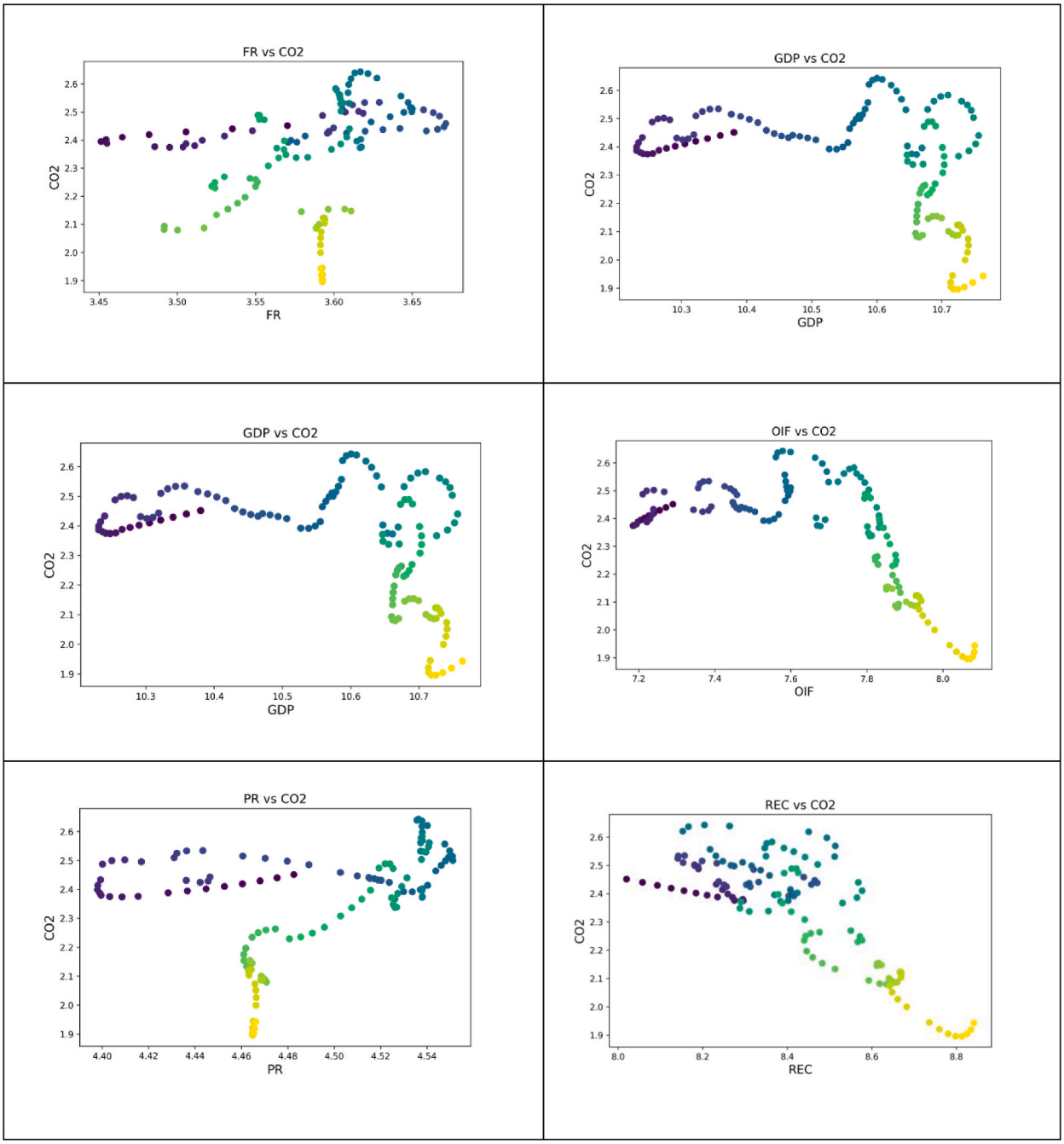


Fig. 9. Illustration of Relationship between each Input variable vs Output Variable.

prediction accuracy. In particular, it is shown that the total R^2 value for MLP-WMFO in CO2 prediction is 0.978. The considerably lesser error rates seen in the MLP-WMFO model showcase its accuracy and efficiency in contrast to the MA models used in the experiment. This highlights the possibility of the MLP-WMFO as a viable approach to improve the MLP's predictive power in CO2 emission prediction. The superior MSE values of the MLP-WMFO model in comparison to other models are especially remarkable. To clarify, the MSE for the MLP-WMFO model is computed to be 0.001085 in the training phase and 0.001865 in the testing phase.

The bar chart in Fig. 13 shows the permutation importance of each feature in an MLP-WMFO model for predicting the occurrence of CO2. Permutation importance is a technique for assessing the importance of features in a predictive model by randomly shuffling the values of each feature and observing the resulting decrease in model accuracy. The more a feature's permutation importance score decreases when its values are shuffled, the more important the feature is to the model [63]. The permutation importance score for each feature is calculated as the difference between the model's performance metric (e.g., accuracy, R-squared) before and after permuting

Table 6
Evaluation metric definition.

