Enhancing Renewable Energy Forecasting Through Adaptive Statistical Modeling and Optimization Strategies

Hua Wang*, Lingwei Wang*‡, Fuyu Zhu*

*Guangdong University of Science and Technology, Dongguan City, Guangdong Province, 523083, China

(15279089598@163.com, wanglingwei9498@163.com, 13113196882@163.com)

‡Corresponding Author; Lingwei Wang, 523083, wanglingwei9498@163.com

Received: 14.07.2025 Accepted:05.08.2025

Abstract- Managing energy resources effectively involves a profound awareness of factors governing consumption behavior, specifically in renewable integration scenarios. With the transition towards the cleaner power grid, demand and supply balancing become ever more complex, and high-level computational approaches must therefore be embraced. This study addresses renewable energy consumption forecast using machine learning (ML) models optimized with the Red Fox Optimization Algorithm (RFOA) with a view to enhancing predictive generalizability and stability for optimal operation under changing scenarios. For this propose, the study compare the performance of six optimized ML algorithms Elastic Net, Extreme Gradient Boosting (XGBoost), Random Forest (RF), Light Gradient Boosting Machine (LightGBM), Categorical Boosting (CatBoost), and Histogram-Based Gradient Boosting (HGB). Comparative analysis demonstrates that RFOA-ElasticNet, with the high R^2 of 0.9859 on the training set and 0.9481 on the test set, indicates excellent explanatory power and generalization. Furthermore, this model with MAE and MASE of 93.24 and 0.2603 on the test data confirms its reliable predictive accuracy with minimal prediction errors. Aside from predictive performance, an understanding of driving factors in consumption is critical in improving model interpretability. Based on the feature importance analysis, it is determined that certain external factors, including carbon intensity in electricity generation with an impact factor of around 140, have been discovered to act as principal drivers of consumption trends. These findings yield useful information for operators, professionals, and policymakers for planning at a strategic level.

Keywords: Energy supply chain; Renewable energy; Collaborative control mode; Machine learning; Optimization algorithm.

1. Introduction

The global energy scenario is transforming with the growing integration of renewable sources [1]. On the other hand, traditional grid infrastructure is inefficient, unreliable, and not environmentally friendly, so new approaches are in demand. The new energy supply chain service communication power system network collaborative control mode is a new technique for the coordination, control, and optimization of new energy networks. For instance, a novel control strategy is introduced using load shaving, load shifting, and energy storage to ensure voltage and frequency stability in microgrids, reduce peak-hour electricity costs, and optimize system performance [2]. Energy demand is increasing across the world due to industrialization and changing lifestyles, while conventional thermal sources decrease, resulting in power shortages and load shedding. Sustainable energy sources are encouraged by governments, while solar energy has become popular due to its environmental advantages [3]. Therefore, a dynamic, adaptable mode is critical in assuring network resiliency and stability to control mode, which uses cutting-edge communications technology, real-time data analysis, and decentralized decision-making.

The critical component of the new energy supply chain is the management of distributed energy sources, which include renewable generation sources (e.g., solar panels and wind farms), storage systems, and electric vehicles [4] . For efficient management of energy systems, coordinating disparate sources of power in a manner that stabilizes the grid, maintains power in high form, and accurately deals with demand for

energy in a reliable form [5]. High-performance algorithms, powered with artificial intelligence (AI) and machine learning (ML), make help to these systems with the predictions regarding the production and consumption of energy for real-time decision processes [6]. In these smart systems for developing high-performance AI and ML algorithms, dataset serves as the foundational element [7]. The dataset consists of high-resolution time-series data collected from a real-world energy system, capturing key variables related to energy production, consumption, storage, and environmental conditions. The AI and ML model using these gathered data, create patterns, predict behavior, and optimize the operations thus maximizing stored energy usage. Indeed, using real-time and past datasets, these algorithms back autonomous decision process whereby the energy systems can react dynamically to changing situations without relying on human input [8]. This allows grid operators to coordinate demand and supply, modulate energy flows in relation to grid conditions, and make the fullest use of renewable sources.

