AI in Sustainable Energy and Environment

Journal homepage: aisesjournal.com

Vol 1, No. 1, 2025
81-107

Western European energy-related CO2 emissions; Integrating data 1990-2020 with cutting-edge swarm intelligence approaches

Serikzhan Opakhai

Institute of Physics and Technical science, L.N. Gumilyov Eurasian National University, Astana, 010000, Kazakhstan

Article Info

Received 30 May 2025

Received in Revised form 15 June 2025

Accepted 28 June 2025

Published online 22 July 2025

DOI: ...

Keywords

CO2 emission

artificial intelligence

metaheuristic algorithm

Western Europe

Abstract

This study aims to provide valuable insights into future emission trends by utilizing advanced predictive modeling techniques. With global energy consumption continuing to rise, understanding and forecasting carbon dioxide (CO2) emissions from energy sources is crucial for policymakers to design effective mitigation measures and transition towards sustainable energy systems. Predicting energy-related CO2 emissions is vital for informing evidence-based environmental policies and strategies to combat climate change. This project investigates the prediction of energy-related carbon dioxide emissions in Western Europe by merging a neural network with three nature-inspired optimization algorithms: Multiverse Optimization (MVO), League Championship Algorithm (LCA), and Evaporation Rate Water Cycle Algorithm (ERWCA). We assess how much this combined approach improves prediction accuracy using a relevant dataset. Our findings demonstrate that the ensemble model works better than alternative methods and has increased accuracy in estimating carbon dioxide emissions, as evaluated by R-squared (R2) and Root Mean Square Error (RMSE). This research provides helpful information for developing sustainability initiatives and regulations by highlighting the advantages of utilizing various optimization techniques in predictive modeling for environmental applications. The accuracy of the MLP is improved by applying the MVO, LCA, and ERWCA algorithms. It was demonstrated that some hybrid techniques can yield more precise predictions than those derived from the conventional MLP ranking. Subsequent analysis revealed that ERWCA outperforms the other algorithms. Using R2 = 0.9977 and 0.9919, RMSE 17.9936 and 30.1394 for ERWCA, R2 = 0.9962 and 0.9898, RMSE 23.3505 and 33.8724 for MVO, and R2 = 0.9898 and 0.9793, RMSE 38.2511 and 48.1272 for LCA, the CO2 emission was estimated with the highest degree of accuracy.

1. Introduction

For several reasons, predicting carbon dioxide (CO_2) emissions is crucial in the context of climate change and environmental sustainability. CO_2 is a significant greenhouse gas responsible for trapping heat in the Earth's atmosphere, leading to global warming and climate change [1]. By accurately predicting CO_2 emissions, we can better understand and anticipate the impact of human activities on the climate system. This information is essential for implementing effective mitigation strategies to limit the rise in global temperatures and minimize the adverse effects of climate change. Predictive models of CO_2 emissions provide valuable insights for policymakers and decision-makers in developing and implementing climate policies and regulations [2]. These models help identify sectors and activities that contribute most to CO_2 emissions, enabling targeted interventions to reduce emissions and transition to low-carbon alternatives. Predicting CO_2 emissions is integral to promoting sustainable development practices. By forecasting future emissions trends, policymakers, businesses, and communities can make informed decisions about investments, resource allocation, and infrastructure

Corresponding author: halilgor@hakkari.edu.tr (H. GÖR)

development [3, 4]. This facilitates the transition towards a more sustainable and resilient economy that balances environmental protection with economic growth and social well-being. Accurate prediction of CO_2 emissions allows for assessing the environmental impact of various human activities, such as energy production, transportation, and industrial processes [5, 6]. Understanding the relationship between CO_2 emissions and environmental degradation helps prioritize conservation efforts, protect ecosystems, and preserve biodiversity. CO_2 emissions are a global challenge that requires collaborative action on an international scale. Predictive emissions modeling provides a common framework for countries to track progress toward emissions reduction targets, share best practices, and engage in climate negotiations [7]. By fostering global cooperation, predictive modeling contributes to collective efforts to address climate change and achieve sustainable development goals [8]. Predicting CO_2 emissions plays a vital role in understanding, mitigating, and adapting to climate change, promoting environmental sustainability, and advancing global efforts towards a more resilient and equitable future.

In recent years, integrating neural networks and optimization algorithms has revolutionized predictive modeling across various domains [9, 10]. Neural networks, inspired by the structure and function of the human brain, have emerged as powerful tools for processing complex data and extracting meaningful patterns. Optimization algorithms, on the other hand, provide efficient methods for tuning the parameters of neural networks to enhance their predictive performance [11, 12]. Neural networks, often called artificial neural networks (ANNs), are computational models composed of interconnected nodes, or neurons, organized in layers [13]. Each neuron receives input signals, processes them through an activation function, and passes the output to subsequent layers. Through a process known as training, neural networks learn to adjust the weights and biases of connections between neurons to optimize their performance on a given task, such as classification or regression. The versatility of neural networks lies in their ability to learn complex, nonlinear relationships from data without requiring explicit programming [14]. This makes them well-suited for tasks involving pattern recognition, time-series prediction, image classification, and natural language processing.

Moreover, advancements in deep learning, which involves training neural networks with multiple hidden layers, have led to breakthroughs in speech recognition, autonomous vehicles, and healthcare diagnostics [15]. Optimization algorithms are crucial in training neural networks by iteratively adjusting their parameters to minimize a predefined loss function [16]. These algorithms seek to find the optimal set of weights and biases that minimize the difference between the predicted outputs of the neural network and the actual targets in the training data. Various optimization algorithms have been developed to tackle the challenge of training neural networks efficiently and effectively [17]. Gradient descent, the most widely used optimization technique, updates the network parameters in the direction of the steepest descent of the loss function. Variants of gradient descent, such as stochastic gradient descent (SGD), mini-batch gradient descent, and adaptive learning rate methods (e.g., Adam, RMSprop), offer convergence speed and stability improvements. In addition to traditional optimization methods, nature-inspired optimization algorithms have gained popularity for optimizing the parameters of neural networks [18]. These algorithms, inspired by natural phenomena or biological processes, mimic the behavior of natural systems to search for optimal solutions in complex search spaces. Examples include genetic algorithms, particle swarm optimization, simulated annealing, and ant colony optimization. The integration of neural networks and optimization algorithms represents a powerful paradigm for predictive modeling, offering flexibility, scalability, and robustness in handling diverse datasets and tasks [19, 20]. By leveraging the capabilities of neural networks to learn from data and the efficiency of optimization algorithms to fine-tune model parameters, researchers and practitioners can develop sophisticated predictive models capable of tackling real-world challenges across various domains [21].

The problem addressed in this research is the need to accurately predict energy-related carbon dioxide (CO_2) emissions in Western Europe. With increasing concerns about climate change and environmental sustainability, there is a growing demand for effective methods to forecast CO_2 emissions and understand their drivers and trends. Traditional approaches to prediction may lack the precision and flexibility required to capture the complex relationships and dynamics inherent in energy systems and environmental processes [22]. Therefore, there is a need to explore innovative methods that can enhance the accuracy and reliability of CO_2 emission predictions, thereby informing policy decisions and supporting efforts to mitigate climate change [23].

To develop a predictive model for energy-related CO_2 emissions in Western Europe using a neural network approach. To investigate the effectiveness of integrating nature-inspired optimization algorithms, specifically Multiverse Optimization (MVO), League Championship Algorithm (LCA), and Evaporation Rate Water Cycle Algorithm (ERWCA), with the neural network model to improve prediction accuracy. To compare the performance of the integrated model with traditional prediction methods in terms of R-squared (R^2) and Root Mean Square Error (RMSE). To analyze the key drivers and trends of CO_2 emissions in Western Europe identified by the predictive model and provide insights for policymakers and stakeholders. To assess the potential implications of the research findings for climate change mitigation strategies, environmental policy development, and sustainable energy planning in Western Europe. By addressing these research objectives, this study aims to contribute to advancing predictive modeling techniques for CO_2 emissions and provide valuable insights for addressing the challenges of climate change and promoting environmental sustainability in Western Europe and beyond.

2. Literature Review

Carbon dioxide (CO_2) emission prediction models, neural networks, and optimization algorithms have been extensively studied in the literature, reflecting the importance of understanding and mitigating the impacts of anthropogenic greenhouse gas emissions on the environment and climate. Numerous studies have focused on developing CO_2 emission prediction models to forecast future emissions trends and assess the effectiveness of mitigation strategies [24-26]. These models often integrate various socioeconomic, demographic, and environmental factors to capture the complex dynamics of energy consumption and emissions [27, 28]. Traditional regression-based approaches, such as linear regression and autoregressive models, have been widely used for CO_2 emission prediction [2, 29, 30]. However, they may struggle to capture nonlinear relationships and complex interactions among variables. Recent advancements in machine learning and data-driven techniques, particularly neural networks, have shown promise in improving the accuracy and flexibility of CO_2 emission prediction models [31-33]. Neural network models, including feedforward neural networks, recurrent neural networks (RNNs), and convolutional neural networks (CNNs), have been applied to capture nonlinear patterns in emissions data and make accurate forecasts. These models can handle large datasets, learn complex relationships, and adapt to changing conditions, offering advantages over traditional approaches [19].

Neural networks have emerged as powerful tools for predictive modeling across various domains, including finance, healthcare, and environmental science [34, 35]. Their ability to learn from data and extract intricate patterns makes them well-suited for classification, regression, and time-series forecasting tasks. In the context of CO_2 emission prediction, neural networks have been applied to analyze historical emission data, identify trends and patterns, and forecast future emissions trajectories. However, challenges remain in training and optimizing neural networks, including selecting appropriate architectures, tuning hyperparameters, and addressing overfitting [36]. Additionally, the interpretability of neural network models can be limited, making it challenging to extract actionable insights and understand the underlying mechanisms driving predictions [37].

Optimization algorithms are crucial in training neural networks by minimizing a predefined loss function and fine-tuning model parameters [38]. Gradient-based optimization methods, such as gradient descent and its variants (e.g., stochastic gradient descent, Adam), are commonly used to update network weights and biases iteratively [39]. These methods are efficient and effective for convex optimization problems but may struggle with non-convex loss surfaces and saddle points. Inspired by biological or natural processes, nature-inspired optimization algorithms have gained popularity for optimizing neural network parameters and addressing the limitations of gradient-based methods [40]. Genetic algorithms, particle swarm optimization, simulated annealing, and ant colony optimization are among the most widely studied nature-inspired algorithms [41-43]. These algorithms offer alternative search strategies, explore diverse regions of the parameter space, and can escape local optima, making them suitable for complex, high-dimensional optimization problems.

The literature review highlights the growing interest in CO_2 emission prediction models, neural networks, and optimization algorithms as tools for understanding and mitigating climate change.

While traditional regression-based models have been prevalent, machine learning techniques, particularly neural networks, offer opportunities for improving prediction accuracy and capturing complex relationships in emissions data [2]. Gradient-based and nature-inspired optimization algorithms are critical in training neural networks and enhancing performance. Future research should focus on developing hybrid models that integrate neural networks with optimization algorithms to improve CO_2 emission predictions further and support informed decision-making for climate change mitigation and environmental sustainability.

While existing research on carbon dioxide (CO_2) emission prediction models, neural networks, and optimization algorithms has made significant progress, several gaps remain that our study aims to address:

1. **Integration of Nature-Inspired Optimization Algorithms with Neural Networks:** While there is ample research on the application of neural networks and optimization algorithms separately, there is a lack of comprehensive studies that explore the integration of nature-inspired optimization algorithms, such as MVO, LCA, and ERWCA, with neural networks for CO_2 emission prediction. Our study seeks to fill this gap by investigating the effectiveness of combining these algorithms with neural networks to enhance prediction accuracy.
2. **Evaluation of Multiple Optimization Algorithms:** Many existing studies focus on a single optimization algorithm or compare only a few alternatives. Our research aims to broaden the scope by evaluating the performance of three distinct nature-inspired optimization algorithms in combination with neural networks. By comparing the effectiveness of MVO, LCA, and ERWCA, we can provide insights into the relative strengths and weaknesses of different optimization strategies for CO_2 emission prediction.
3. **Assessment of Predictive Performance Metrics:** While some studies evaluate prediction accuracy using standard metrics such as R-squared (R^2) and Root Mean Square Error (RMSE), there is a need for more comprehensive evaluation frameworks that consider additional performance metrics and assess model robustness across different datasets and scenarios. Our study aims to address this gap by rigorously evaluating the predictive performance of the integrated model using multiple metrics and conducting sensitivity analyses to assess model stability and generalization capabilities.
4. **Analysis of Policy Implications:** Despite the importance of CO_2 emission prediction for informing climate policies and mitigation strategies, many existing studies focus primarily on technical aspects of modeling without considering the broader policy implications of their findings. Our research seeks to bridge this gap by analyzing our predictive model's policy implications, providing insights for policymakers and stakeholders in designing effective climate policies and promoting sustainable energy practices in Western Europe.

By addressing these gaps in current research, our study aims to contribute to advancing predictive modeling techniques for CO_2 emission prediction and provide valuable insights for decision-makers in addressing the challenges of climate change and environmental sustainability.

3. Materials and methods

In our methodology, we employed a hybrid approach that combines the strengths of two distinct predictive modeling techniques: neural networks and optimization algorithms. This hybrid method harnesses the power of neural networks to capture complex relationships and patterns in the data while simultaneously leveraging optimization algorithms to fine-tune model parameters and enhance predictive accuracy. Specifically, we utilized MVO, LCA, and ER-WCA as optimization algorithms, each offering unique advantages in optimizing the neural network structure. By integrating these algorithms into the training process, we aimed to improve the model's ability to learn from data and generate accurate predictions of energy-related CO_2 emissions in Western Europe. This hybrid methodology represents a novel approach to predictive modeling, offering a promising avenue for enhancing the performance and robustness of CO_2 emission prediction models. Only modeling methodologies, modeling method validation, and optimization algorithm analysis may lead to achieving the abovementioned objectives. Figure 1 shows these phases, which are explained in greater detail below.