Metric	Formula	Definition
NNSE	$NSE = 1 - \frac{\sum_{i=1}^{n}\left(Y_i^{Exp} - \overline{Y}\right)^2}{\sum_{i=1}^{n}\left(Y_i^{Exp} - Y_i^{MLP}\right)^2 + \left(Y_i^{Exp} - \overline{Y}\right)^2}$	Normalized Nash-Sutcliffe Efficiency Coefficient
R^2	$\frac{\sum_{i=1}^{N}\left(Y_i^{Exp} - \overline{Y}^{Exp}\right)^2 - \sum_{i=1}^{N}\left(Y_i^{Exp} - Y_i^{MLP}\right)^2}{\sum_{i=1}^{N}\left(Y_i^{Exp} - \overline{Y}^{Exp}\right)^2}$	Coefficient of Determination
MSE	$\frac{1}{N}\sum_{i=1}^{N}\left(Y_i^{MLP} - Y_i^{ExP}\right)^2$	Mean square error
MSLE	$\frac{1}{N}\sum_{i=0}^{N}\left(\log\left(Y_i^{Exp}+1\right) - \log\left(Y_i^{MLP}+1\right)\right)^2$	Mean Squared Log Error
MAE	$\frac{1}{N}\sum_{i=1}^{n}\left\|Y_i^{Exp} - Y_i^{MLP}\right\|$	Mean absolute error

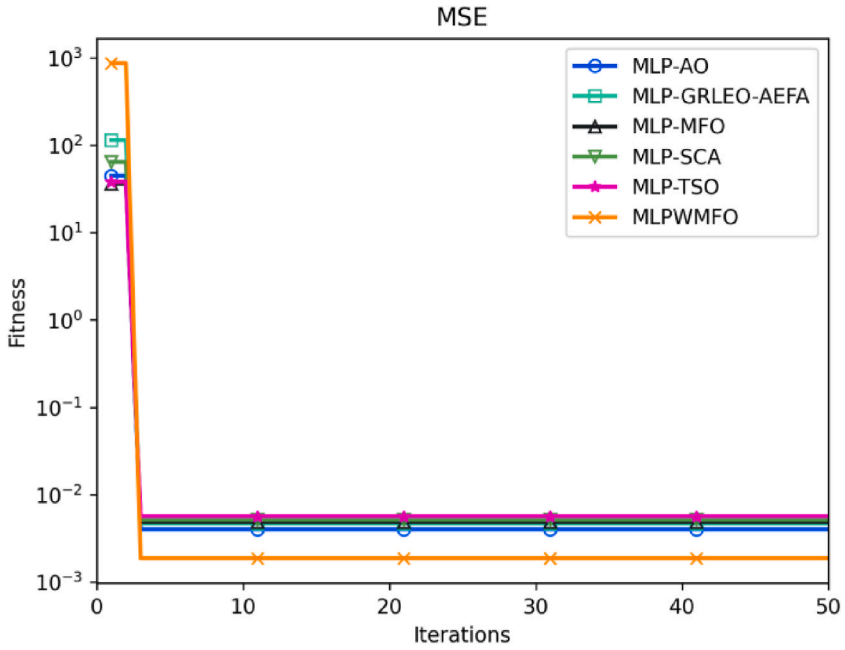


Fig. 10. Convergence trajectories of hybrid MLP models.

the values of that feature. In this case, the most important features are OIF and GDP, which have permutation importance scores of 0.4098 and 0.4062, respectively. This means that shuffling the values of GDP has the largest impact on the model's accuracy. The next most important features are GASF, REC, FR, and PR, which have permutation importance scores of 0.2656, 0.1365, 0.0808, and 0.0655, respectively.

The permutation importance scores can be interpreted as the percentage of the model's predictive power that can be attributed to each feature. In this case, OIF and GDP explain about 40 % respectively of the model's predictive power, GASF explains about 26.5 %, REC explains about 13.6 %, FR explains about 8 %, and PR explains about 6.5 %. The remaining percentage of the model's predictive power is likely due to interactions between the features. For example, the impact of GDP on CO2 may depend on the values of GASF, REC, FR, and PR. Finally, Fig. 13 shows that OIF GDP, GASF, REC, FR, and PR are the most important features for predicting CO2. These features can be used to develop more accurate models for predicting CO2 and Further environmental degradation.

Stakeholders and researchers should consider that improving the efficiency of gas and oil utilization can reduce energy consumption and associated emissions. Implementing energy-efficient technologies, optimizing industrial processes, and promoting energy conservation practices can significantly lower the environmental impact of fossil fuel use since both variables show relatively high importance in the dataset. Also, renewable energy sources such as solar, wind, and hydropower emit significantly fewer greenhouse gases than fossil fuels, thereby reducing CO2 emissions and mitigating climate change. Governments, industries, and individuals can promote renewable energy adoption through incentives, investments, and public awareness campaigns. Finally, political instability and corruption can lead to environmental neglect and unsustainable development. Fostering good governance, promoting environmental policies, and strengthening international cooperation can create a stable and supportive environment for environmental