Existing literature describes a move towards new, distributive, and smart systems, emphasizing the role of the smart grid in real-time data exchange and flexible control. A critical role is also placed on high-tech networks in the development of new energy systems [9–12]. In this context, Al-Mayyahi et al. [13] explored the use of LSTM neural networks with the Nadam optimizer to predict solar PV system power output, addressing the challenges of intermittency. Their findings show that the LSTM model, when combined with the Nadam optimizer, provides more accurate predictions than traditional models like SARIMA and ARIMA, thereby improving the reliability and profitability of solar energy systems. Ahmad and Nduka [14] addressed the challenges associated with traditional power flow analysis methods, which involve solving nonlinear equations to determine the steady-state operating conditions of the power grid. These methods, although accurate, are computationally expensive and may struggle with convergence, leading to high time and space complexity. To overcome these limitations, the authors explore the application of Machine Learning (ML) algorithms for solving power flow equations. The paper provides a comprehensive comparative analysis of different ML algorithms, with experimental simulations conducted on IEEE 3-bus and IEEE 118-bus networks using custom-developed Python codes. The results highlight the effectiveness of ML in improving the efficiency and accuracy of power flow analysis. Further contributing to the optimization of solar energy systems, Chen et al. [15] developed a two-stage metaheuristic framework for parameter identification in solar cell models using the Sea Lion Optimizer (SLO). The results showed superior performance compared to other algorithms. Their improved version, ISLO, incorporating Lévy flight and personal best information, further enhanced accuracy, resulting in a 59.7% improvement in RMSE. The two-stage methodology, which combines initial estimates with refined search ranges, further reduced RMSE by up to 91.3%, demonstrating significant improvements in parameter estimation. Khan et al. [16] introduced an Adaptive Energy Management System (AEMS) for smart grids, leveraging machine learning and IoT to optimize load forecasting and energy management. The system, tested on a mock-up with LSTM-based forecasting, reduced household energy costs by 28.17%, improved renewable energy utilization, and enhanced grid stability by over 70%. Scalability tests showed the system's capability to operate effectively even with large-scale configurations, making it a promising solution for smart grid performance and renewable energy integration. Brahmankar et al. [17] presented a Genetic Algorithm (GA)-based approach for optimizing load scheduling in residential energy systems. This approach minimized energy costs by 36.9%, increased PV energy use by 5.1%, and enhanced battery storage by 65.9%. The study highlights the potential of GA in improving home energy management efficiency and sustainability, emphasizing the importance of parameter tuning for optimal results.

Recent advancements in ML have exhibited a high potential for predicting energy production/consumption trends and optimizing the distribution of resources. Some of the ML algorithms, including time series analysis, neural networks, and ensemble approaches, have been adopted for high-accuracy prediction of renewable energy production and consumption. A novel method of community energy supply management was introduced in [18], which showed lower community electricity costs and less dependence on centralized distribution grids by forecasting electricity consumption via ML, specifically Extreme Gradient Boosting (XGBoost), along with a decision algorithm for energy trading with the public grid. Mbasso et al. [19] presented a model of energy consumption estimation, created by applying a hybrid Sinusoid-polynomial to model seasonality and effects of non-linearities. The observed results showed that temperature has a considerable influence on energy demand and production. A Long Short-Term Memory (LSTM)-based controller, is used for inverter control, which showed superior dynamic performance. Grève et al. [20] developed ensemble model combining two neural networks (Bidirectional LSTM (BLSTM) and Multilayer Perceptron (MLP)) and two tree-based methods (Gradient Boosting (GB) and Random Forest (RF)) models to help communities with their energy generation and consumption scheduling. Intravaia et al. [21] introduced a Autoencoders strategy for generating realistic hourly energy consumption profiles using data typically available from energy bills. Bogner et al. [22] analyzed predictive uncertainty and accuracy of energy consumption and production forecasts with a variety of ML approaches, with a strong focus in southern Switzerland. Dimitropoulos et al. [23] presented a solution to forecast short-term generation of energy at community solar plants via usage of a combination of monitoring and infrastructure tools such as XGBoost. Molu et al. [24] presented a novel approach for short-term solar irradiance forecasting by integrating Bayesian Optimized Attention-Dilated LSTM with Savitzky-Golay filtering using data from a solar irradiance probe in