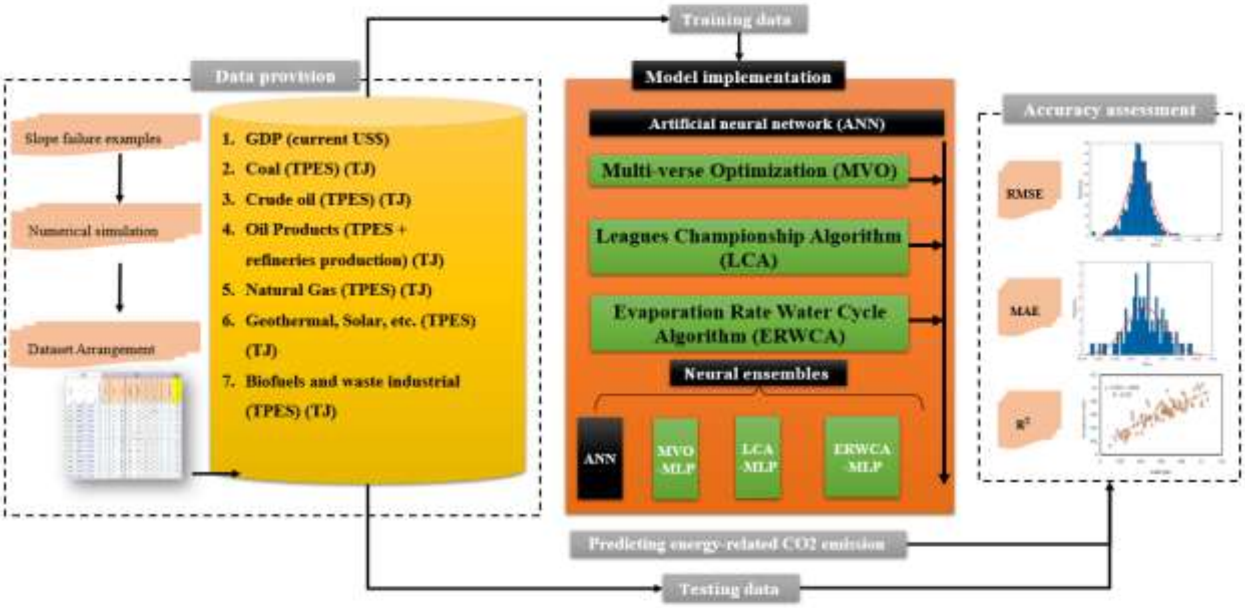


Figure 1. An outline of the modeling procedure

3.1. Artificial Neural Network

A two-layered feedforward neural network from the Matlab ANN Toolbox was used to predict carbon dioxide emissions. The Levenberg-Marquardt method from the Matlab ANN Toolbox was used to train the ANN network. Three components comprise the ANN: an output layer with a linear output function, a hidden layer with a sigmoid activation function, and an input layer. Using a random initialization increased the accuracy of the forecasts. The buried layer's sigmoid transfer function allows for handling nonlinear data. The outcome is condensed to lie between 0 and 1, with the input ranging from plus to negative infinity [44]. The sigmoid activation function is shown in equation (1).

$$f(x) = 1/(1 + exp^{-x}) \tag{1}$$

The output neuron determined the amount of carbon dioxide released, while the input neurons watched the variable data. It was shown that the best model structure was obtained when the number of hidden neurons was increased from 1 to 10. Thirty percent of the data set went into creating training and test data sets (70 percent). The network has been trained to determine the ideal weights for cost-effectiveness. A cost function system was implemented to ascertain the best fit across the model iterations. The training was stopped after the error reduction failed six times a row to avoid overfitting. The primary outcome of this study was the estimation of carbon dioxide emission, which was ultimately projected using the best predictive network. Four statistical indicators were calculated for the models for both training and testing. The R^2, mean squared error (MSE), root mean square error (RMSE), and mean absolute error (MAE) are all described in equations (2)-(5). These statistical criteria were used to estimate the derived models' accuracy. For instance, the model's accuracy was evaluated using RMSE, while its robustness was evaluated using R^2. In this case, the measured value is y_k, the predicted value is $\widehat{y_k}$, and the mean value of, y_k is $\bar{y}$. There are n samples in total.

$$\text{RMSE} = \sqrt{\left(\sum_{k=1}^{n}(\widehat{y_k} - y_k)^2\right)/n} \tag{2}$$

$$\text{R}^2 = 1 - \left(\frac{(\sum_{k=1}^{n}(\widehat{y_k} - y_k)^2)}{(\sum_{k=1}^{n}(y_k - \bar{y})^2)}\right) \tag{3}$$

$$\text{MSE} = \left(\sum_{k=1}^{n}(\widehat{y_k} - y_k)^2\right)/n \tag{4}$$

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}(\widehat{y_k} - y_k) \tag{5}$$

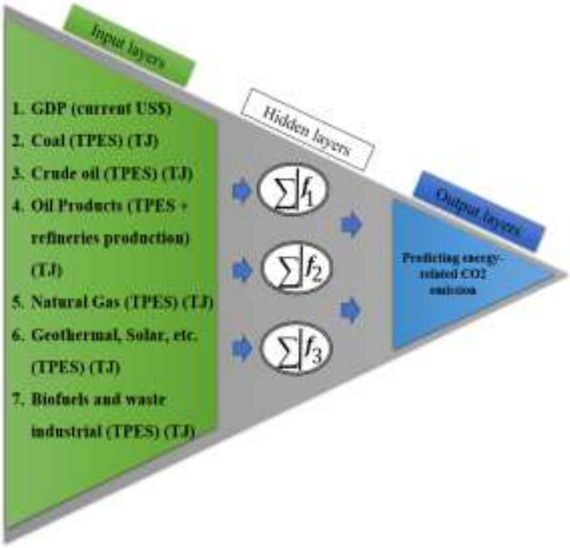


Figure 2. Diagram illustrating the ANN algorithm

3.1. Multiverse Optimization (MVO)

According to [45], Wormholes, white holes, and black holes are the three main pillars of the multiverse theory in physics, and they are the mathematical models created using the MVO approach. Let each variable in the optimization problem reflect one of the following universes concerning laws. A variation in the inflation rate affects some but not all items in the universe through wormholes that lead to the ideal state. Objects are more likely to pass via white holes in universes with higher inflation rates and those with lower inflation rates through black holes. Higher inflation rates are linked to white holes, whereas lower inflation rates are linked to black holes. The MVO algorithm is described as follows in brief:

Step 1: Set the universe's initial values, maximum repetitions, maximum iterations, interval variable *[lb, ub]*, and universe location.

Step 2: To locate a white hole based on the inflation rate of the universe, use a roulette wheel selection approach.

$$x_i^j = \{x_k^j \;\; r1 < NI(U_i) \;\; x_i^j \;\; r1 \geq NI(U_i) \tag{6}$$

where $r1$ is a randomly generated number from the interval [0, 1]; U_i is the *i*th universe; x_i^j is the *i*th universe's *j*th parameter; x_k^j is the *k*th universe's *j*th parameter selected by the roulette process; and $NI(U_i)$ is the universe's normative inflation rate.

Step 3. It is time for a wormhole existence probability (WEP), a travel distance rate (TDR) update, and a boundary check.

$$WEP = min + l \cdot \left(\frac{max - min}{L}\right) \tag{7}$$

$$TDR = 1 - \frac{l^{\frac{1}{p}}}{L^{\frac{1}{p}}} \tag{8}$$

The numbers l for the current iteration, L for the maximum number of repetitions, and p for the accuracy of the exploitation stand for the highest and lowest *WEP* values, respectively. In the MVO model, low *WEP* and high *TDR* encourage exploration and the avoidance of local optima, whereas high WEP and low *TDR* enhance exploitation [46].

Step 4: Find the current inflation rate in the universe. The cosmos shifts if the rate of inflation rises over its present value. In all other circumstances, the cosmos seems to continue existing.

Step 5: Update the position of the universe as provided by Equation (13).

$$x_i^j = \begin{cases} \begin{cases} X_j + TDR\left((ub_j - lb_j)r4 + lb_j\right) & r3 < 0.5 \\ X_j - TDR\left((ub_j - lb_j)r4 + lb_j\right) & r3 \geq 0.5 \end{cases} & r2 < WEP \\ x_i^j & r2 \geq WEP(i) \end{cases} \tag{9}$$

where *r2, r3,* and *r4* are random values chosen from the range [0, 1]; ub_j is the *j*th variable's upper bound; and, lb_j is its lower bound. Where X_j is the *j*th parameter of the best universe at that instant.

Step 6: criteria for termination. If the prerequisites for termination are met, the required output is produced. If not, an extra iteration is performed, and Step 2 of the procedure is followed.

3.3. League Championship Algorithm (LCA)

Like other evolutionary algorithms, the LCA operates on a population of people [47]. As a result, during the initialization stage, a league (population) of *L* (the league size) teams (solutions) is formed, and their playing characteristics (fitness values) are evaluated. Every team will have *n* players if we analyze a function with *n* variables, where *n* is the number of variables. For now, the setups that work well for the teams take advantage of the starting settings. The competition is the next stage. According to the league schedule, the clubs play each other in pairs for $S \times (L - 1)$ weeks, where S is the number of seasons and t is the week. There is no tie regarding the results of the games or matches between teams I and J. Wins and losses are shown for each outcome. The performance of

each squad determines this. Every team designs a new configuration during the recuperation time based on what performed well in the previous week's play and what is currently its finest formation. The selection process in LCA is voracious. It swaps out the current configuration for the best one with a more powerful and efficient one. Stated differently, the new configuration is the best alternative for the team. It should be considered the fittest one if it is the best answer discovered thus far for the *i*th member of the population. Upon meeting the halting criterion, the algorithm terminates.
A few terms we used in explaining the LCA technique need to be defined and thoroughly explained. These concepts include creating the league schedule and determining whether the team is winning or losing. Further information on these ideas is provided in the sections that follow.

3.3.1. Generating a league schedule

Creating a schedule containing every game for every season is the first step in creating the illusion of a championship setting, complete with teams vying for supremacy. Throughout the season, each team plays each other once in a round-robin style. Since $L/2$ matches would be played in parallel during each of the $(L-1)$ weeks if there are L (an even number of teams), there will be $L(L-1)/2$ matches (if L is an odd number, there would be L weeks with $(L-1)/2$ matches and one team would play no games during any given week). After that, the championship lasts for S more seasons [47].

3.3.2. Determining winner/loser

Each squad participates in the LCA and plays against other squads; no team may win or lose a game. After a game, a team's result is determined stochastically using the playing strength criterion as long as the likelihood of a team winning is commensurate with its fit level. According to [47], The degree of fit is determined by the distance with an ideal reference point and is associated with the team's playing strength.

3.4. Evaporation Rate Water Cycle Algorithm (ERWCA)

Sadollah, et al. [48] introduced a novel search strategy called the evaporation rate-water cycle algorithm (ER-WCA). This approach modifies the WCA technique as originally proposed [48]. Two instances of how nature influences the WCA algorithm are the water cycle and water flowing toward the ocean. During the hydrological cycle, water from streams evaporates and is used by plants for photosynthesis. Once the vapor enters the atmosphere, it condenses as clouds. Depending on the weather, water re-enters the earth in various states. According to this system, rivers are excellent persons, while other water flows are called streams. In the event when K represents the issue's magnitude, the potential streams are $x_1, x_2, \dots, x_k$. The initial population is created at random, as seen below:

$$\text{Total population} = \begin{bmatrix} Sea \\ River_1 \\ River_2 \\ \vdots \\ Stream_{K_{sr+1}} \\ Stream_{K_{sr+2}} \\ \vdots \\ Stream_{K_{pop}} \end{bmatrix} \tag{10}$$

$$= \begin{bmatrix} x_1^1 & x_2^1 & \dots & x_k^1 \\ x_1^2 & x_2^2 & \dots & x_k^2 \\ \vdots & \vdots & \vdots & \vdots \\ x_1^{K_{pop}} & x_2^{K_{pop}} & \cdots & x_N^{K_{pop}} \end{bmatrix}$$

where the swarm size is indicated by K_{pop}. The intensity of flow for each approach is computed using Equation 10:

$$\text{Cost}_i = f\left(x_1^i, x_2^i, \dots, x_K^i\right) \quad I = 1, 2, \dots, K_{pop} \tag{11}$$

Because rivers and seas are made up of the best-performing individuals, $K_{Streams}$ denotes the portion of the population that may yet flow into rivers or the sea. The amount of water drawn from the sea or river varies depending on the strength of the flow. The bellows show the estimated distribution of streams to each river and the sea.

$$C_n = Cost_n - Cost_{K_{sr+1}} \\ n = 1, 2, \dots, K_{sr} \tag{12}$$

$$NS_n = round\{\left|\frac{C_n}{\sum_{n=1}^{K_{sr}} C_n} \times K_{Streams}\right| \tag{13}$$

The symbols indicate the number of streams flowing toward a particular river or sea NS_n. The fitness function distributes streams proportionally between rivers and the sea since more streams flow into the sea. In the natural world, specific streams unite to create new rivers. Figure 3 illustrates the path a stream travels in the direction of a river when there is only one sea and K_{sr-1} rivers among a population of K_{pop} people. Additional information on the proposed methodology may be found in related papers [49].

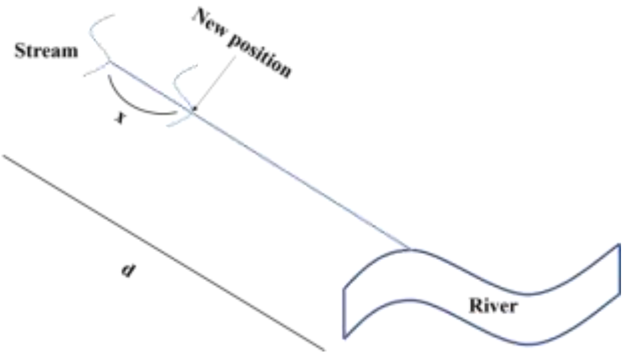
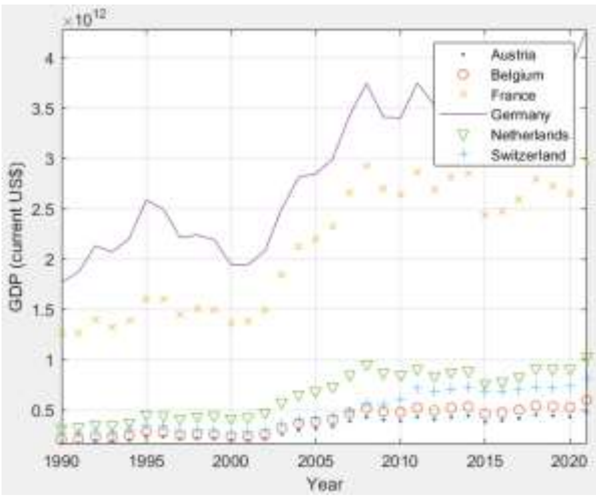


Figure 3. The direction in which a stream flows toward a particular river [48].

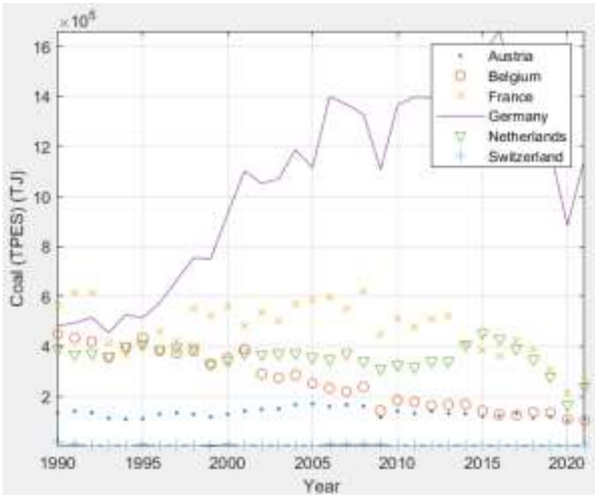
4. Established database

Based on these inputs, neural networks and optimization techniques were employed to predict the nations of Western Europe's carbon dioxide emissions. By considering these variables as inputs, this study aims to identify the relationship between fuel use, economic activity (as shown by GDP), and carbon dioxide emissions. This facilitates the analysis and understanding of the factors affecting greenhouse gas emissions in the countries of Western Europe throughout the selected period. Fuel consumption parameters allow the models to consider each nation's various energy sources—non-renewable and renewable. This acknowledges the need to account for various energy sources in determining carbon dioxide emissions.