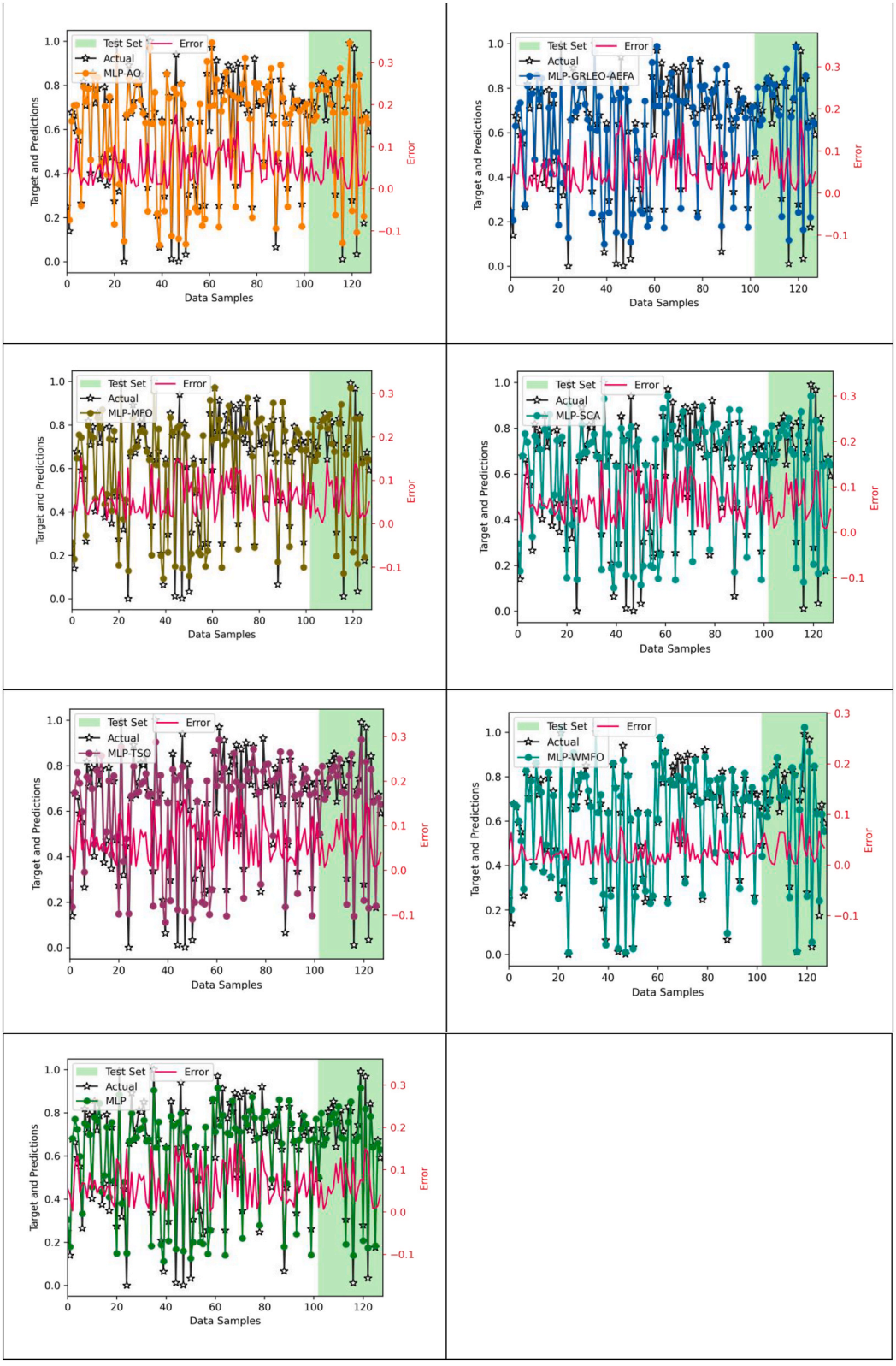


Fig. 11. Observed vs Predicted Data with Error Rate of Hybrid Model and MLP.