Douala, Cameroon. Table 1 summarizes the studies conducted to utilize ML techniques for particular applications in energy system. Based on this literature review in energy forecasting models, existing researches present inability to generalize, the tendency to overfit, and lack of interpretability especially where several interacting, multifaceted factors influence energy systems.

Table 1. Summary of the ML techniques used to forecast the production/consumption of energy

Studies	Context	Best performance ML model	Metric
Ref [20]	Wind power prediction	Ensemble	RMSE = 2327 (1 day ahead)
Ref [21]	Communities' consumption forecast	Autoencoders	RMSE= 0.1 (average of all users/year)
Ref [22]	Production forecast	Multivariate Adaptive Regression Splines and Gradient Boosting Machines	R^2 = 0.61
Ref [22]	Consumption forecast	Quantile Regression Forest and Gradient Boosting Machines	R^2 = 0.83
Ref [23]	Solar power production forecast	Extreme Gradient Boosting	RMSE = 1.834 (6 hour ahead)
Ref [24]	Solar irradiance forecasting	Bayesian Optimized Attention-Dilated LSTM	RMSE = 22.94 (5 day ahead)
Ref [25]	Consumption forecast	Support Vector Machine	RMSE = 759.78 (15 min ahead)
Ref [26]	Energy sources categorization forecast	Random Forest	Accuracy = 81.1%

The nonlinear and complex relationships found in energy systems are frequently not sufficiently captured by traditional methods, which results in decreased accuracy and adaptability. Optimizing ML model performance through efficient hyperparameter tuning is a crucial tactic for overcoming these constraints. In this study, these issues are addressed by combining ML models with an advanced optimization algorithm to improve predictive transparency and robustness. The study present comparative analysis between six optimized ML algorithms Elastic Net, XGBoost, RF, Light Gradient Boosting Machine (LightGBM), Categorical Boosting (CatBoost), and Histogram-Based Gradient Boosting (HGB). The main innovation of this research is the model optimization using hyperparameter tuning with the help of the Red Fox Optimization Algorithm (RFOA), which makes it possible to determine the best parameter setups for every model. Furthermore, in order to validate RFOA's optimization capability, a comparative evaluation is carried out with the Bayesian Optimization (BO) on the top-performing model hyperparameter tuning. Finally, the study bridges the gap between predictive performance and model transparency by analyzing and visualizing the impact of each input feature on the model outputs using Pareto plots and SHAP (Shapley Additive Explanations) values to ensure interpretability.

This study is designed in several sections in an attempt to present a thorough study of the estimation of renewable energy consumption. Section 2 is the methodology, in which preprocessing and collection of data, ML approaches, and performance evaluation for comparing the performance of models are discussed. Section 3 is results and discussion, in which hyperparameters tuning results, model comparison, and feature importance analysis are discussed. Section 4, in conclusion, presents key findings, real-life implications, and future avenues for research.

2. Methodology

2.1. Modelling Procedure

This study in order to increase the forecasting accuracy of renewable energy systems, provide the comparative analysis between optimized Elastic Net, RF, XGBoost, LightGBM, CatBoost, and HGB models, which was chosen for thier unique capabilities in modeling intricate and nonlinear data patterns. For instances, ensemble-based models like RF offer the capacity to handle high-dimensional datasets, while Elastic Net, offers transparency and simplicity through regularization. Furthermore, Gradient boosting methods like XGBoost, LightGBM, HGB and CatBoost, improve learning effectiveness and capture complex feature interactions, making the framework extremely flexible to the complex nature of renewable energy data. As seen in Fig. 1, the suggested methodology forecasts renewable energy starts with comprehensive data preparation. Following preprocessing, input features and output targets are determined and the dataset are separated into training and testing sets. This study employe a k-fold cross-validation strategy to improve the models' generalizability and reliability. Therefore, the intricate, nonlinear relationships between input and output variables are learned using a variety of selected ML models. In this step, RFOA algorithm facilitates an effective and automated exploration of the hyperparameter space, using tune hyperparameters in order to guarantee optimal model performance. In order to determine the most accurate and