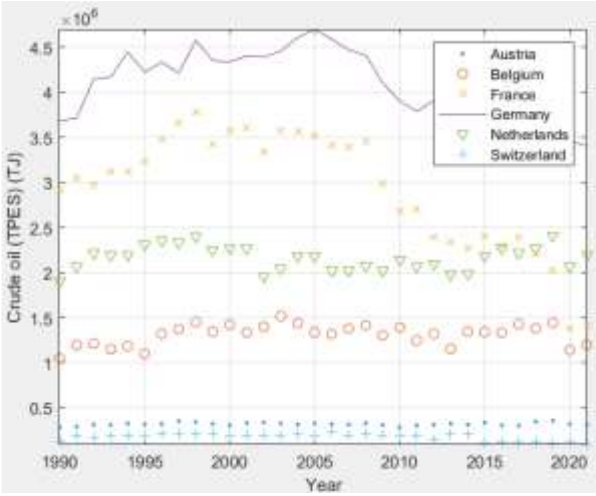
Furthermore, by using GDP as an input variable, the models may incorporate the economic activity of any country. GDP may be used as a stand-in for factors such as energy use, industrial output, and overall economic expansion that affect carbon dioxide emissions. Using fuel consumption and GDP as inputs, the models may examine the complex relationship between energy usage, economic development, and carbon dioxide emissions in Western European countries. It provides a more thorough analysis of the factors influencing emission patterns. Figure 4 shows the intakes and outputs of the Western European nations over several years.



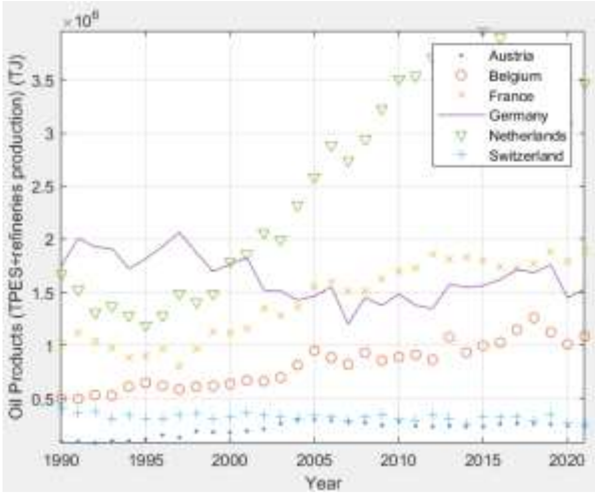
a) GDP (current US$)



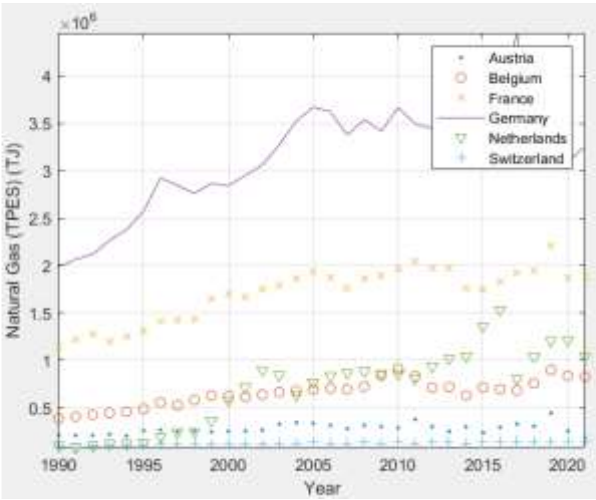
b) Coal (TPES) (TJ)



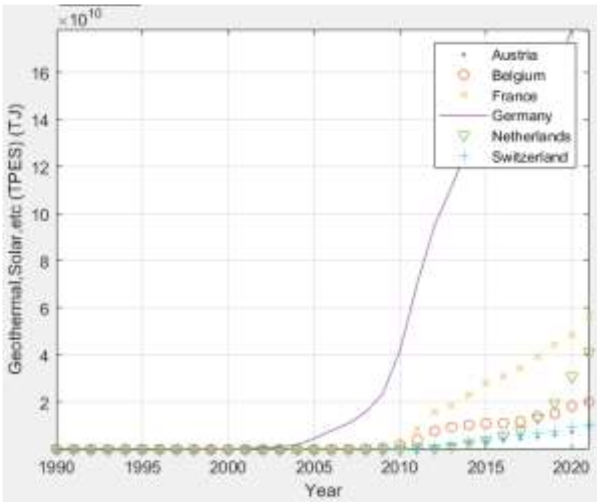
c) Crude oil (TPES) (TJ)



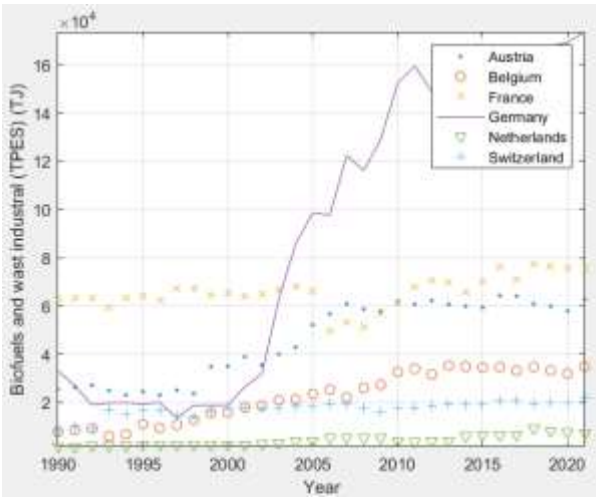
d) Oil products (TPES + refineries production) (TJ)



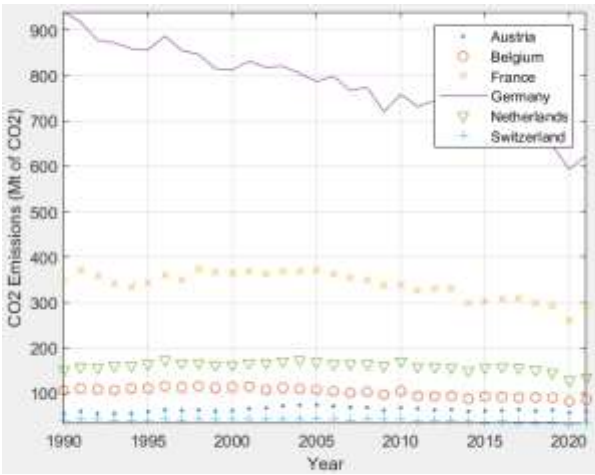
e) Natural gas (TPES) (TJ)



f) Geothermal, Solar, etc (TPES) (TJ)



g) Biofuels and waste industrial (TPES) (TJ)

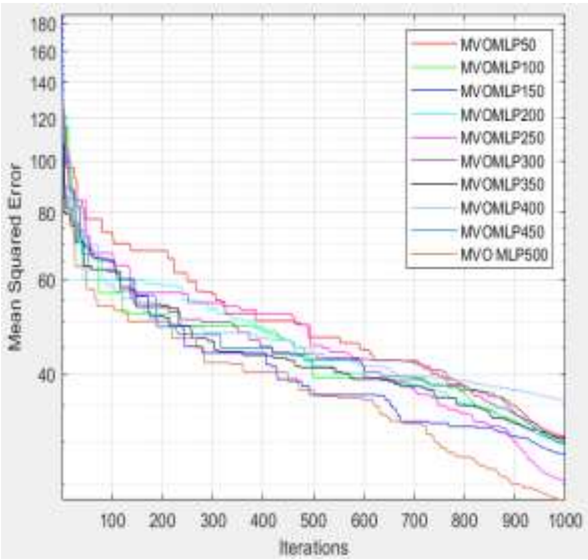


h) CO_2 emissions (Mt of CO_2)

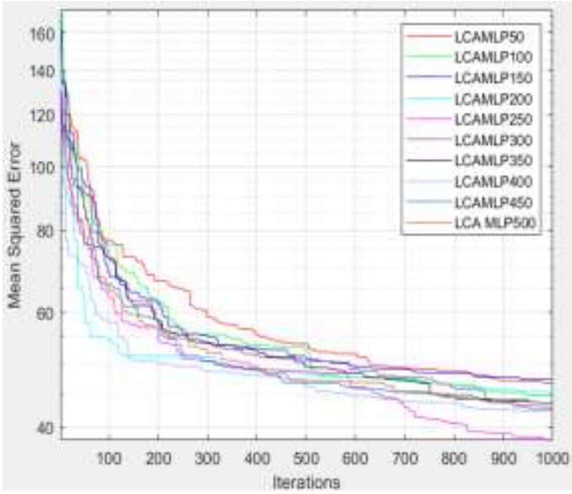
Figure 4: parameters for inputs and outputs.

5. Results and discussion

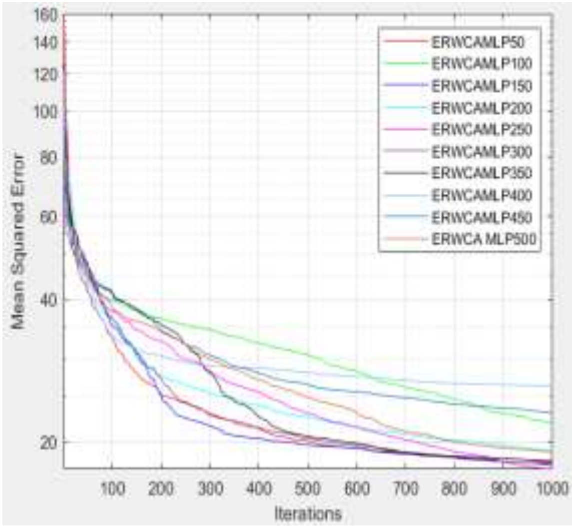
Many networks with different numbers of layers and types of neurons have been constructed to determine the optimal configuration. Modifying the number of layers and neurons in a conventional ANN also affects the models' accuracy. The ideal network was constructed employing a feedforward back-propagation technique and an average of five hidden units based on the RMSE and R^2 metrics. Many optimization methodologies begin with early optimization findings. With the highest score, the model offers the best prediction network. Fascinatingly, the ratings were based on the model's forecast accuracy. For example, a reduced RMSE results in a higher score for the stated model. The R^2 increases with the score. The outcomes of these networks are therefore utilized in the following sections. Figure 5 displays the MSE fluctuations for each strategy. The initial optimization discovery phase will be the foundation for the subsequent optimization tactics. As a result, the outputs of these networks are utilized in the following sections. Predictable accuracy is higher in structures with a reduced MSE. The proposed model's predicted values can be used more precisely to solve regression and classification problems. The MSE variations between carbon dioxide emission prediction system estimations for the combined MVO, LCA, and ERWCA constructions are shown in Figure 5 throughout several iterations. Based on these facts, MVO, LCA, and ERWCA have determined that 500, 250, and 150 (N_{pop}) are the best possibilities.



(a) MVO-MLP



(b) LCA-MLP



(c) ERWCA-MLP

Figure 5. MSE technique variation.

5.1. Statistical Accuracy Assess

A scoring system assigns a number based on an object's or person's performance or qualities. Different ranking strategies may be required for different circumstances and objectives. One well-liked method is the total score rank strategy, which involves adding up each object or person's scores and assigning a score based on their total score. An alternative method—a term not frequently employed—is the color-scoring rank system. It may, however, reference a color-coded rating system that uses levels or categories. In the current study, for example, population sizes have been graded based on their R^2 and RMSE values, with different colors indicating different performance levels. Throughout the grading process, R^2 and RMSE are used to choose the top hybrid designs (Table 1). In the best hybrid approach for carbon dioxide emission, 500 swarm populations are utilized for training and evaluating predictive modeling outputs (i.e., how effectively the algorithm could estimate carbon dioxide emission). It also demonstrates how closely step two's results adhere to phase one's. The network results for the different MVO–MLP, LCA–MLP, and ERWCA–MLP models are shown in Tables 1-3.

Table 1 shows the results of using MLP neural networks in combination with MVO to anticipate CO_2 emissions in Western Europe. The table displays the performance metrics for multiple MVO-MLP model configurations over a range of population sizes, including R^2 and RMSE. Each configuration's training and testing datasets are assessed independently, and respective scores are given for each dataset. Furthermore, each configuration's overall performance is shown by its total score and rank. This table offers a thorough summary of the MVO-MLP models' predictive power for various population sizes, offering important context for understanding how well this method works for CO_2 emission forecasting.

The population size of 500 exhibits the most excellent performance among the population sizes examined in the MVO-MLP setups for estimating CO_2 emissions in Western Europe. With this setup, the testing dataset's MVO-MLP model yields the lowest RMSE of 23.35054 and the most excellent R^2 value of 0.9962. This suggests that the model, trained on a population of 500, exhibits remarkable precision in forecasting CO_2 emissions, accounting for a significant amount of the data's volatility. On the other side, a population of 400 people exhibits the lowest performance. The MVO-MLP model has the lowest R^2 value of 0.9910 and the greatest RMSE of 35.89284 for the testing dataset in this setup. Even though this model is the least successful configuration, it is noteworthy that it has a reasonably good predicted accuracy, indicating the overall effectiveness of the MVO-MLP strategy. In conclusion, among the configurations examined, the MVO-MLP configuration with a population size of 500 proves to be the most successful in forecasting CO_2 emissions, while the configuration with a population size of 400 is the least successful. These results emphasize how crucial population size selection is to maximizing the MVO-MLP model's effectiveness in CO_2 emission prediction.

Table 1. The network results for several MVO-MLP setups.

Population size	Training dataset		Testing dataset		Scoring				Total Score	Rank
	RMSE	R^2	RMSE	R^2	Training		Testing			
50	30.69459	0.9934	40.17346	0.98567	2	2	6	6	16	7
100	29.69849	0.9938	39.67891	0.98602	7	7	7	7	28	3
150	28.42438	0.9944	37.95892	0.98722	8	8	9	9	34	2
200	29.8752	0.9938	38.84735	0.98661	5	5	8	8	26	5
250	25.38279	0.9955	40.97803	0.98508	9	9	5	5	28	3
300	30.17235	0.9936	43.77443	0.98296	4	4	3	3	14	8
350	30.45397	0.9935	47.18794	0.98017	3	3	2	2	10	9
400	35.89284	0.9910	47.72456	0.97971	1	1	1	1	4	10
450	29.73643	0.9938	42.28497	0.98411	6	6	4	4	20	6
500	23.35054	0.9962	33.8724	0.98983	10	10	10	10	40	1

Table 2 shows the best and worst population sizes for predicting CO_2 emissions in Western Europe using the LCA and MLP configurations. The best-performing configuration is observed with a population size of 250, achieving the lowest RMSE of 38.251 and the highest R^2 value of 0.9898 for both the training and testing datasets. This configuration demonstrates exceptional predictive accuracy and robustness, indicating that a moderate population size allows for compelling solution space exploration, leading to superior model performance. Conversely, the worst-performing configuration is associated with a population size of 150, exhibiting the highest RMSE of 47.02609 and the lowest R^2 value of 0.9845 for both the training and testing datasets. Despite a smaller population size, this configuration fails to capture the complexities of the CO_2 emission prediction task adequately. The limited population size may restrict the exploration of potential solutions, resulting in suboptimal model performance and decreased predictive accuracy. Comparing the best and worst population sizes highlights the importance of population size selection in the LCA for MLP neural networks. Optimal performance is achieved with a moderate population size, enabling compelling exploration of the solution space and yielding accurate predictions of CO_2 emissions in Western Europe.

Table 2. The network results for several LCA-MLP setups.