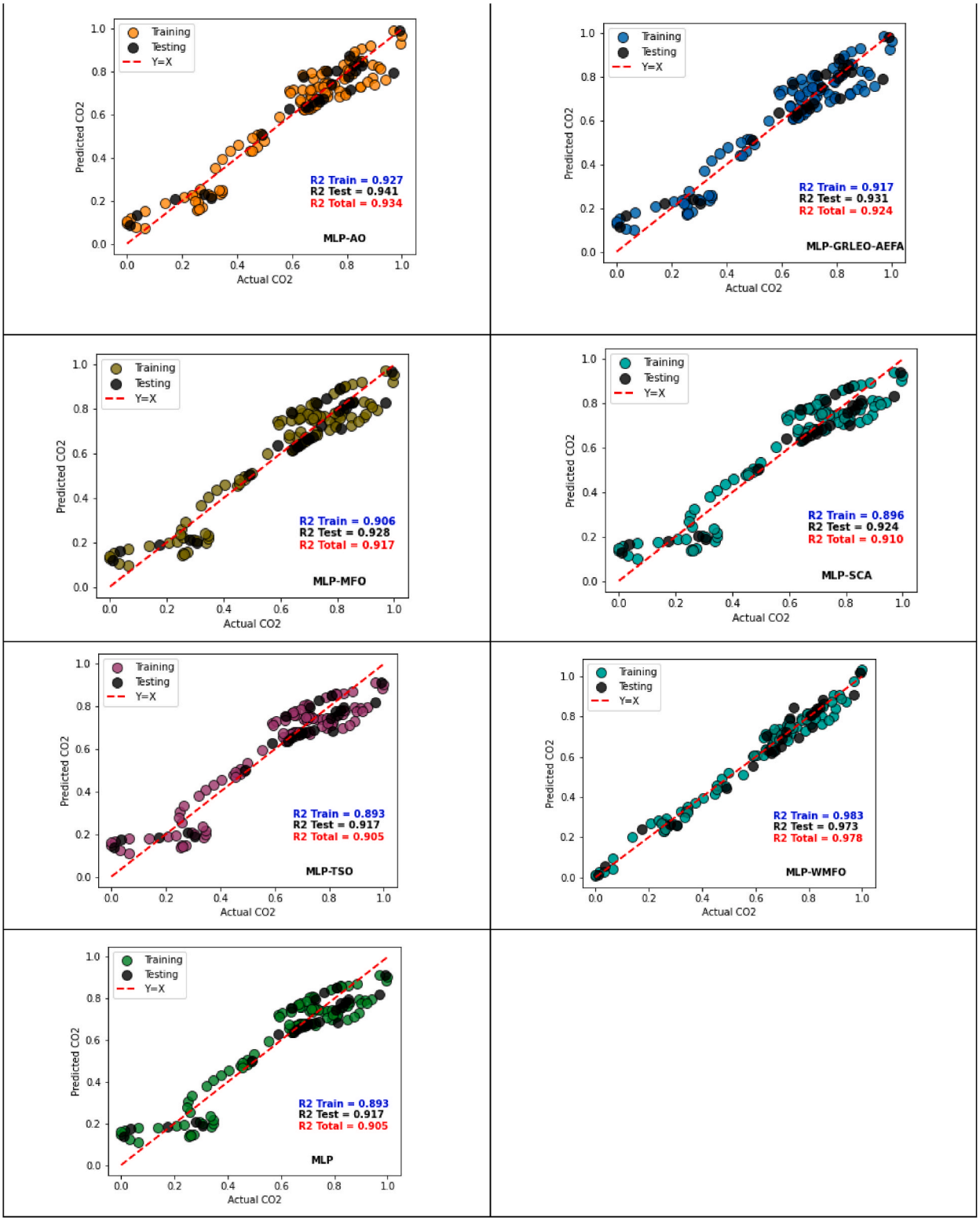


Fig. 12. Scatter Plot of Observed vs Predicted Values of Models.

Table 7
Results of evaluation metrics on MLP and hybrid models on training data.

Mode	NNSE	R2	MSE	MSLE	MAE
MLP-AO	0.932178	0.927243	0.004686	0.002005	0.055434
MLP-GRLEO-AEFA	0.923228	0.916844	0.005355	0.00228	0.059735
MLP-MFO	0.914149	0.906087	0.006048	0.002661	0.065168
MLP-SCA	0.905882	0.896103	0.006691	0.003008	0.068731
MLP-TSO	0.903364	0.893026	0.006889	0.003086	0.069169
MLP-WMFO	**0.983429**	**0.98315**	**0.001085**	**0.00043**	**0.02486**
MLP	0.903364	0.893026	0.006889	0.003086	0.069169

Table 8
Results of evaluation metrics on MLP and hybrid models on testing data.

Model	NNSE	R2	MSE	MSLE	MAE
MLP-AO	0.944507	0.941246	0.003992	0.001729	0.046902
MLP-GRLEO-AEFA	0.935618	0.931188	0.004676	0.00218	0.052182
MLP-MFO	0.933087	0.928289	0.004873	0.002304	0.054531
MLP-SCA	0.929409	0.924047	0.005161	0.002506	0.058899
MLP-TSO	0.92331	0.91694	0.005644	0.002728	0.061907
MLP-WMFO	**0.973289**	**0.972556**	**0.001865**	**0.00072**	**0.034572**
MLP	0.92331	0.91694	0.005644	0.002728	0.061907

Table 9
Results of evaluation metrics total on MLP and hybrid models.

	NNSE	R2	MSE	MSLE	MAE
MLP-AO	0.938343	0.934245	0.004339	0.001867	0.051168
MLP-GRLEO-AEFA	0.929423	0.924016	0.005016	0.00223	0.055959
MLP-MFO	0.923618	0.917188	0.005461	0.002483	0.05985
MLP-SCA	0.917646	0.910075	0.005926	0.002757	0.063815
MLP-TSO	0.913337	0.904983	0.006267	0.002907	0.065538
MLP-WMFO	**0.978359**	**0.977853**	**0.001475**	**0.000575**	**0.029716**
MLP	0.913337	0.904983	0.006267	0.002907	0.065538

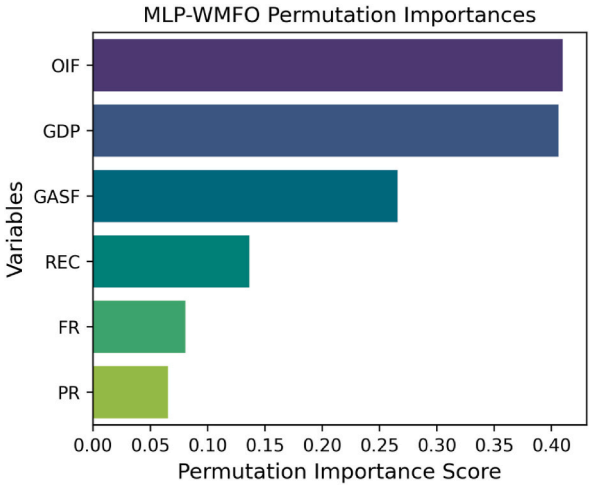


Fig. 13. Feature importance of MLP-WMFO

protection. The proposed WMFO-MLP model for carbon emission prediction holds significant practical implications in real-world scenarios.

1. Improved Prediction Accuracy: By combining the WMFO optimization algorithm with the Multi-Layer Perceptron (MLP) model, the WMFO-MLP model can achieve higher prediction accuracy compared to traditional methods. Enhanced accuracy in carbon

emission prediction is crucial for policymakers, industries, and environmental agencies to make informed decisions and implement effective strategies for mitigating climate change.
2. Better Understanding of Carbon Emission Drivers: The feature importance analysis provided by the WMFO-MLP model offers insights into the relative importance of different factors influencing carbon emissions, such as gas efficiency, economic growth, renewable energy usage, and political risk. Understanding the drivers of carbon emissions enables stakeholders to prioritize interventions and policies aimed at reducing emissions in sectors with the highest impact.
3. Optimized Resource Allocation: The optimization capabilities of WMFO allow for efficient parameter tuning of the MLP model, leading to optimized resource allocation in carbon emission prediction efforts. By allocating resources effectively, organizations can maximize the efficiency of carbon reduction initiatives and minimize associated costs.
4. Early Warning System for Policy Evaluation: The WMFO-MLP model can serve as an early warning system by providing timely and accurate predictions of future carbon emissions based on current socio-economic and environmental factors. Policymakers can use these predictions to evaluate the effectiveness of existing policies and adjust strategies as needed to meet emission reduction targets.

However, it's important to acknowledge the limitations and potential challenges of implementing the WMFO-MLP model in practical applications.