dependable forecasting model, all developed models are thoroughly assessed and contrasted using a variety of performance metrics in the final stage. Furthermore, the impact of each input feature on the model outputs are analyzed using Pareto plots and SHAP.

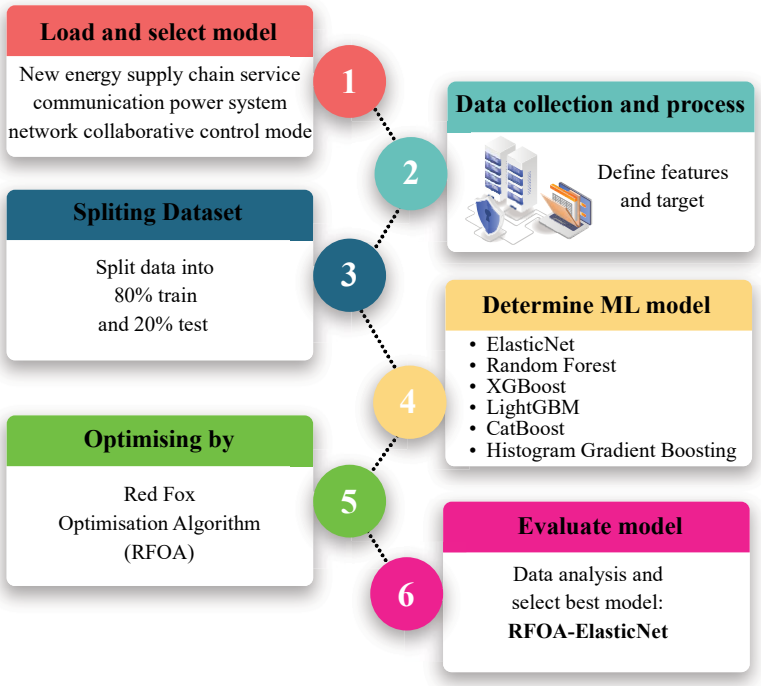


Fig. 1. Pipeline of proposed methodology for renewable energy forecasting.

2.2. Dataset

The study's dataset, which includes key global indicators affecting energy consumption, is taken from [27] and and available in https://github.com/owid/energydata/tree/master?tab=readme-ov-file. The dataset offers a long-term view of global energy and environmental indicators since it includes annual data that records one entry per location per year. Using ISO 3166-1 alpha-3 codes and aggregated regional categories, it follows OWID's standardized naming conventions and offers extensive geographic coverage, encompassing nations and regions worldwide. The time period varies by indicator: population data is available from 1950, Gross Domestic Product (GDP) from 1960, coal consumption from 1900, gas and renewables consumption from 1965, electricity generation and carbon intensity from 1985, and greenhouse gas emissions from 1990 onward.

In general, it includes roughly 1,848 data points, each of which represents a distinct observation that adds to a comprehensive understanding of global trends in energy usage. The population, GDP, carbon intensity of electricity production, coal and gas consumptions, electricity generation, renewable consumption, and greenhouse gas emissions are some of the important variables included in the dataset that significantly influence patterns of energy consumption. Since the main objective of this analysis is to forecast the amount of renewable energy use worldwide, renewable energy consumption is chosen as the target variable among these.