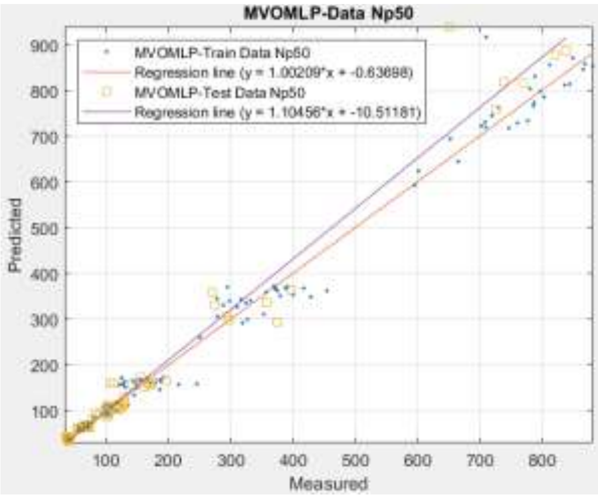
Population size	Training dataset		Testing dataset		Scoring				Total Score	Rank
	RMSE	R^2	RMSE	R^2	Training		Testing			
50	46.62401	0.9848	55.4547	0.97251	2	2	4	4	12	8
100	44.45728	0.9862	58.39934	0.96947	4	4	3	3	14	7
150	47.02609	0.9845	60.23934	0.96748	1	1	1	1	4	10
200	45.23428	0.9857	58.74578	0.9691	3	3	2	2	10	9
250	38.25115	0.9898	48.12728	0.97937	10	10	10	10	40	1
300	42.72072	0.9872	52.4199	0.97547	8	8	9	9	34	2
350	43.68125	0.9866	52.75155	0.97516	5	5	8	8	26	5
400	42.30209	0.9875	53.34104	0.97459	9	9	6	6	30	3
450	43.62747	0.9867	54.80796	0.97316	6	6	5	5	22	6
500	42.84902	0.9871	52.96821	0.97495	7	7	7	7	28	4

Table 3 shows the best and worst population sizes for predicting CO2 emissions in Western Europe using the ERWCA in conjunction with MLP configurations. The best-performing configuration is associated with a population size of 150, achieving the lowest RMSE of 17.99364 and the highest R-squared (R^2) value of 0.9977 for both the training and testing datasets. This configuration demonstrates exceptional predictive accuracy and robustness, indicating that a moderate population size allows for practical solution space exploration, leading to superior model performance. On the other hand, the worst-performing configuration is observed with a population size of 400, exhibiting the highest RMSE of 26.28317 and the lowest R^2 value of 0.9952 for both the training and testing datasets. Despite being a larger population size, this configuration fails to adequately capture the complexities of the CO_2 emission prediction task. The excessive population size may lead to overfitting or inefficient solution space exploration, resulting in suboptimal model performance and decreased predictive accuracy. Comparing the best and worst population sizes highlights the importance of population size selection in the ERWCA for Multi-layer Perceptron (MLP) neural networks. Optimal performance is achieved with a moderate population size, enabling compelling exploration of the solution space and yielding accurate predictions of CO_2 emissions in Western Europe.

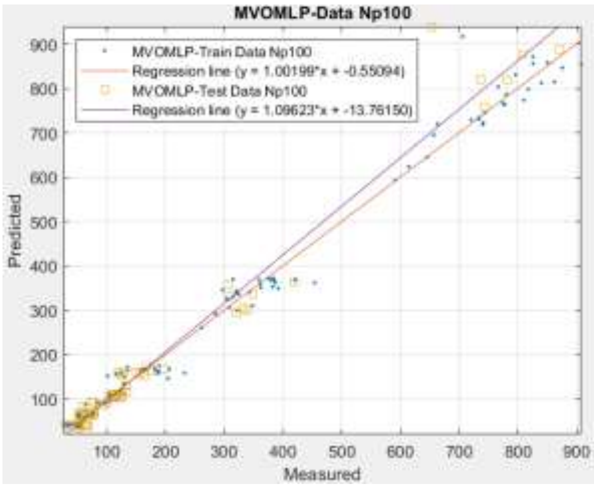
Table 3. The network results for several ERWCA-MLP setups.

Population size	Training dataset		Testing dataset		Scoring				Total Score	Rank
	RMSE	R^2	RMSE	R^2	Training		Testing			
50	18.13245	0.9977	33.58319	0.99001	7	7	4	4	22	7
100	22.01718	0.9966	42.6357	0.98384	3	3	1	1	8	8
150	17.99364	0.9977	30.13947	0.99196	8	8	10	10	36	1
200	19.19307	0.9974	31.53503	0.99119	4	4	9	9	26	4
250	17.52497	0.9979	33.45542	0.99008	10	10	5	5	30	3
300	17.85191	0.9978	32.29247	0.99076	9	9	7	7	32	2
350	18.22919	0.9977	33.26106	0.9902	6	6	6	6	24	6
400	26.28317	0.9952	41.03856	0.98504	1	1	3	3	8	8
450	23.04591	0.9963	42.19242	0.98418	2	2	2	2	8	8
500	19.06265	0.9975	31.55152	0.99118	5	5	8	8	26	4

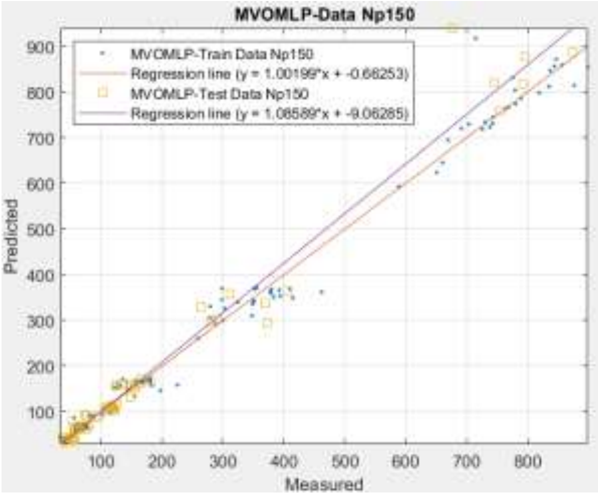
The outcomes of the second stage are derived by contrasting the actual data with the hybrid design's anticipated values. The R^2 is a popular technique for determining which hybrid design is optimal. As previously said, the graph illustrates how a binary classifier system's diagnostic capabilities are affected when the discriminating threshold is changed. The model's ability to distinguish between positive and negative categories improves as R^2 increases. The best-fit structural R^2 plots for the hybrid MVO-MLP, LCA-MLP, and ERWCA-MLP models are displayed in Figure 6-8. The best prediction model (based on the recommended hybrid MVO-MLP, LCA-MLP, and ERWCA-MLP models) was developed for population sizes of 500, 250, and 150 based on the results of the iteration phase.



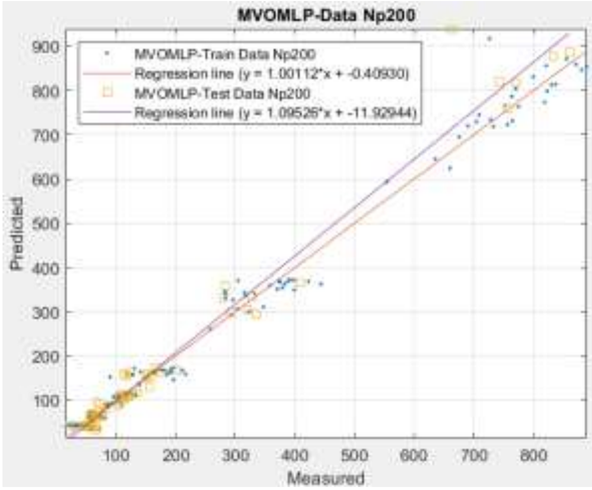
(a) MVOMLP-N_p=50



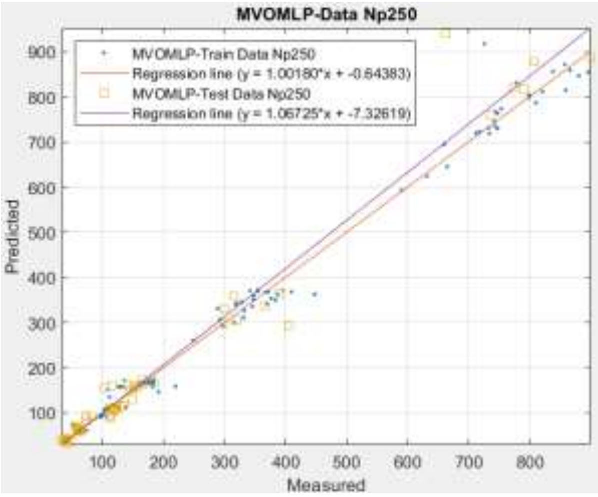
(b) MVOMLP -N_p=100



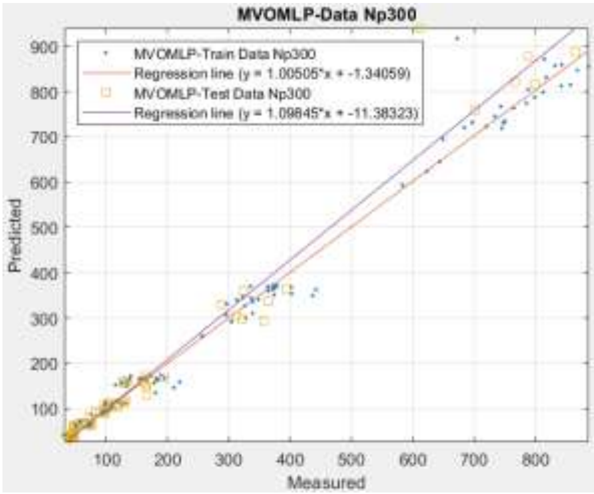
(c) MVOMLP -N_p=150



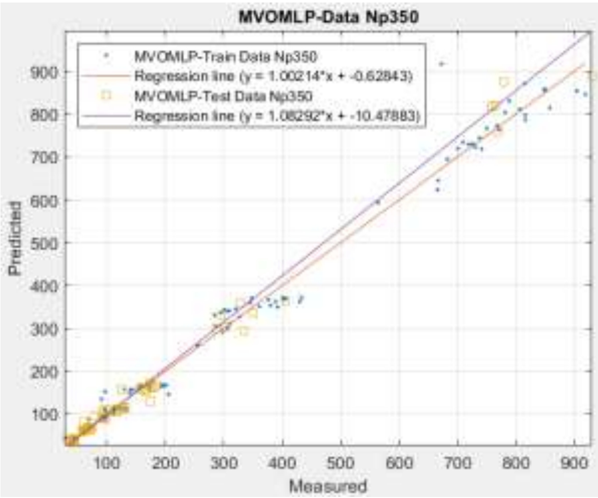
(d) MVOMLP -N_p=200



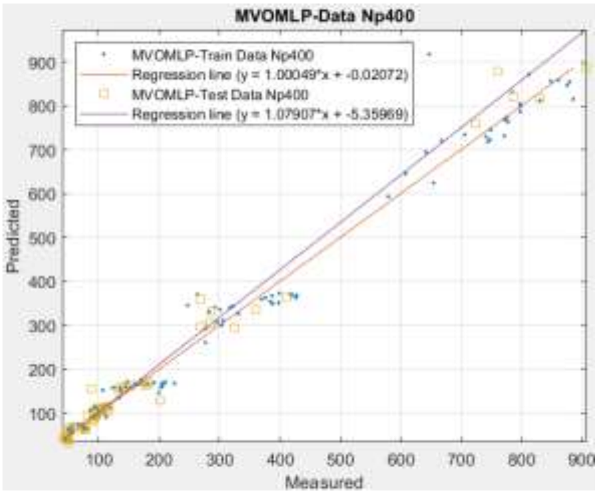
(e) MVOMLP -N_p=250



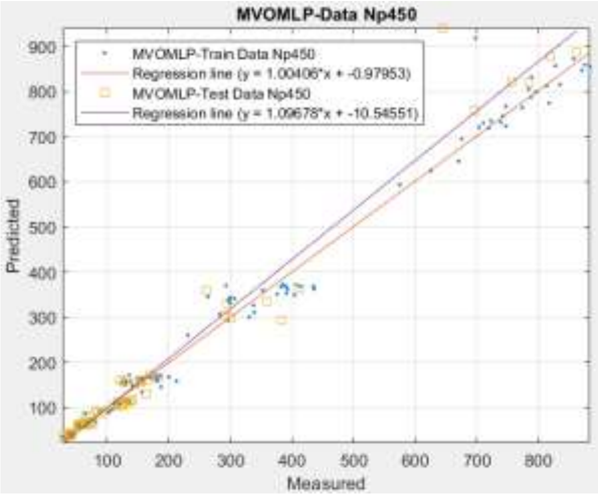
(f) MVOMLP -N_p=300

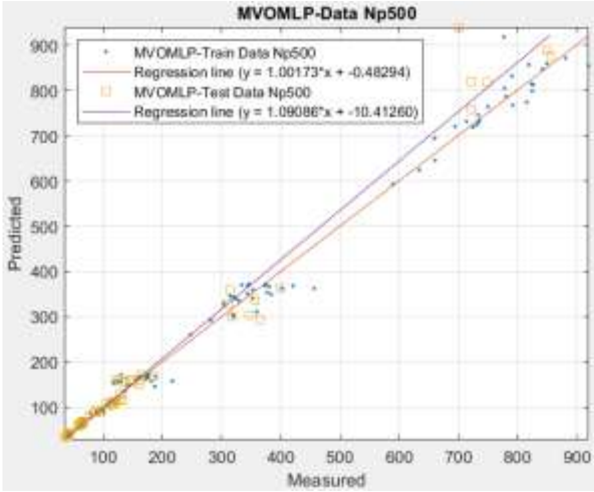


(g) MVOMLP -N_p=350



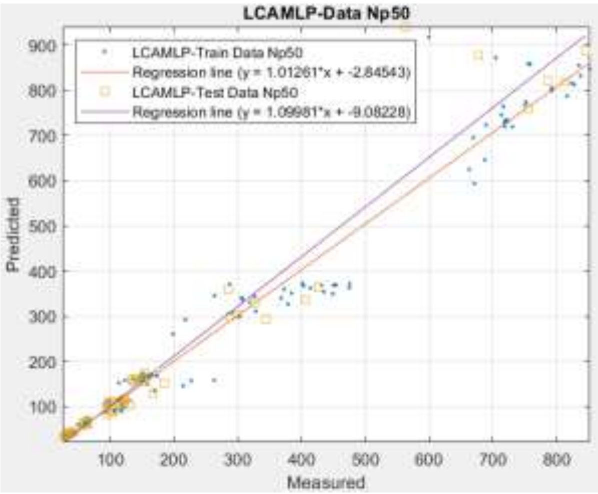
(h) MVOMLP -N_p=400



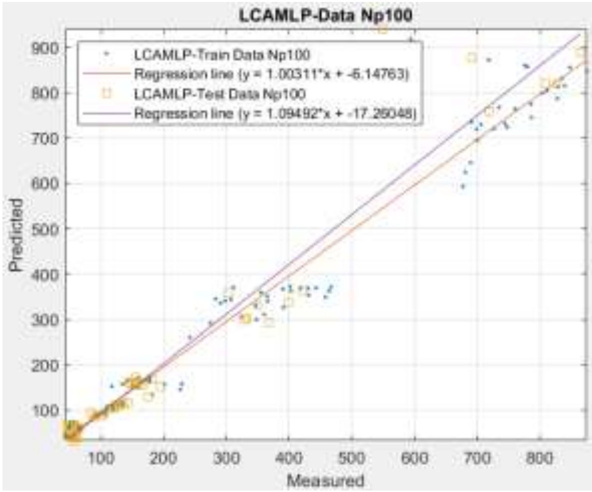


(i) MVOMLP -N_p=450 **(j) MVOMLP -N_p=500**

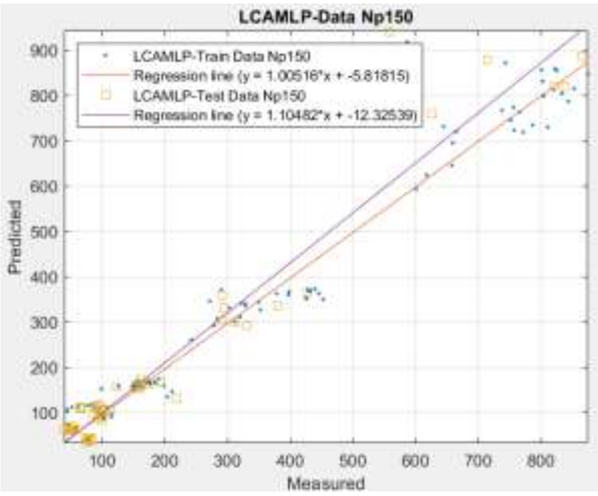
Figure 6: shows the accuracy results for MVO-MLP model-based best-fit architectures.