1. Data Availability and Quality: The accuracy and reliability of carbon emission predictions heavily depend on the availability and quality of data on various socio-economic and environmental factors. Data gaps or inaccuracies can introduce bias and uncertainty into the model predictions, limiting its effectiveness in real-world applications.
2. Model Complexity and Interpretability: The complexity of the WMFO-MLP model may pose challenges in terms of model interpretability and transparency. Stakeholders may find it difficult to understand and trust the predictions generated by a complex black-box model, hindering its adoption and acceptance in decision-making processes.
3. Computational Resources: Training and optimizing the WMFO-MLP model will require substantial computational resources, particularly for large-scale datasets and complex optimization tasks. Organizations with limited computational infrastructure or expertise may face challenges in implementing and maintaining the model effectively.
4. Generalization and Robustness: The ability of the WMFO-MLP model to generalize to unseen data and adapt to changing conditions is critical for its practical utility. Overfitting to training data or failure to capture emerging trends and patterns could reduce the model's robustness and reliability in real-world scenarios.

5. Conclusion

CO2 prediction has attracted considerable focus, giving rise to a robust research domain primarily for its capacity to negatively impact the environment. Consequently, this study seeks to predict CO2 emission by employing a novel MLP hybridized WMFO as an extrapolative methodology. In this context, the Multilayer Perceptron (MLP) is chosen as the fundamental machine learning model to analyze training data and recognize correlations among various CO2 factors. To enhance the efficiency of the MLP, a hybridization approach incorporating WMFO into the MLP is implemented. This combination facilitates adept tuning and strategic selection of the most efficient parameters, specifically the weights and biases, for the network. Following this, an extensive evaluation is undertaken to assess the performance of the MLP-WMFO model. Experiment results unveil that the MLP-WMFO hybrid model showcases exceptional performance, attaining the utmost levels of precision, reliability, and efficiency. More specifically, the MLP-WMFO attains noteworthy total R^2 values, reaching 0.978 for the CO2 prediction. These values exceed that of other techniques, signaling that the MLP-WMFO surpasses its counterparts in predictive accuracy. The MLP-WMFO feature importance analysis shows that OIF and GDP have a high impact of up to 40 % on the predictive power of MLP-WMFO, while GASF explains about 26.5 %, REC explains about 13.6 %, FR explains about 8 %, and PR explains about 6.5 %. Regarding limitations, the MLP-WMFO model's efficacy and practicality may be restricted due to a shortfall in accessibility to detailed and high-quality information. Furthermore, additional research is necessary to determine whether the proposed model can be adapted to various locations and climates. Lastly, it is critical to recognize the model's resource requirements and its mathematical complexity. Significant computing resources and duration is required for the training and optimization procedures. To improve the accuracy and resilience of prediction models, future research will look at expanding the scope of data collecting to include larger, more varied datasets from other climate zones. Furthermore, adding continuous information streaming devices might offer an extra fluid and current knowledge base. To improve clarity and streamline the process of establishing policies that would have a good environmental impact, we will also look into methods for clarifying and interpreting the forecasts generated by MLP-WMFO. Furthermore, we aim to address these challenges by developing an open-source library for WMFO-MLP, thereby facilitating accessibility to the methodology for researchers and stakeholders. This endeavor aligns with our commitment to fostering open science practices and promoting transparency in research. By making the WMFO-MLP methodology readily available, we aim to empower researchers and stakeholders to utilize this powerful tool for CO2 prediction and environmental decision-making.

Data availability

The data obtained through the experiments are available upon request.

CRediT authorship contribution statement

Oluwatayomi Rereloluwa Adegboye: Writing – original draft, Resources, Conceptualization. **Ezgi Deniz Ülker:** Methodology, Formal analysis, Conceptualization. **Afi Kekeli Feda:** Software, Resources, Investigation. **Ephraim Bonah Agyekum:** Validation, Supervision, Resources. **Wulfran Fendzi Mbasso:** Writing – review & editing, Methodology, Formal analysis. **Kamel Salah:** Writing – review & editing, Supervision, Resources.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] W. Qiao, H. Lu, G. Zhou, M. Azimi, Q. Yang, W. Tian, A hybrid algorithm for carbon dioxide emissions forecasting based on improved lion swarm optimizer, J. Clean. Prod. 244 (Jan. 2020) 118612, https://doi.org/10.1016/j.jclepro.2019.118612.

[2] A. Rehman, et al., Carbonization and agricultural productivity in Bhutan: investigating the impact of crops production, fertilizer usage, and employment on CO2 emissions, J. Clean. Prod. 375 (Nov. 2022) 134178, https://doi.org/10.1016/j.jclepro.2022.134178.

[3] J. Sillmann, et al., Combined impacts of climate and air pollution on human health and agricultural productivity, Environ. Res. Lett. 16 (9) (Sep. 2021) 093004, https://doi.org/10.1088/1748-9326/ac1df8.

[4] K.O. Yoro, M.O. Daramola, Chapter 1 - CO2 emission sources, greenhouse gases, and the global warming effect, in: M.R. Rahimpour, M. Farsi, M.A. Makarem (Eds.), Advances in Carbon Capture, Woodhead Publishing, 2020, pp. 3–28, https://doi.org/10.1016/B978-0-12-819657-1.00001-3.

[5] S.A. Ward, Assessment of optimum SF/sub 6/-air, SF/sub 6/-N/sub 2/, SF/sub 6/-CO/sub 2/according to particle contamination sensitivity, in: 1999 Annual Report Conference on Electrical Insulation and Dielectric Phenomena (Cat. No.99CH36319), vol. 1, Oct. 1999, pp. 415–418, https://doi.org/10.1109/CEIDP.1999.804676.