The dataset goes through a rigorous preprocessing step to guarantee its quality and suitability for analysis prior to the application of any ML techniques. For this propose, data cleaning is used to eliminate or fix errors and inconsistencies, normalization is used to scale numerical values into a consistent range, and imputation is used to make sure that no information gaps jeopardize the modeling process. In order to assess model performance on unseen data, the dataset is then split into two subsets: 20% (370 samples) for testing and 80% (1,478 samples) for training. Moreover, a five-fold cross-validation technique is used to further guarantee the ML models' dependability and generalizability. Using this method, the data is divided into five equal folds that four-folds are used for training and one-fold is set aside for validation in each iteration.

The correlation matrix in Fig. 2 reveals important information regarding variable relations. There are high positive relations between factors that have an economic and an energy-related basis, including renewables consumption, fossil fuels consumption, Gross Domestic Product (GDP), and electricity generation, suggesting a high level of interdependence between them. Conversely, the carbon intensity of electricity is less correlated with most features, and its variation must therefore depend on factors other than these. Most salient, greenhouse gas emissions have a significant relation with the consumption of fossil fuels, corroborating its important role in environmental impact assessments. This analysis guides model development and feature selection, forming a reliable platform for optimizing energy systems.

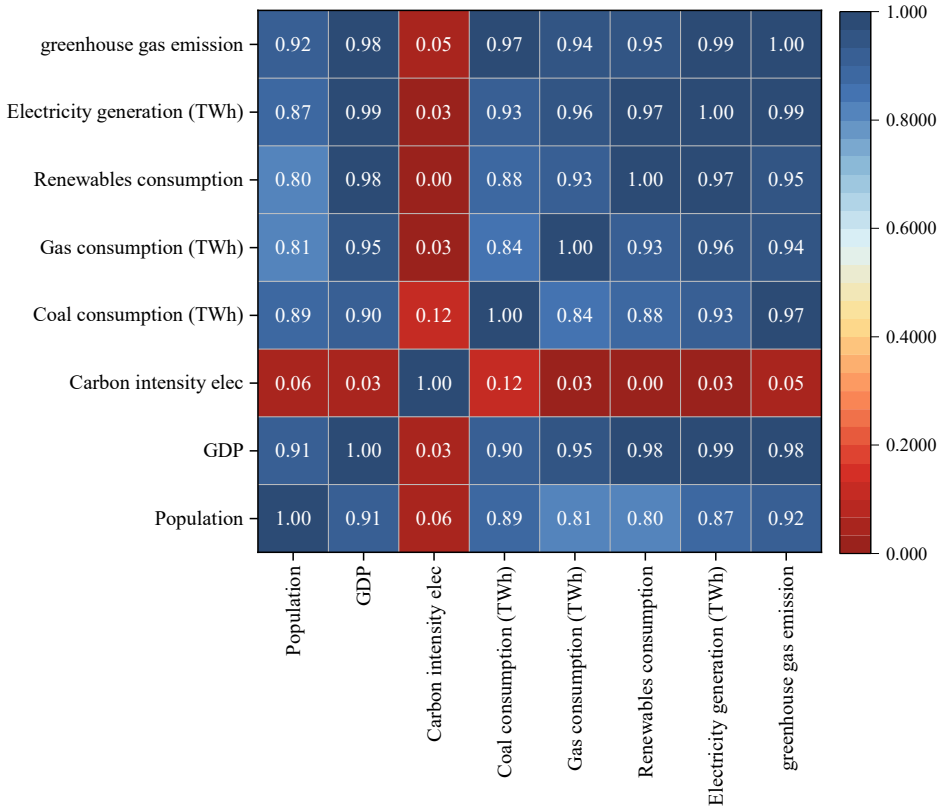


Fig. 2. Correlation matrix of key features.

2.3. Machine Learning Techniques

2.3.1. Elastic net

Elastic Net regression [28,29] is a hybrid statistical model for the regularization and selection of important predictor variables with a high impact on the response variable, as well as addressing multicollinearity among predictor variables. It integrates both l1 (also known as Lasso) and l2 (also known as Ridge) penalties, which enables a trade-off between regularization and feature selection. This makes it especially valuable in multiple domains, where datasets usually contain several connected factors. By tuning hyperparameters regulating the l1 and l2 penalty mix, Elastic Net becomes adaptable in choosing important features and preventing overfitting, hence, it becomes a model of preference for most predictive modeling cases.