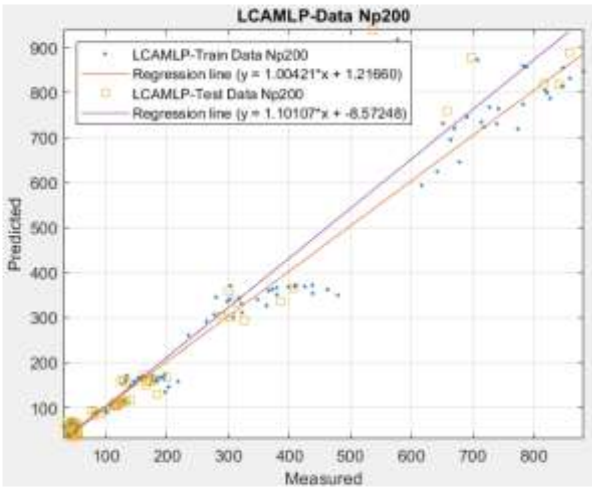
(a) LCAMLP-N_p=50



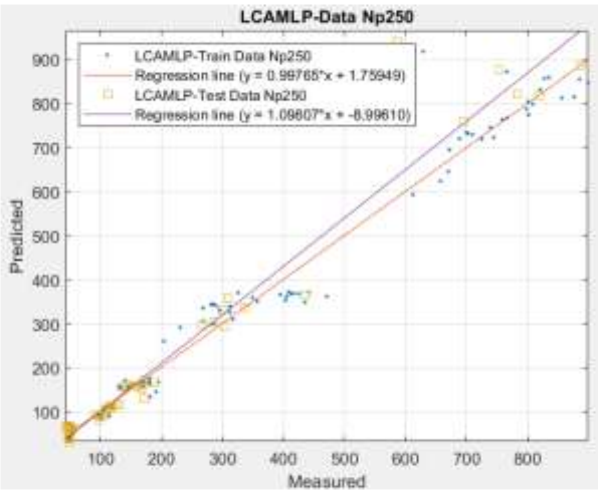
(b) LCAMLP -N_p=100



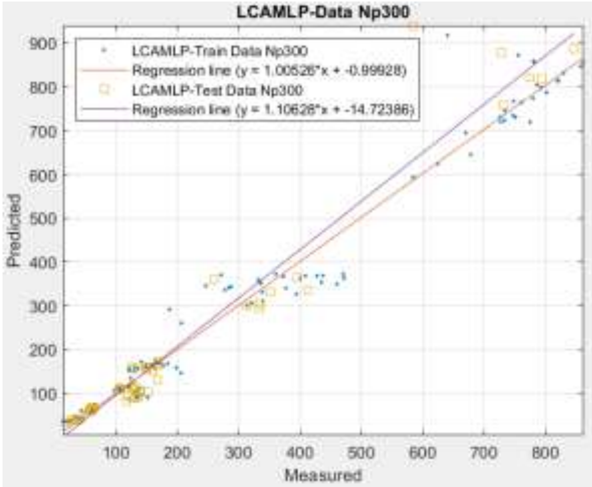
(c) LCAMLP -N_p=150



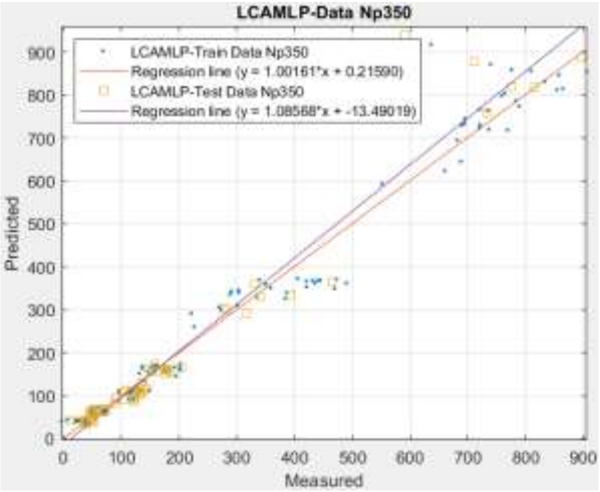
(d) LCAMLP -N_p=200



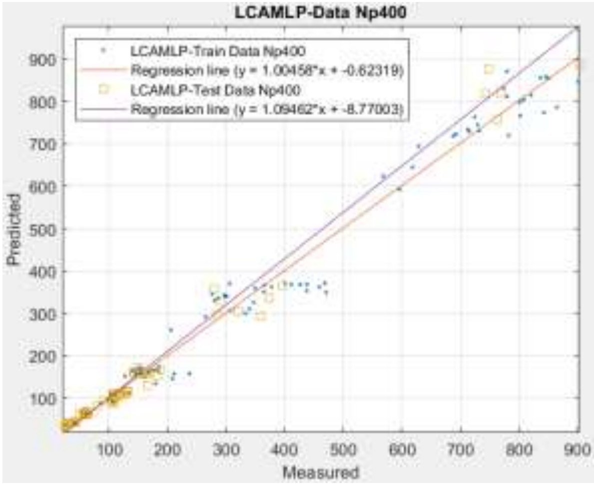
(e) LCAMLP -N_p=250



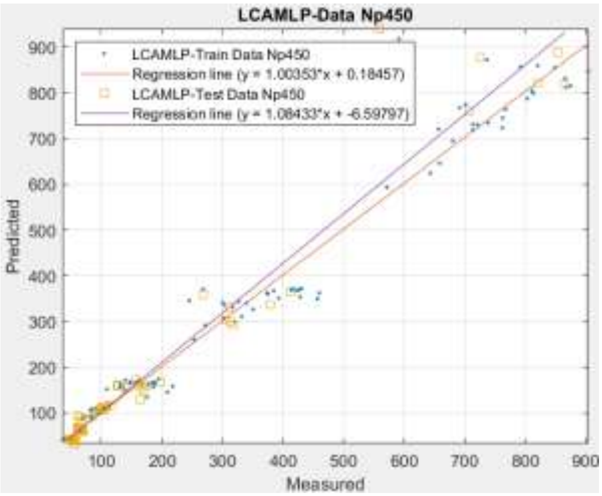
(f) LCAMLP -N_p=300



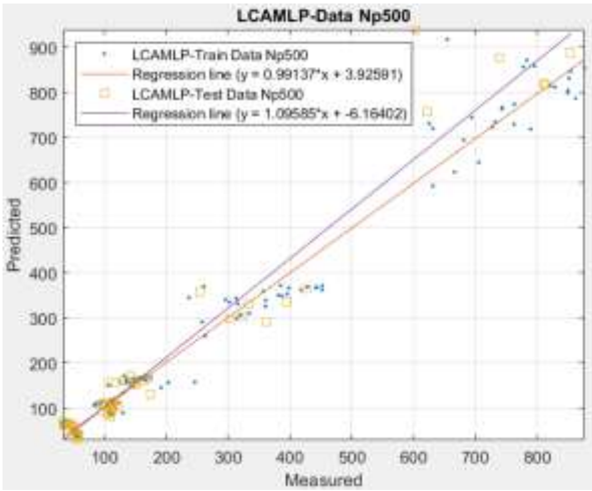
(g) LCAMLP -N_p=350



(h) LCAMLP -N_p=400

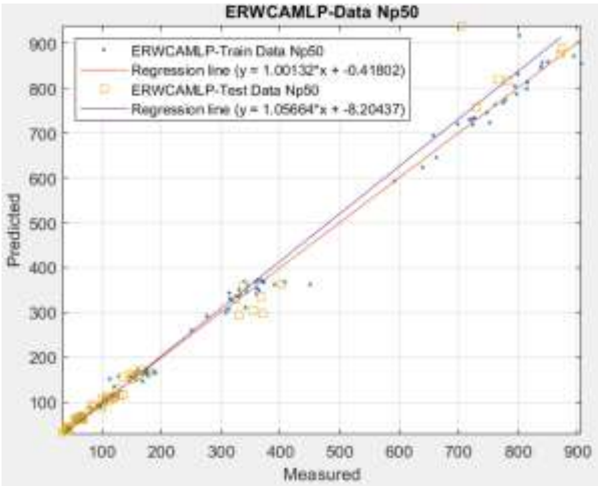


(i) LCAMLP -N_p=450

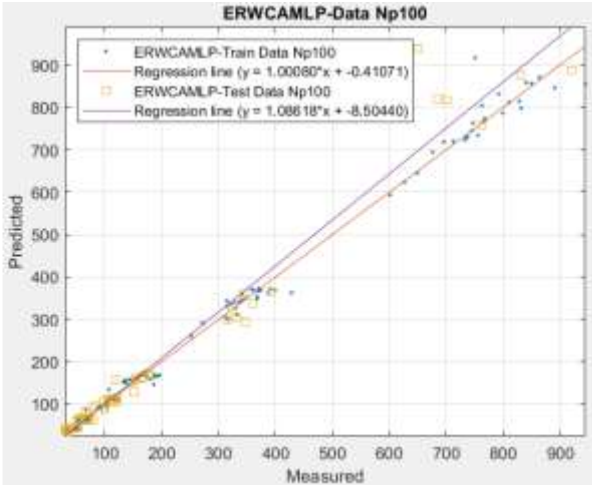


(j) LCAMLP -N_p=500

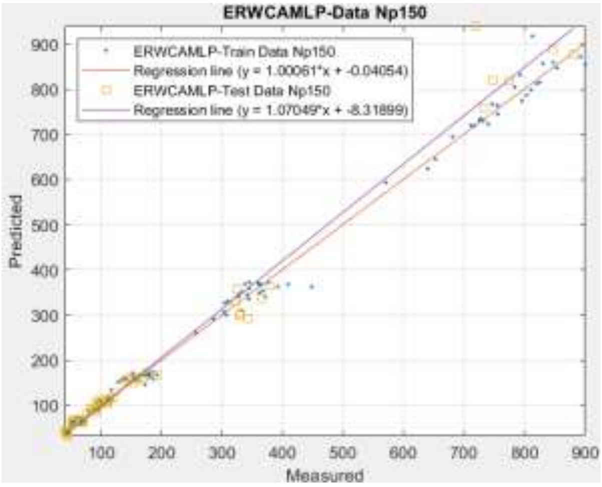
Figure 7. shows the accuracy results for LCA-MLP model-based best-fit architectures.