[6] Z. Cui, et al., Recent progresses, challenges and proposals on SF6 emission reduction approaches, Sci. Total Environ. 906 (Jan. 2024) 167347, https://doi.org/10.1016/j.scitotenv.2023.167347.

[7] J. Zhang, Y. Zhang, Examining the economic and environmental effects of emissions policies in China: a Bayesian DSGE model, J. Clean. Prod. 266 (Sep. 2020) 122026, https://doi.org/10.1016/j.jclepro.2020.122026.

[8] T. Goh, B.W. Ang, Quantifying CO2 emission reductions from renewables and nuclear energy – some paradoxes, Energy Pol. 113 (Feb. 2018) 651–662, https://doi.org/10.1016/j.enpol.2017.11.019.

[9] S.A. Ward, Optimum SF/sub 6/-N/sub 2/, SF/sub 6/-air, SF/sub 6/-CO/sub 2/mixtures based on particle contamination, in: Conference Record of the 2000 IEEE International Symposium on Electrical Insulation (Cat. No.00CH37075), Apr. 2000, pp. 292–295, https://doi.org/10.1109/ELINSL.2000.845510.

[10] J. Bi, G. Zhou, Y. Zhou, Q. Luo, W. Deng, Artificial electric field algorithm with greedy state transition strategy for spherical multiple traveling salesmen problem, Int. J. Comput. Intell. Syst. 15 (1) (Jan. 2022) 5, https://doi.org/10.1007/s44196-021-00059-0.

[11] M.S. Bakay, Ü. Ağbulut, Electricity production based forecasting of greenhouse gas emissions in Turkey with deep learning, support vector machine and artificial neural network algorithms, J. Clean. Prod. 285 (Feb. 2021) 125324, https://doi.org/10.1016/j.jclepro.2020.125324.

[12] Y.-S. Park, S. Lek, Chapter 7 - artificial neural networks: multilayer perceptron for ecological modeling, in: S.E. Jørgensen (Ed.), Developments in Environmental Modelling, vol. 28, Elsevier, 2016, pp. 123–140, https://doi.org/10.1016/B978-0-444-63623-2.00007-4. Ecological Model Types, vol. 28.

[13] I. Lorencin, N. Anđelić, J. Španjol, Z. Car, Using multi-layer perceptron with Laplacian edge detector for bladder cancer diagnosis, Artif. Intell. Med. 102 (Jan. 2020) 101746, https://doi.org/10.1016/j.artmed.2019.101746.

[14] E. Rozos, P. Dimitriadis, K. Mazi, A.D. Koussis, A multilayer perceptron model for stochastic synthesis, Hydrology 8 (2) (Jun. 2021) 2, https://doi.org/10.3390/hydrology8020067.

[15] X. Feng, G. Ma, S.-F. Su, C. Huang, M.K. Boswell, P. Xue, A multi-layer perceptron approach for accelerated wave forecasting in Lake Michigan, Ocean Eng. 211 (Sep. 2020) 107526, https://doi.org/10.1016/j.oceaneng.2020.107526.

[16] S. Samadianfard, et al., Wind speed prediction using a hybrid model of the multi-layer perceptron and whale optimization algorithm, Energy Rep. 6 (Nov. 2020) 1147–1159, https://doi.org/10.1016/j.egyr.2020.05.001.

[17] B. Mohammadi, Y. Guan, R. Moazenzadeh, M.J.S. Safari, Implementation of hybrid particle swarm optimization-differential evolution algorithms coupled with multi-layer perceptron for suspended sediment load estimation, Catena 198 (Mar. 2021) 105024, https://doi.org/10.1016/j.catena.2020.105024.

[18] F. Ecer, S. Ardabili, S.S. Band, A. Mosavi, Training multilayer perceptron with genetic algorithms and particle swarm optimization for modeling stock price index prediction, Entropy 22 (11) (Nov. 2020) 11, https://doi.org/10.3390/e22111239.

[19] G. Raman Mr, N. Somu, A.P. Mathur, A multilayer perceptron model for anomaly detection in water treatment plants, International Journal of Critical Infrastructure Protection 31 (Dec. 2020) 100393, https://doi.org/10.1016/j.ijcip.2020.100393.

[20] H. Hong, P. Tsangaratos, I. Ilia, C. Loupasakis, Y. Wang, Introducing a novel multi-layer perceptron network based on stochastic gradient descent optimized by a meta-heuristic algorithm for landslide susceptibility mapping, Sci. Total Environ. 742 (Nov. 2020) 140549, https://doi.org/10.1016/j.scitotenv.2020.140549.

[21] W.A.H.M. Ghanem, A. Jantan, A new approach for intrusion detection system based on training multilayer perceptron by using enhanced Bat algorithm, Neural Comput. Appl. 32 (15) (Aug. 2020) 11665–11698, https://doi.org/10.1007/s00521-019-04655-2.

[22] L.B. Almeida, Multilayer perceptrons, in: Handbook of Neural Computation, CRC Press, 1996.

[23] Q.B. Pham, et al., Hybrid model to improve the river streamflow forecasting utilizing multi-layer perceptron-based intelligent water drop optimization algorithm, Soft Comput. 24 (23) (Dec. 2020) 18039–18056, https://doi.org/10.1007/s00500-020-05058-5.

[24] E. Dokur, N. Erdogan, M.E. Salari, U. Yuzgec, J. Murphy, An integrated methodology for significant wave height forecasting based on multi-strategy random weighted grey wolf optimizer with swarm intelligence, IET Renew. Power Gener. 18 (3) (2024) 348–360, https://doi.org/10.1049/rpg2.12961.