2.3.2. Random forest

Random Forest (RF) [30,31] is a standard non-parametric model that develops forecast rules using various types of predictor variables with no consideration of their relationship with the target variable. It generates an ensemble of decision trees during training and takes an average over them to produce more reliable and stable predictions. By using a technique known as bootstrap aggregation (bagging), RF introduces randomization into both feature selection and data sampling, reducing the risk of overfitting that can occur with a single decision tree. Besides, via tuning the hyperparameters, the model can become even more robust for various types of datasets.

2.3.3. Extreme gradient boosting

Extreme Gradient Boosting, or XGBoost [32,33], is an improved version of the gradient boosting machine that belongs to the decision tree boosting models which employs regularization to successfully counteract overfitting. In its iterative process, the loss function decreases in the direction of its gradient, increasing the adaptability of the model. XGBoost adopts level-wise growth, a computationally expensive strategy. Chen and Guestrin [34] claimed that data scientists frequently employ XGBoost to get cutting-edge outcomes on a variety of machine learning problems. Users may tune the hyperparameters to attain the best results.

2.3.4. Light gradient boosting machine

Light Gradient Boosting Machine (LightGBM) [35,36] is another type of gradient boosting decision tree model involving exclusive feature bundling (EFB) and gradient-based one-side sampling (GOSS). Trees in LightGBM are expanded leaf-wise, producing deeper and more computationally efficient forests with less processing than level-wise tree expansion found in traditional boosting techniques. As a result, it is immensely scalable, with capabilities to work with datasets with a high instance count and yet utilizing less memory. LightGBM is also known to be incredibly simple to use, with inbuilt hyperparameter tuning, allowing one to tune performance with ease.

2.3.5. *Categorical boosting*

Categorical Boosting, known as CatBoost [37,38], is a new boosting model capable of handling categorical, numerical, and even textual data. The model constructs a series of decision trees, with each new tree being built at a smaller loss when compared to the tree before it. CatBoost is preferred in most real-life tasks due to its capability to handle missing variables and reduce prediction bias. However, even with the stability and ease of use of CatBoost, much attention is still necessary to the choice of hyperparameters in order to achieve the highest accuracy.

2.3.6. *Histogram gradient boosting*

Histogram Gradient Boosting (HGB) [39,40] is a boosting model designed to address the main weakness of the gradient boosting algorithm, which is its lengthy duration of training when applied to large datasets. This challenge is handled by discretizing or binning the continuous input variables into numerous distinct values. Like LightGBM, HGB utilizes a leaf-wise tree growing algorithm that overcomes overfitting with inbuilt regularization techniques and enables deeper trees and a higher model fidelity. By discovering the best sets of hyperparameters for a specific dataset, HGB optimizes performance, overcomes overfitting, and enhances prediction accuracy.

2.3.7. *Optimization algorithm*

In the fast-growing computational intelligence context, optimization algorithms become vital for the tuning of ML models and efficiency improvement. To improve the accuracy and capability of the ML models, this work adopts the Red Fox Optimization Algorithm (RFO) [41], a novel method inspired by the adaptive intelligence and hunting strategy of red foxes. The RFOA is a metaheuristic nature-inspired algorithm that mimics red foxes' intelligent foraging behavior. The approach efficiently balances exploration and exploitation by continuously modifying the motion patterns of search agents in a way that enables them to avoid local optima while converging toward a global solution. The method includes adaptive techniques to narrow search spaces, making it ideal for high-dimensional and difficult optimization problems (see [42,43]). Fig. 3 shows how the RFOA is incorporated in an ML pipeline, where it is used for hyperparameter tuning of the model. RFO combined with the learning model enhances the predictive performance and generalization, indicating the effectiveness of this approach to ML workflow optimization.