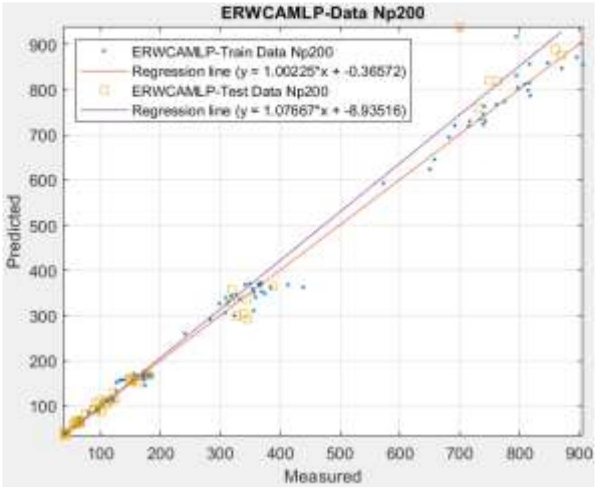
(a) ERWCAMLP -N_p=50



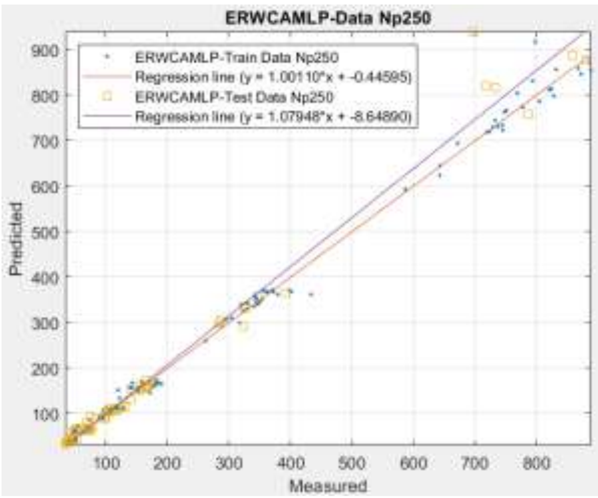
(b) ERWCAMLP -N_p=100



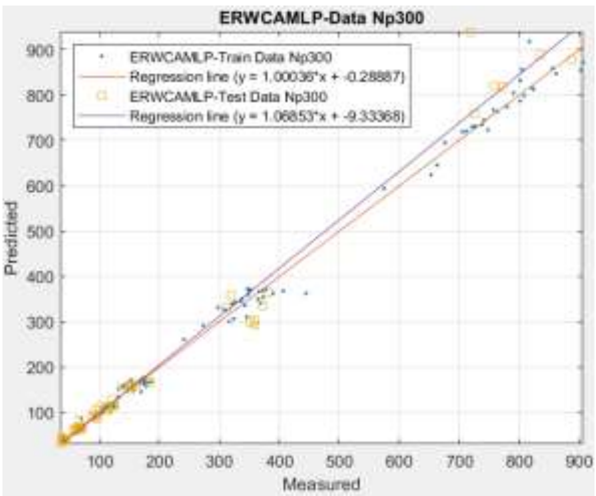
(c) ERWCAMLP -N_p=150



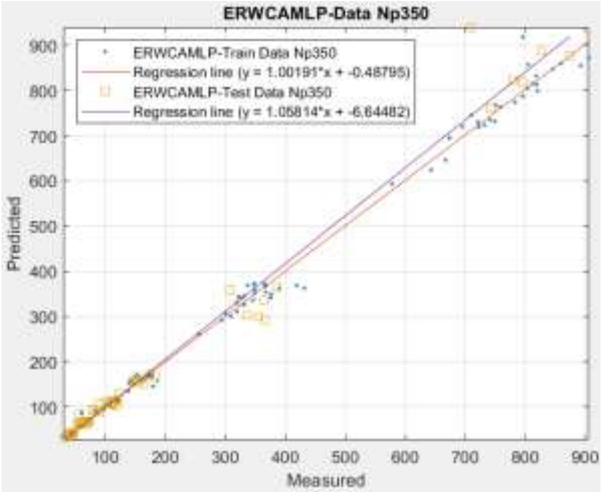
(d) ERWCAMLP -N_p=200



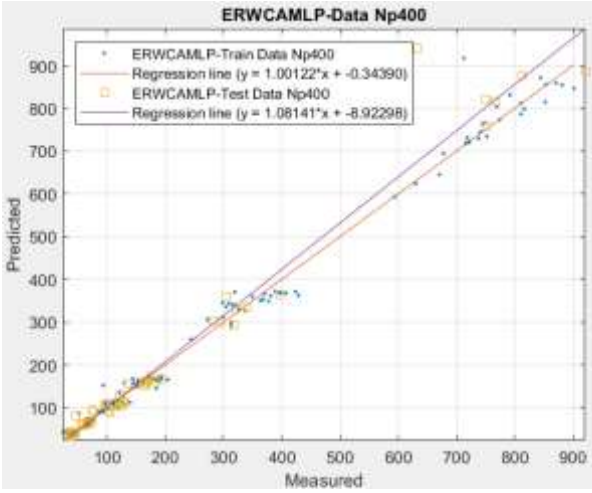
(e) ERWCAMLP -N_p=250



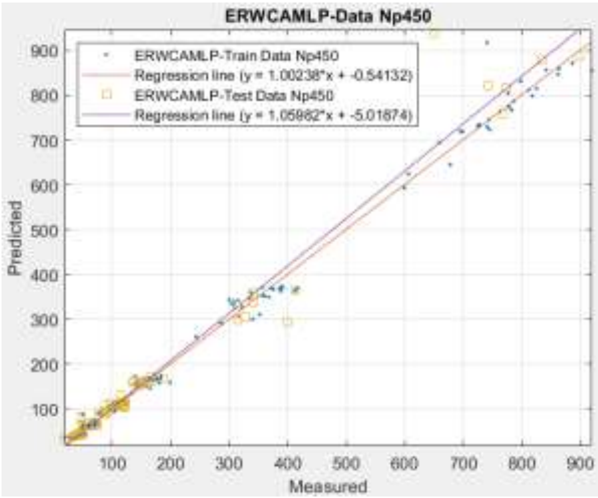
(f) ERWCAMLP -N_p=300



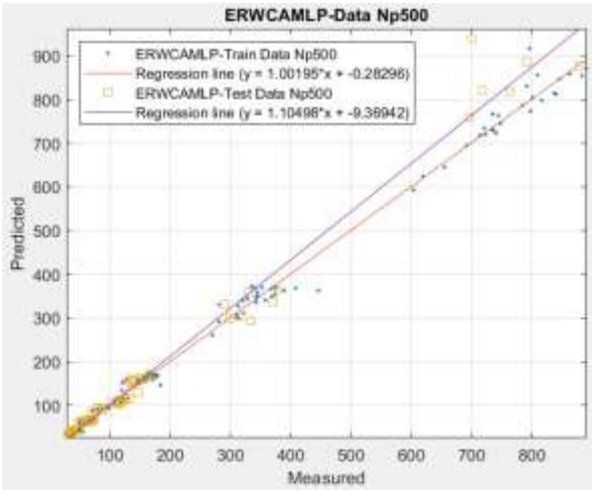
(g) ERWCAMLP -N_p=350



(h) ERWCAMLP -N_p=400



(i) ERWCAMLP -N_p=450

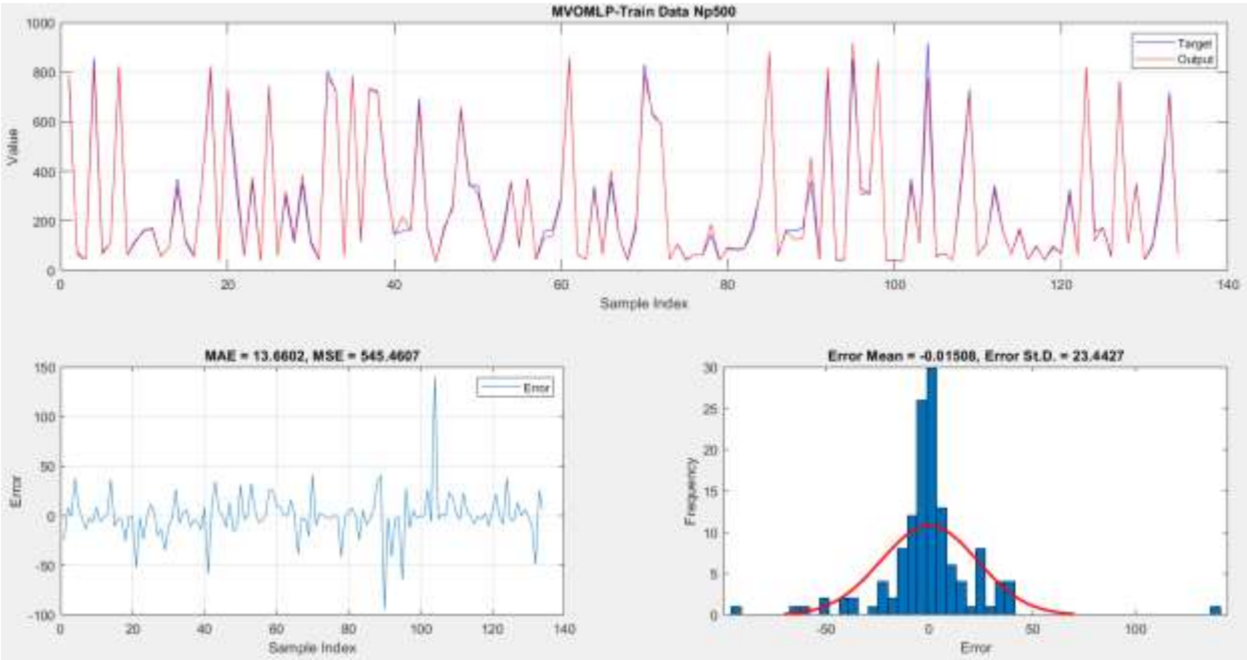


(j) ERWCAMLP -N_p=500

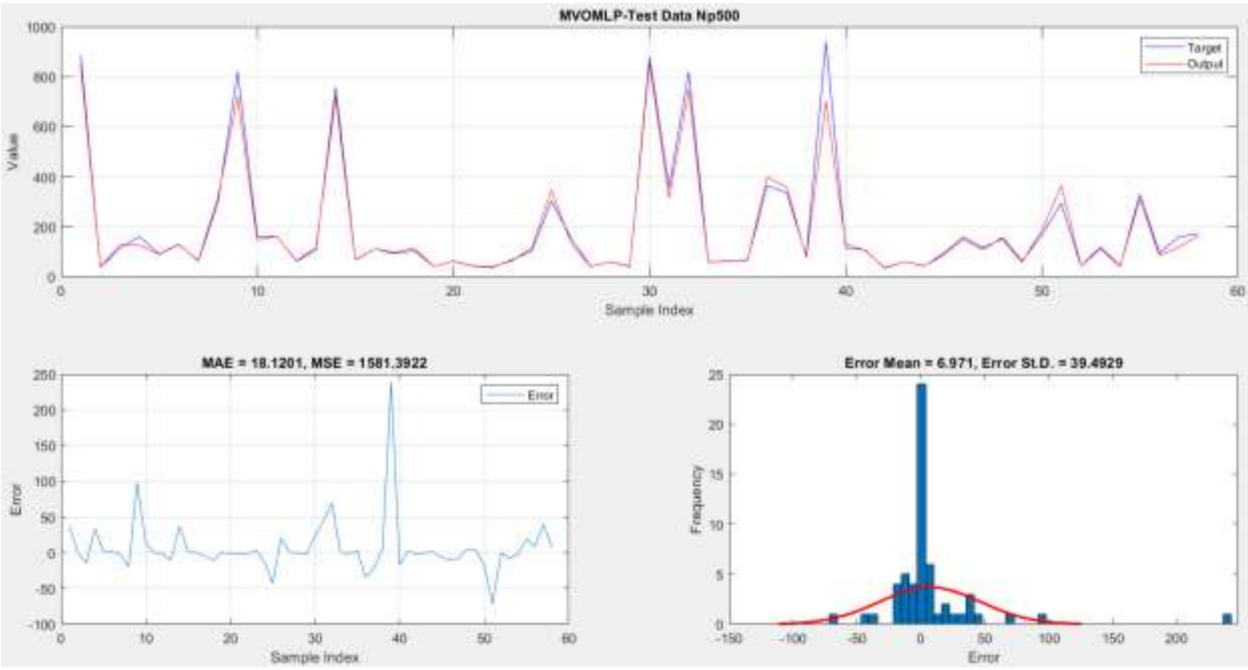
Figure 8. shows the accuracy results for ERWCA-MLP model-based best-fit architectures.

5.2. Error analysis

Figures 9-11 display the frequency in the best-fitted structures for MVO-MLP, LCA-MLP, and ERWCA-MLP. The results from the training and testing datasets show an exceptionally high degree of agreement between the estimated and observed carbon dioxide emission measurements. Based on the findings of the training and testing datasets, the study concludes that there is a very high degree of agreement between the calculated and observed carbon dioxide emission measurements. This suggests that the models, which employ various techniques, such as MVO-MLP, LCA-MLP, and ERWCA-MLP, are helpful for accurately estimating carbon dioxide emissions in the context of Western Europe. The remarkable level of agreement between the estimated and observed data indicates that the models match the underlying dynamics and patterns of carbon dioxide emissions quite well. It suggests that reliable emission estimates may be generated using this work's neural network models and optimization strategies.