[25] E. Ashraf, A.E. Kabeel, Y. Elmashad, S.A. Ward, W.M. Shaban, Predicting solar distiller productivity using an AI Approach: modified genetic algorithm with Multi-Layer Perceptron, Sol. Energy 263 (Oct. 2023) 111964, https://doi.org/10.1016/j.solener.2023.111964.

[26] S. Ahmed, A software framework for predicting the maize yield using modified multi-layer perceptron, Sustainability 15 (4) (Jan. 2023) 4, https://doi.org/10.3390/su15043017.

[27] Meenu Sreedharan, A.M. Khedr, M. El Bannany, A multi-layer perceptron approach to financial distress prediction with genetic algorithm, Automat. Control Comput. Sci. 54 (6) (Nov. 2020) 475–482, https://doi.org/10.3103/S0146411620060085.

[28] A.A. Ewees, M.A. Elaziz, Z. Alameer, H. Ye, Z. Jianhua, Improving multilayer perceptron neural network using chaotic grasshopper optimization algorithm to forecast iron ore price volatility, Resour. Pol. 65 (Mar. 2020) 101555, https://doi.org/10.1016/j.resourpol.2019.101555.

[29] M. Doaei, S.A. Mirzaei, M. Rafigh, Hybrid multilayer perceptron neural network with grey wolf optimization for predicting stock market index, Advances in Mathematical Finance and Applications 6 (4) (Oct. 2021) 883–894, https://doi.org/10.22034/amfa.2021.1903474.1452.

[30] A. Goli, H. Khademi-Zare, R. Tavakkoli-Moghaddam, A. Sadeghieh, M. Sasanian, R. Malekalipour Kordestanizadeh, An integrated approach based on artificial intelligence and novel meta-heuristic algorithms to predict demand for dairy products: a case study, Netw. Comput. Neural Syst. 32 (1) (Jan. 2021) 1–35, https://doi.org/10.1080/0954898X.2020.1849841.

[31] E.-S.M. El-kenawy, et al., Feature selection in wind speed forecasting systems based on meta-heuristic optimization, PLoS One 18 (2) (Feb. 2023) e0278491, https://doi.org/10.1371/journal.pone.0278491.
[32] S. Mirjalili, Moth-flame optimization algorithm: a novel nature-inspired heuristic paradigm, Knowl. Base Syst. 89 (Nov. 2015) 228–249, https://doi.org/10.1016/j.knosys.2015.07.006.
[33] G.I. Sayed, A.E. Hassanien, A hybrid SA-MFO algorithm for function optimization and engineering design problems, Complex Intell. Syst. 4 (3) (Oct. 2018) 195–212, https://doi.org/10.1007/s40747-018-0066-z.
[34] Y.A. Shah, H.A. Habib, F. Aadil, M.F. Khan, M. Maqsood, T. Nawaz, CAMONET: moth-flame optimization (MFO) based clustering algorithm for VANETs, IEEE Access 6 (2018) 48611–48624, https://doi.org/10.1109/ACCESS.2018.2868118.
[35] M.H. Nadimi-Shahraki, M. Banaie-Dezfouli, H. Zamani, S. Taghian, S. Mirjalili, B-MFO: a binary moth-flame optimization for feature selection from medical datasets, Computers 10 (11) (Nov. 2021) 11, https://doi.org/10.3390/computers10110136.
[36] M. Zhang, D. Wang, J. Yang, Hybrid-flash butterfly optimization algorithm with logistic mapping for solving the engineering constrained optimization problems, Entropy 24 (4) (Apr. 2022) 4, https://doi.org/10.3390/e24040525.
[37] A. Helmi, A. Alenany, An enhanced Moth-flame optimization algorithm for permutation-based problems, Evol. Intel. 13 (4) (Dec. 2020) 741–764, https://doi.org/10.1007/s12065-020-00389-6.
[38] F. Wang, X. Liao, N. Fang, Z. Jiang, Optimal scheduling of regional combined heat and power system based on improved MFO algorithm, Energies 15 (9) (Jan. 2022) 9, https://doi.org/10.3390/en15093410.
[39] M.A. Taher, S. Kamel, F. Jurado, M. Ebeed, An improved moth-flame optimization algorithm for solving optimal power flow problem, International Transactions on Electrical Energy Systems 29 (3) (2019) e2743, https://doi.org/10.1002/etep.2743.
[40] V. Jaiswal, V. Sharma, S. Varma, MMFO: modified moth flame optimization algorithm for region based RGB color image segmentation, Int. J. Energy a Clean Environ. (IJECE) 10 (1) (Feb. 2020) 196, https://doi.org/10.11591/ijece.v10i1.pp196-201.
[41] M. Shehab, H. Alshawabkah, L. Abualigah, N. AL-Madi, Enhanced a hybrid moth-flame optimization algorithm using new selection schemes, Eng. Comput. 37 (4) (Oct. 2021) 2931–2956, https://doi.org/10.1007/s00366-020-00971-7.
[42] M.H. Nadimi-Shahraki, A. Fatahi, H. Zamani, S. Mirjalili, L. Abualigah, M. Abd Elaziz, Migration-based moth-flame optimization algorithm, Processes 9 (12) (Dec. 2021) 12, https://doi.org/10.3390/pr9122276.
[43] L. Ma, C. Wang, N. Xie, M. Shi, Y. Ye, L. Wang, Moth-flame optimization algorithm based on diversity and mutation strategy, Appl. Intell. 51 (8) (Aug. 2021) 5836–5872, https://doi.org/10.1007/s10489-020-02081-9.
[44] R.K. Singh, S. Gangwar, D.K. Singh, V.K. Pathak, A novel hybridization of artificial neural network and moth-flame optimization (ANN–MFO) for multi-objective optimization in magnetic abrasive finishing of aluminium 6060, J. Braz. Soc. Mech. Sci. Eng. 41 (6) (Jun. 2019) 270, https://doi.org/10.1007/s40430-019-1778-8.
[45] A. Fathy, H. Rezk, H.S. Mohamed Ramadan, Recent moth-flame optimizer for enhanced solid oxide fuel cell output power via optimal parameters extraction process, Energy 207 (Sep. 2020) 118326, https://doi.org/10.1016/j.energy.2020.118326.
[46] X.-N. Bui, H. Nguyen, Q.-H. Tran, D.-A. Nguyen, H.-B. Bui, Predicting blast-induced ground vibration in quarries using adaptive Fuzzy inference neural network and moth–flame optimization, Nat. Resour. Res. 30 (6) (Dec. 2021) 4719–4734, https://doi.org/10.1007/s11053-021-09968-5.
[47] W. Yamany, M. Fawzy, A. Tharwat, A.E. Hassanien, Moth-flame optimization for training multi-layer perceptrons, in: 2015 11th International Computer Engineering Conference (ICENCO), Dec. 2015, pp. 267–272, https://doi.org/10.1109/ICENCO.2015.7416360.
[48] N. Bacanin, K. Alhazmi, M. Zivkovic, K. Venkatachalam, T. Bezdan, J. Nebhen, Training multi-layer perceptron with enhanced brain storm optimization metaheuristics, Comput. Mater. Continua (CMC) 70 (2) (2021) 4199–4215, https://doi.org/10.32604/cmc.2022.020449.
[49] S. Qiao, et al., Individual disturbance and neighborhood mutation search enhanced whale optimization: performance design for engineering problems, Journal of Computational Design and Engineering 9 (5) (Oct. 2022) 1817–1851, https://doi.org/10.1093/jcde/qwac081.
[50] A. Yadav Anita, N. Kumar, Artificial electric field algorithm for engineering optimization problems, Expert Syst. Appl. 149 (Jul. 2020) 113308, https://doi.org/10.1016/j.eswa.2020.113308.
[51] L. Abualigah, A. Diabat, S. Mirjalili, M. Abd Elaziz, A.H. Gandomi, The arithmetic optimization algorithm, Comput. Methods Appl. Mech. Eng. 376 (Apr. 2021) 113609, https://doi.org/10.1016/j.cma.2020.113609.
[52] O.R. Adegboye, E. Deniz Ülker, Gaussian mutation specular reflection learning with local escaping operator based artificial electric field algorithm and its engineering application, Appl. Sci. 13 (7) (Jan. 2023) 7, https://doi.org/10.3390/app13074157.
[53] S. Mirjalili, SCA: a Sine Cosine Algorithm for solving optimization problems, Knowl. Base Syst. 96 (Mar. 2016) 120–133, https://doi.org/10.1016/j.knosys.2015.12.022.
[54] M.H. Qais, H.M. Hasanien, S. Alghuwainem, Transient search optimization: a new meta-heuristic optimization algorithm, Appl. Intell. 50 (11) (Nov. 2020) 3926–3941, https://doi.org/10.1007/s10489-020-01727-y.
[55] M. Qaraad, S. Amjad, N.K. Hussein, S. Mirjalili, N.B. Halima, M.A. Elhosseini, Comparing SSALEO as a scalable large scale global optimization algorithm to high-performance algorithms for real-world constrained optimization benchmark, IEEE Access 10 (2022) 95658–95700, https://doi.org/10.1109/ACCESS.2022.3202894.
[56] O.R. Adegboye, E. Deniz Ülker, Hybrid artificial electric field employing cuckoo search algorithm with refraction learning for engineering optimization problems, Sci. Rep. 13 (1) (Mar. 2023) 1, https://doi.org/10.1038/s41598-023-31081-1.
[57] O.R. Adegboye, A.K. Feda, O.R. Ojekemi, E.B. Agyekum, B. Khan, S. Kamel, DGS-SCSO: enhancing sand cat swarm optimization with dynamic pinhole imaging and golden sine algorithm for improved numerical optimization performance, Sci. Rep. 14 (1) (Jan. 2024) 1491, https://doi.org/10.1038/s41598-023-50910-x.
[58] O.R. Adegboye, A.K. Feda, O.S. Ojekemi, E.B. Agyekum, A.G. Hussien, S. Kamel, Chaotic opposition learning with mirror reflection and worst individual disturbance grey wolf optimizer for continuous global numerical optimization, Sci. Rep. 14 (1) (Feb. 2024) 4660, https://doi.org/10.1038/s41598-024-55040-6.
[59] World Bank, World Developmental Indicator, 2022 [Online]. Available: https://data.worldbank.org/country/china. (Accessed 7 August 2022).
[60] O.W. in Data, M. Roser, OWID Homepage," *Our World In Data*, Mar. 2024 [Online]. Available: https://ourworldindata.org. (Accessed 30 March 2024).
[61] H. Shakibi, A. Shokri, B. Sobhani, M. Yari, Numerical analysis and optimization of a novel photovoltaic thermal solar unit improved by Nano-PCM as an energy storage media and finned collector, Renew. Sustain. Energy Rev. 179 (Jun. 2023) 113230, https://doi.org/10.1016/j.rser.2023.113230.
[62] A. Rastgoo, H. Khajavi, A novel study on forecasting the airfoil self-noise, using a hybrid model based on the combination of CatBoost and Arithmetic Optimization Algorithm, Expert Syst. Appl. 229 (Nov. 2023) 120576, https://doi.org/10.1016/j.eswa.2023.120576.
[63] B. Gregorutti, B. Michel, P. Saint-Pierre, Correlation and variable importance in random forests, Stat. Comput. 27 (3) (May 2017) 659–678, https://doi.org/10.1007/s11222-016-9646-1.