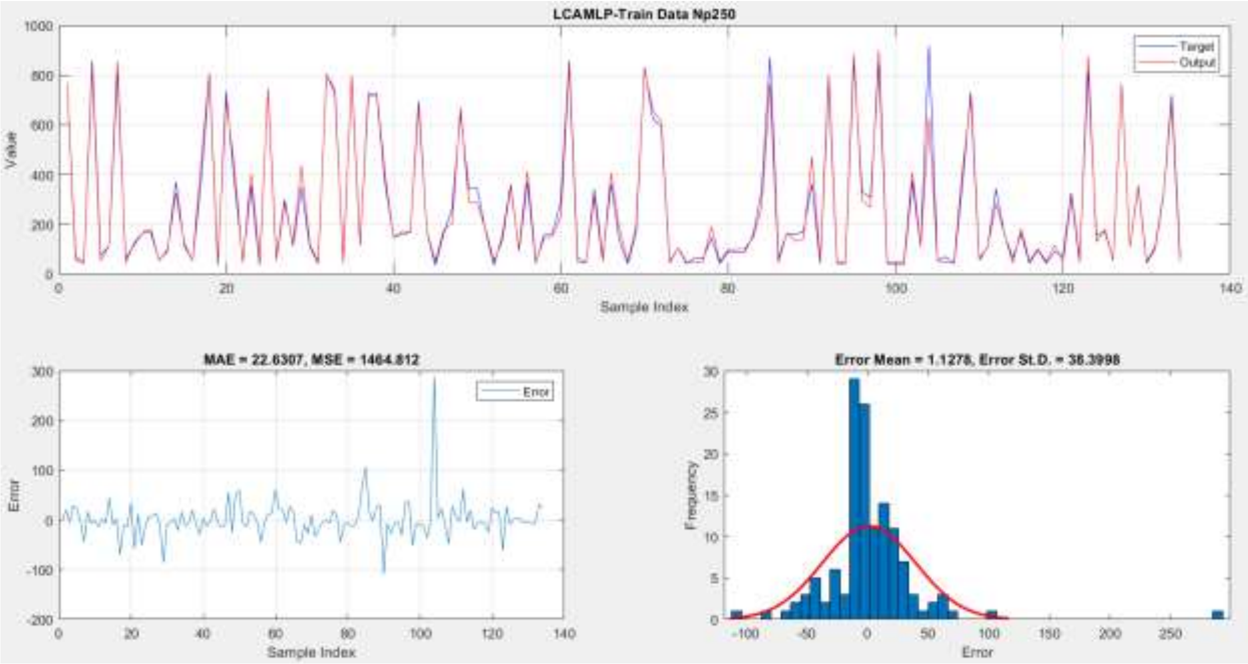


a) Training-500

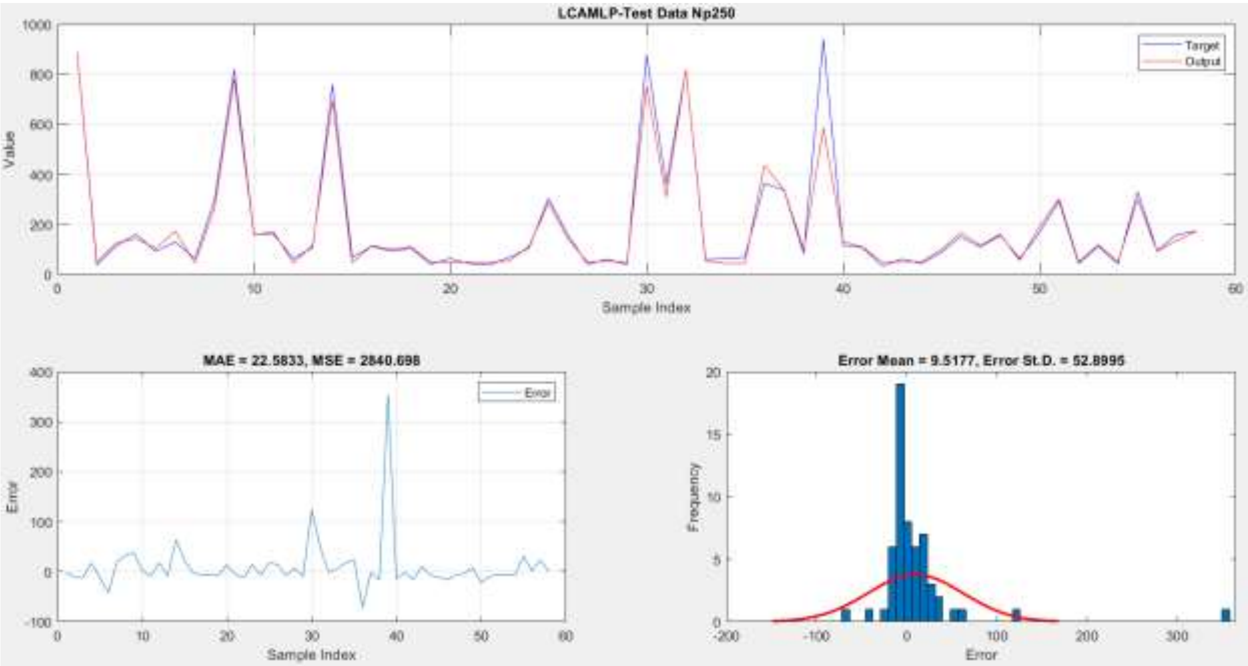


b) Testing-500

Figure 9. The suggested frequency for the MVO-MLP method that works best

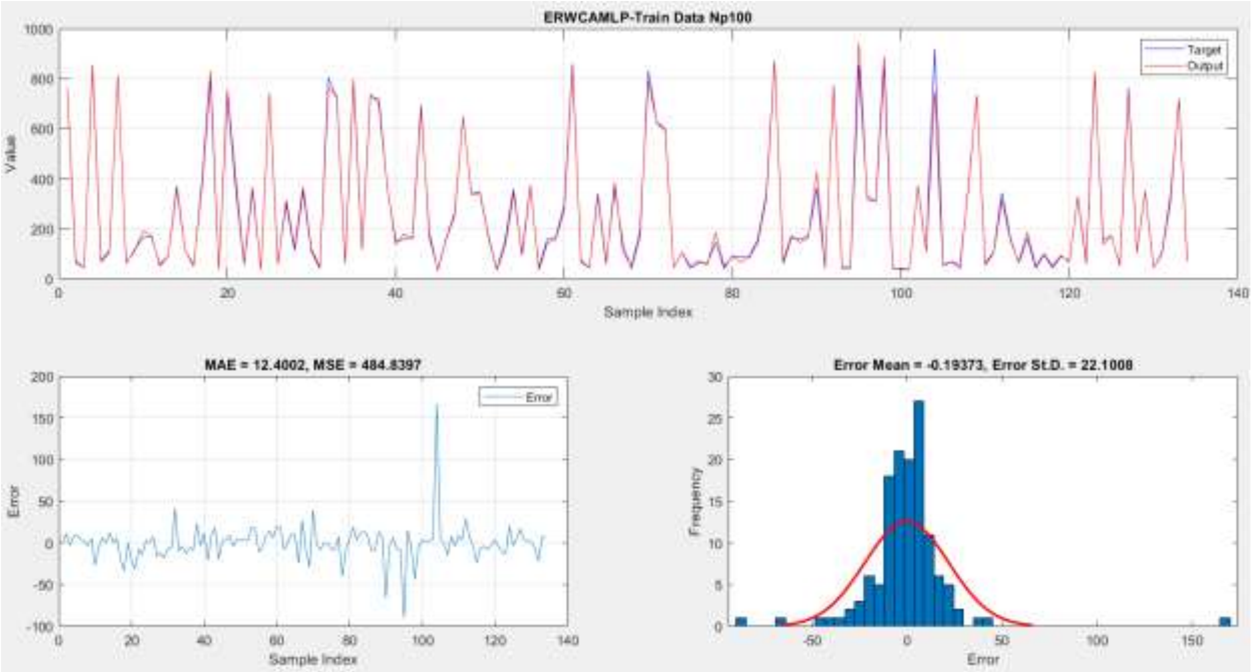


a) Training-250

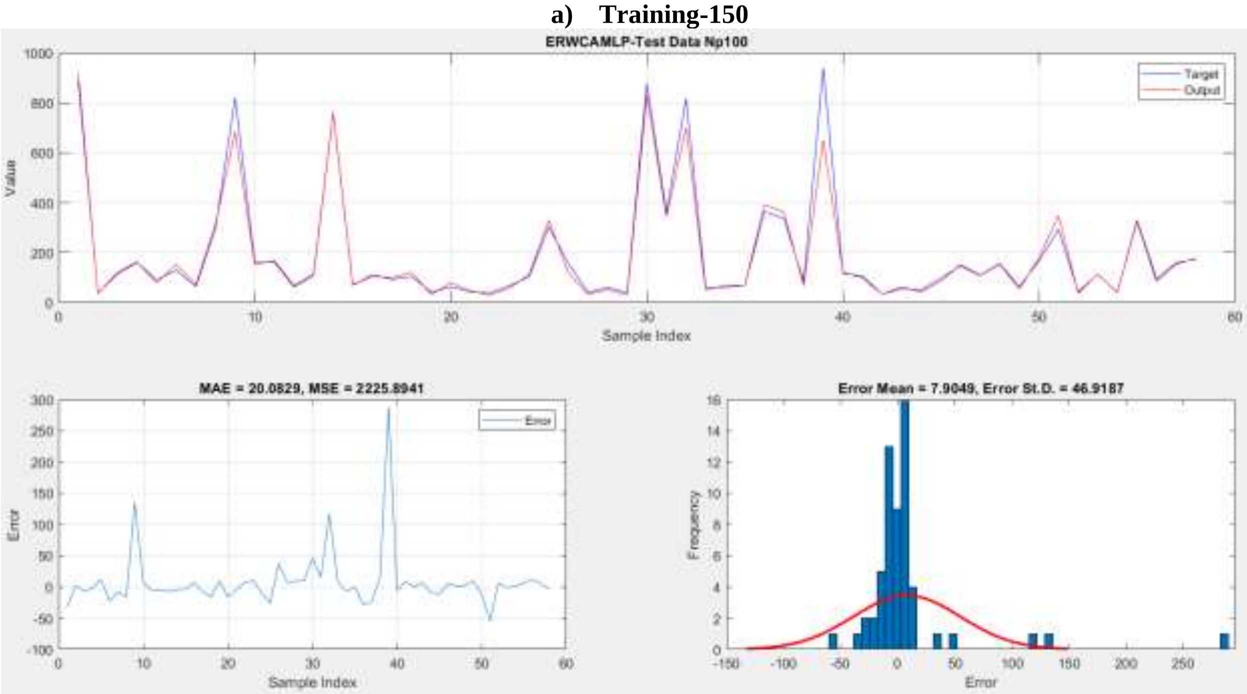


b) Testing-250

Figure 10. The suggested frequency for the LCA-MLP method that works best



a) Training-150



b) Testing-150

Figure 11: The suggested frequency for the ERWCA-MLP method that works best

5.3. Measurement of carbon dioxide emissions by optimization algorithms

Table 4 shows the best and worst population sizes for predicting CO_2 emissions in Western Europe across the three model structures. The best-performing population size is associated with the ERWCAMLP model, utilizing the ERWCA with a population size of 150. This configuration achieves the lowest RMSE of 17.99364 and the highest R^2 value of 0.9977 for both the training and testing datasets. The moderate population size of 150 allows for effective solution space exploration, leading to superior predictive accuracy and robustness. Conversely, the worst-performing population size is observed with the LCAMLP model, utilizing the LCA with a population size of 250. This configuration exhibits the highest RMSE of 38.251 and the lowest R^2 value of 0.9898 for both the training and testing datasets. Despite a larger population size, the

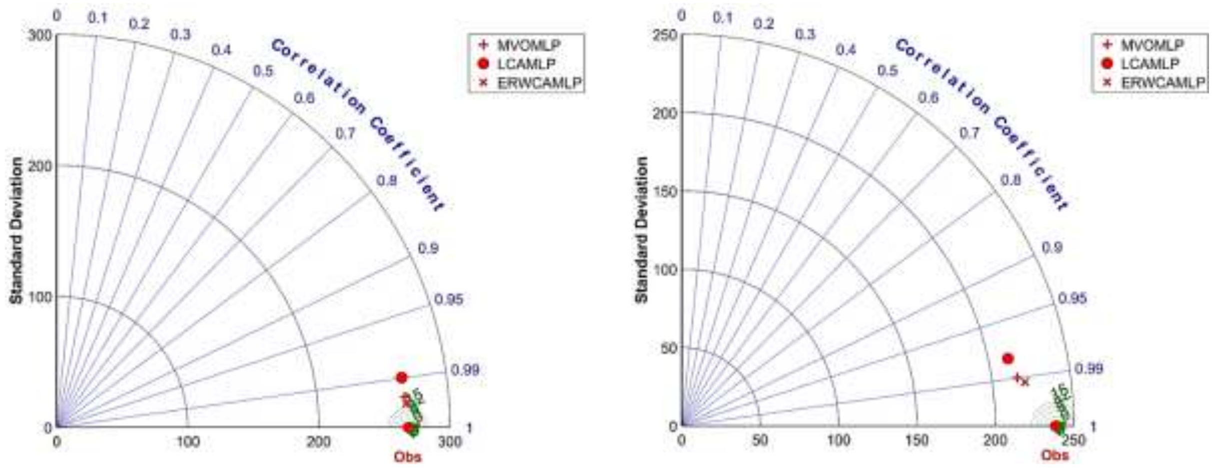
LCAMLP model failed to adequately capture the complexities of the CO_2 emission prediction task, resulting in suboptimal model performance and decreased predictive accuracy. Overall, comparing the best and worst population sizes underscores the importance of population size selection in optimizing predictive model performance for CO_2 emission prediction in Western Europe. Moderate population sizes allow for practical solutions for space exploration and yield accurate predictions. In contrast, excessively large or small population sizes may lead to overfitting or inefficient exploration, resulting in decreased model performance.

Table 4. The MVO-MLP, LCA-MLP, and ERWCA-MLP structures' network results

Proposed models	Swarm size	Training dataset		Testing dataset		Scoring				Total Score	Rank
		RMSE	R^2	RMSE	R^2	Training		Testing			
MVOMLP	500	23.35054	0.9962	33.8724	0.98983	2	2	2	2	8	2
LCAMLP	250	38.25115	0.9898	48.12728	0.97937	1	1	1	1	4	3
ERWCAMLP	150	17.99364	0.9977	30.13947	0.99196	3	3	3	3	12	1

5.4. Taylor Diagrams

A Taylor diagram is a graphical representation commonly used in meteorology and climate science to assess the skill of models or observational datasets relative to a reference dataset. It plots the standard deviation of the model (or prediction) against the correlation coefficient with the reference dataset, with each model represented by a point on the diagram. The closer the point is to the reference dataset, the better the agreement between the model and the observations in terms of both variability and pattern correlation Taylor [50]. In the context of your paper, a Taylor diagram could be used to compare the performance of different predictive models (e.g., MVO-MLP, LCA-MLP, ERWCA-MLP in terms of their ability to capture the variability and pattern correlation of energy-related CO_2 emissions compared to observed data. This visualization could provide a comprehensive assessment of model skills and help identify the most reliable and accurate model for predicting CO_2 emissions in Western Europe. The pattern correlation coefficients for the MVO-MLP, LCA-MLP, and ERWCA-MLP are 0.999.



(a) Training **(b) Testing**

Figure 12. Taylor Diagram for the CO_2 emision

5.5. Discussion

Interpreting the results in the context of the research objectives provides valuable insights into the effectiveness of different optimization algorithms and population sizes for predicting CO_2 emissions in Western Europe using neural network models. The results demonstrate that the choice of optimization algorithm significantly impacts the predictive performance of the neural network models. The ERWCA-MLP model, utilizing the Evaporation Rate Water Cycle Algorithm (ERWCA), outperforms the MVOMLP and LCAMLP models regarding both RMSE and R-squared (R^2) values for training and testing datasets. This suggests that ERWCA is more effective in optimizing the neural network parameters for accurate CO_2 emission prediction in Western Europe than Multiverse Optimization (MVO) and LCA. The analysis reveals that the population size plays a crucial role in determining the predictive accuracy of the neural network models. The ERWCAMLP model with a population size of 150 achieves the best overall performance, indicating that a moderate population size allows for practical solution space exploration and yields accurate predictions. Conversely, the LCAMLP model with a population size of 250 exhibits the worst performance, suggesting that substantial population sizes may lead to suboptimal model performance. The comparison between the MVO-MLP, LCAMLP, and ERWCAMLP models highlights the superiority of nature-inspired optimization algorithms, particularly ERWCA, over traditional methods. The ERWCAMLP model achieves the highest rank and total score, indicating its effectiveness in predicting CO_2 emissions in Western Europe compared to MVO and LCA. This underscores the importance of leveraging advanced optimization techniques to enhance the accuracy and reliability of predictive models for environmental applications. Overall, interpreting the results aligns with the research objectives by providing valuable insights into the effectiveness of different optimization algorithms and population sizes in predicting CO_2 emissions in Western Europe. The findings contribute to advancing predictive modeling techniques for environmental science and provide practical guidance for policymakers and stakeholders in addressing climate change and promoting sustainable energy practices.

Integrating neural networks with optimization algorithms leads to enhanced predictive accuracy compared to traditional methods. The results demonstrate that the combined models, such as ERWCAMLP, achieve lower Root Mean Square Error (RMSE) and higher R-squared (R^2) values for both training and testing datasets, indicating improved model performance in capturing the underlying patterns and trends in CO_2 emissions data. The combined models show improved generalization capabilities, as evidenced by their consistent performance on training and testing datasets. This suggests that integrating optimization algorithms helps mitigate overfitting and improves the model's ability to generalize to unseen data, resulting in more reliable predictions of CO_2 emissions in Western Europe.

Optimization algorithms play a crucial role in fine-tuning the parameters of neural networks to optimize their performance. By exploring the solution space and iteratively adjusting the network weights and biases, optimization algorithms facilitate the convergence of the model to an optimal solution, leading to improved predictive accuracy and robustness.

The flexibility and adaptability of neural networks allow them to learn complex patterns and relationships from data, while optimization algorithms provide efficient methods for training and optimizing the model parameters. This combination enables the development of flexible and adaptable predictive models that capture the nonlinear dynamics and uncertainties inherent in CO_2 emission prediction. While the results demonstrate the effectiveness of combining neural networks with optimization algorithms, there is potential for further exploration and optimization. Future research could investigate additional optimization algorithms, hybrid approaches, and ensemble techniques to enhance predictive accuracy and robustness for CO_2 emission prediction and other environmental applications.

Integrating neural networks with optimization algorithms offers a promising approach for improving the accuracy and reliability of CO_2 emission prediction models. By leveraging the complementary strengths of both techniques, researchers can develop more effective predictive models that contribute to addressing climate change and promoting environmental sustainability. The findings suggest that integrating neural networks with optimization algorithms, particularly nature-inspired algorithms like the ERWCA can significantly improve the accuracy of CO_2 emission prediction models. Policymakers can leverage these advanced predictive models to understand current emission trends better, forecast future emissions trajectories, and identify key drivers of CO_2 emissions in Western Europe. This information can inform the development of evidence-based policies and regulations to reduce emissions and transition to more sustainable energy sources.

The enhanced predictive accuracy of the combined models enables policymakers to identify specific sectors, regions, and activities that contribute most to CO_2 emissions. By targeting these high-emission areas with tailored mitigation strategies, such as incentives for renewable energy adoption, energy efficiency improvements, and carbon pricing mechanisms, policymakers can maximize the effectiveness of their interventions and accelerate progress toward emission reduction targets. Accurate CO_2 emission predictions facilitate informed decision-making regarding resource allocation and investment in low-carbon technologies and infrastructure. By anticipating future emissions trends and identifying areas with the most significant potential for emission reduction, policymakers and investors can prioritize investments in clean energy, sustainable transportation, and climate-resilient infrastructure, thereby driving economic growth while reducing greenhouse gas emissions.

Developing advanced predictive models for CO_2 emission prediction enables ongoing monitoring and evaluation of the effectiveness of climate policies and mitigation measures. By regularly updating and refining the predictive models based on new data and insights, policymakers can track progress toward emission reduction goals, assess the impact of policy interventions, and make timely adjustments to ensure alignment with long-term climate objectives.

The findings underscore the importance of international collaboration in addressing climate change and reducing CO_2 emissions. By sharing best practices, data, and predictive modeling techniques, countries in Western Europe and beyond can collaborate to develop more accurate and robust emission prediction models, harmonize climate policies, and achieve collective emission reduction targets outlined in international agreements such as the Paris Agreement. The findings from integrating neural networks with optimization algorithms for CO_2 emission prediction have far-reaching implications for environmental policy and decision-making. By harnessing the power of advanced predictive modeling techniques, policymakers can develop more effective strategies for mitigating climate change, fostering sustainable development, and safeguarding the health and well-being of current and future generations.

6. Conclusions

The study explores the integration of neural networks with three nature-inspired optimization algorithms—MVO, LCA, and ERWCA—for predicting CO_2 emissions in Western Europe. The results indicate that ERWCA outperforms MVO and LCA in terms of predictive accuracy, achieving the lowest RMSE and highest R^2 values for training and testing datasets. This suggests ERWCA is more effective in optimizing neural network parameters for CO_2 emission prediction.

- Moderate population sizes, such as 150 for ERWCA, lead to superior model performance compared to excessively large or small population sizes. ERWCA-MLP models with a population size of 150 consistently exhibit the best performance, indicating the importance of population size selection in optimizing predictive model performance.
- The study compares the performance of the integrated models MVOMLP, LCAMLP, and ERWCAMLP) with traditional methods. ERWCAMLP achieves the highest rank and total score, indicating its superiority in predicting CO_2 emissions in Western Europe compared to MVO and LCA.
- The findings have significant implications for environmental policy and decision-making. Accurate CO_2 emission prediction models can inform policy formulation, targeted mitigation strategies, resource allocation, and international collaboration efforts to address climate change and promote environmental sustainability.

Overall, the study highlights the effectiveness of integrating neural networks with optimization algorithms for CO_2 emission prediction and underscores the importance of selecting appropriate optimization techniques and population sizes for optimizing model performance.

Reflecting on the significance of our research contributions and envisioning potential avenues for future exploration, our study represents a vital step forward in advancing predictive modeling techniques for CO_2 emission prediction. By investigating the integration of neural networks with nature-inspired optimization algorithms, such as MVO, LCA, and ERWCA, we shed light on the effectiveness of different optimization techniques and population sizes in optimizing predictive model performance. Our findings offer practical implications for informing environmental policy and decision-making and pave the way for future research endeavors. Moving forward, potential avenues for exploration include the development of hybrid models, ensemble techniques, and spatially explicit modeling approaches to improve predictive accuracy and capture dynamic emission trends. Moreover, there is a need for uncertainty analysis, sensitivity testing, and interdisciplinary collaboration to address methodological challenges and foster innovation in the quest to mitigate climate change and promote environmental sustainability. Through collaborative efforts and interdisciplinary approaches, we can continue to push the boundaries of knowledge, develop holistic solutions, and make meaningful strides toward a more sustainable future.

7. References

[1] Yoro, Kelvin O, Michael O Daramola, CO2 emission sources, greenhouse gases, and the global warming effect, Advances in carbon capture, Elsevier, 2020, pp. 3-28.

[2] Chang, Lei, Muhammad Mohsin, Amir Hasnaoui, Farhad Taghizadeh-Hesary, Exploring carbon dioxide emissions forecasting in China: A policy-oriented perspective using projection pursuit regression and machine learning models, Technological Forecasting and Social Change, 197 (2023) 122872,

[3] Thacker, Scott, Daniel Adshead, Marianne Fay, Stéphane Hallegatte, Mark Harvey, Hendrik Meller, Nicholas O'Regan, Julie Rozenberg, Graham Watkins, Jim W Hall, Infrastructure for sustainable development, Nature Sustainability, 2 (2019) 324-331,

[4] Ofodile, Onyeka Chrisanctus, Adedoyin Tolulope Oyewole, Chinonye Esther Ugochukwu, Wilhelmina Afua Addy, Omotayo Bukola Adeoye, Chinwe Chinazo Okoye, Predictive analytics in climate finance: Assessing risks and opportunities for investors, GSC Advanced Research and Reviews, 18 (2024) 423-433,

[5] Van Fan, Yee, Simon Perry, Jiří Jaromír Klemeš, Chew Tin Lee, A review on air emissions assessment: Transportation, Journal of cleaner production, 194 (2018) 673-684,

[6] Nuţă, Florian Marcel, Alina Cristina Nuţă, Cristina Gabriela Zamfir, Stefan-Mihai Petrea, Dan Munteanu, Dragos Sebastian Cristea, National carbon accounting—analyzing the impact of urbanization and energy-related factors upon CO2 emissions in central–eastern European countries by using machine learning algorithms and panel data analysis, Energies, 14 (2021) 2775,

[7] Hermwille, Lukas, Wolfgang Obergassel, Hermann E Ott, Christiane Beuermann, UNFCCC before and after Paris–what's necessary for an effective climate regime?, Climate policy, 17 (2017) 150-170,

[8] Rane, Nitin, Roles and challenges of ChatGPT and similar generative artificial intelligence for achieving the sustainable development goals (SDGs), Available at SSRN 4603244, (2023),

[9] Emmert-Streib, Frank, Zhen Yang, Han Feng, Shailesh Tripathi, Matthias Dehmer, An introductory review of deep learning for prediction models with big data, Frontiers in Artificial Intelligence, 3 (2020) 4,

[10] Movassagh, Ali Akbar, Jafar A Alzubi, Mehdi Gheisari, Mohamadtaghi Rahimi, Senthilkumar Mohan, Aaqif Afzaal Abbasi, Narjes Nabipour, Artificial neural networks training algorithm integrating invasive weed optimization with differential evolutionary model, Journal of Ambient Intelligence and Humanized Computing, (2023) 1-9,

[11] Abdolrasol, Maher GM, SM Suhail Hussain, Taha Selim Ustun, Mahidur R Sarker, Mahammad A Hannan, Ramizi Mohamed, Jamal Abd Ali, Saad Mekhilef, Abdalrhman Milad, Artificial neural networks based optimization techniques: A review, Electronics, 10 (2021) 2689,

[12] Hamdia, Khader M, Xiaoying Zhuang, Timon Rabczuk, An efficient optimization approach for designing machine learning models based on genetic algorithm, Neural Computing and Applications, 33 (2021) 1923-1933,

[13] Tino, Peter, Lubica Benuskova, Alessandro Sperduti, Artificial neural network models, Springer Handbook of Computational Intelligence, (2015) 455-471,

[14] Fan, Zhenjiang, Kate F Kernan, Aditya Sriram, Panayiotis V Benos, Scott W Canna, Joseph A Carcillo, Soyeon Kim, Hyun Jung Park, Deep neural networks with knockoff features identify nonlinear causal relations and estimate effect sizes in complex biological systems, GigaScience, 12 (2023) giad044,

[15] Taye, Mohammad Mustafa, Understanding of machine learning with deep learning: architectures, workflow, applications and future directions, Computers, 12 (2023) 91,

[16] Mehmood, Faisal, Shabir Ahmad, Taeg Keun Whangbo, An efficient optimization technique for training deep neural networks, Mathematics, 11 (2023) 1360,

[17] Darwish, Ashraf, Dalia Ezzat, Aboul Ella Hassanien, An optimized model based on convolutional neural networks and orthogonal learning particle swarm optimization algorithm for plant diseases diagnosis, Swarm and evolutionary computation, 52 (2020) 100616,

[18] Gupta, Tarun Kumar, Khalid Raza, Optimization of ANN architecture: a review on nature-inspired techniques, Machine learning in bio-signal analysis and diagnostic imaging, (2019) 159-182,

[19] Zhao, Jingyuan, Xuebing Han, Minggao Ouyang, Andrew F Burke, Specialized deep neural networks for

battery health prognostics: Opportunities and challenges, Journal of Energy Chemistry, (2023),

[20] Jahin, Md Abrar, Md Sakib Hossain Shovon, Md Saiful Islam, Jungpil Shin, MF Mridha, Yuichi Okuyama, QAmplifyNet: pushing the boundaries of supply chain backorder prediction using interpretable hybrid quantum-classical neural network, Scientific Reports, 13 (2023) 18246,

[21] Zhang, Hemeng, Pengcheng Wang, Mohammad Rahimi, Hung Vo Thanh, Yongjun Wang, Zhenxue Dai, Qian Zheng, Yong Cao, Catalyzing net-zero carbon strategies: Enhancing CO2 flux Prediction from underground coal fires using optimized machine learning models, Journal of Cleaner Production, 441 (2024) 141043,

[22] Zhang, Yi, Xiaohan Shi, Hengxu Zhang, Yongji Cao, Vladimir Terzija, Review on deep learning applications in frequency analysis and control of modern power system, International Journal of Electrical Power & Energy Systems, 136 (2022) 107744,

[23] Kaack, Lynn H, Priya L Donti, Emma Strubell, George Kamiya, Felix Creutzig, David Rolnick, Aligning artificial intelligence with climate change mitigation, Nature Climate Change, 12 (2022) 518-527,

[24] Zheng, Xiaosong, Dalia Streimikiene, Tomas Balezentis, Abbas Mardani, Fausto Cavallaro, Huchang Liao, A review of greenhouse gas emission profiles, dynamics, and climate change mitigation efforts across the key climate change players, Journal of Cleaner Production, 234 (2019) 1113-1133,

[25] Wang, Di, Wei He, Ruyi Shi, How to achieve the dual-control targets of China's CO2 emission reduction in 2030? Future trends and prospective decomposition, Journal of cleaner production, 213 (2019) 1251-1263,

[26] Yu, Yang, Yu Du, Impact of technological innovation on CO2 emissions and emissions trend prediction on 'New Normal'economy in China, Atmospheric Pollution Research, 10 (2019) 152-161,

[27] Liu, Xiaoping, Jinpei Ou, Yimin Chen, Shaojian Wang, Xia Li, Limin Jiao, Yaolin Liu, Scenario simulation of urban energy-related CO2 emissions by coupling the socioeconomic factors and spatial structures, Applied energy, 238 (2019) 1163-1178,

[28] Capellán-Pérez, Iñigo, Ignacio de Blas, Jaime Nieto, Carlos de Castro, Luis Javier Miguel, Oscar Carpintero, Margarita Mediavilla, Luis Fernando Lobejón, Noelia Ferreras-Alonso, Paula Rodrigo, MEDEAS: A new modeling framework integrating global biophysical and socioeconomic constraints, Energy & environmental science, 13 (2020) 986-1017,

[29] Fang, Debin, Xiaoling Zhang, Qian Yu, Trenton Chen Jin, Luan Tian, A novel method for carbon dioxide emission forecasting based on improved Gaussian processes regression, Journal of cleaner production, 173 (2018) 143-150,

[30] Shabani, Elham, Babollah Hayati, Esmaeil Pishbahar, Mohammad Ali Ghorbani, Mohammad Ghahremanzadeh, A novel approach to predict CO2 emission in the agriculture sector of Iran based on Inclusive Multiple Model, Journal of Cleaner Production, 279 (2021) 123708,

[31] Ahmad, Tanveer, Rafal Madonski, Dongdong Zhang, Chao Huang, Asad Mujeeb, Data-driven probabilistic machine learning in sustainable smart energy/smart energy systems: Key developments, challenges, and future research opportunities in the context of smart grid paradigm, Renewable and Sustainable Energy Reviews, 160 (2022) 112128,

[32] Strielkowski, Wadim, Andrey Vlasov, Kirill Selivanov, Konstantin Muraviev, Vadim Shakhnov, Prospects and challenges of the machine learning and data-driven methods for the predictive analysis of power systems: A review, Energies, 16 (2023) 4025,

[33] Pollice, Robert, Gabriel dos Passos Gomes, Matteo Aldeghi, Riley J Hickman, Mario Krenn, Cyrille Lavigne, Michael Lindner-D'Addario, AkshatKumar Nigam, Cher Tian Ser, Zhenpeng Yao, Data-driven strategies for accelerated materials design, Accounts of Chemical Research, 54 (2021) 849-860,

[34] Zhong, Shifa, Kai Zhang, Majid Bagheri, Joel G Burken, April Gu, Baikun Li, Xingmao Ma, Babetta L Marrone, Zhiyong Jason Ren, Joshua Schrier, Machine learning: new ideas and tools in environmental science and engineering, Environmental Science & Technology, 55 (2021) 12741-12754,

[35] Thakkar, Ankit, Kinjal Chaudhari, A comprehensive survey on deep neural networks for stock market: The need, challenges, and future directions, Expert Systems with Applications, 177 (2021) 114800,

[36] Santos, Claudio Filipi Gonçalves Dos, João Paulo Papa, Avoiding overfitting: A survey on regularization methods for convolutional neural networks, ACM Computing Surveys (CSUR), 54 (2022) 1-25,

[37] Oviedo, Felipe, Juan Lavista Ferres, Tonio Buonassisi, Keith T Butler, Interpretable and explainable machine learning for materials science and chemistry, Accounts of Materials Research, 3 (2022) 597-607,

[38] Akay, Bahriye, Dervis Karaboga, Rustu Akay, A comprehensive survey on optimizing deep learning models by metaheuristics, Artificial Intelligence Review, 55 (2022) 829-894,

[39] Haji, Saad Hikmat, Adnan Mohsin Abdulazeez, Comparison of optimization techniques based on gradient descent algorithm: A review, PalArch's Journal of Archaeology of Egypt/Egyptology, 18 (2021) 2715-2743,

[40] Kumar, Manish, Sudhansu K Mishra, A comprehensive review on nature inspired neural network based adaptive filter for eliminating noise in medical images, Current medical imaging, 16 (2020) 278-287,

[41] Odili, Julius Beneoluchi, Ahmad Noraziah, Radzi Ambar, Mohd Helmy Abd Wahab, A critical review of major nature-inspired optimization algorithms, The Eurasia proceedings of science technology engineering and mathematics, (2018) 376-394,

[42] Rahman, Md Ashikur, Rajalingam Sokkalingam, Mahmod Othman, Kallol Biswas, Lazim Abdullah, Evizal Abdul Kadir, Nature-inspired metaheuristic techniques for combinatorial optimization problems: Overview and recent advances, Mathematics, 9 (2021) 2633,

[43] Korani, Wael, Malek Mouhoub, Review on nature-inspired algorithms, Operations research forum, Springer, 2021, pp. 36.

[44] Hagan, Martin T, Howard B Demuth, Orlando De Jesús, An introduction to the use of neural networks in control systems, International Journal of Robust and Nonlinear Control: IFAC- Affiliated Journal, 12 (2002) 959-985, https://doi.org/10.1002/rnc.727.

[45] Mirjalili, Seyedali, Seyed Mohammad Mirjalili, Abdolreza Hatamlou, Multi-verse optimizer: a nature-inspired algorithm for global optimization, Neural Computing and Applications, 27 (2016) 495-513,

[46] Zhao, Huiru, Xiaoyu Han, Sen Guo, DGM (1, 1) model optimized by MVO (multi-verse optimizer) for annual peak load forecasting, Neural Computing and Applications, 30 (2018) 1811-1825,

[47] Kashan, Ali Husseinzadeh, League Championship Algorithm (LCA): An algorithm for global optimization inspired by sport championships, Applied Soft Computing, 16 (2014) 171-200,

[48] Sadollah, Ali, Hadi Eskandar, Ardeshir Bahreininejad, Joong Hoon Kim, Water cycle algorithm with evaporation rate for solving constrained and unconstrained optimization problems, Applied Soft Computing, 30 (2015) 58-71,

[49] Wedyan, Ahmad, Jacqueline Whalley, Ajit Narayanan, Hydrological Cycle Algorithm for Continuous Optimization Problems, Journal of Optimization, 2017 (2017) 1-25, 10.1155/2017/3828420.

[50] Taylor, Karl E, Summarizing multiple aspects of model performance in a single diagram, Journal of geophysical research: atmospheres, 106 (2001) 7183-7192, https://doi.org/10.1029/2000JD900719.