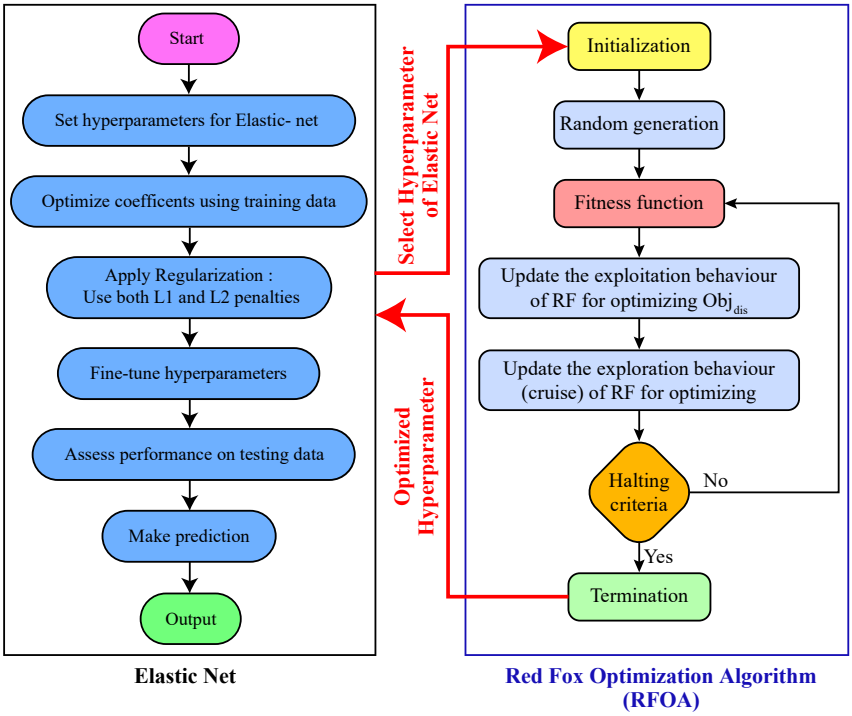


Fig. 3. Dynamic structure of the model hyperparameters tuning via RFOA.

As mentioned previously, for better results, some hyperparameters must be adjusted for certain ML models. Regarding Elastic Net, RFOA balances the contribution of L1 and L2 regularization by optimizing alpha and l1_ratio hyperparameters. In RF, decision tree efficiency is increased by tuning important hyperparameters such as max depth, min samples split, and n-estimators. Furthermore, the algorithm optimizes learning rate, max depth, and n-estimators for XGBoost and LightGBM to improve gradient boosting performance. RFOA also seeks the optimum setting of the hyperparameters of the CatBoost and HGB models, including learning rate and depth, along with the optimal number of iterations. These models will obtain optimum settings via the flexibility and intelligence of the RFOA to improve accuracy and generalization on the datasets.

2.4. *Evaluation Metrics*

In assessing predictive model performance, strong evaluation metrics must be utilized, capable of accurately

estimating accuracy and dependability. Such metrics enable one to understand model generalizability over the dataset and its performance in real-world use cases. There are several key metrics that have been considered in this study including coefficient of determination (R^2), mean absolute error (MAE), mean squared error (MSE), and mean absolute scaled error (MASE), Max Error, Huber Loss, Wilcoxon Stat and Wilcoxon p-value, each provides a specific view of error and goodness of fit. For instance, the R^2 value indicates how effectively the model explains variance in the target variable [44]. The MAE is the widely implemented loss function in regression issues that takes the average of the absolute variations between the forecasted and actual values [45]. The MSE calculates the average of the squared variations among the forecasted and actual values [44]. MASE is another generally used metric for comparing the accuracies of prediction errors at various scales [46]. The worst-case prediction is highlighted by the Max Error, which is the greatest absolute difference between the expected and actual values [47]. By combining the advantages of MSE and MAE, Huber Loss is a robust loss function that is less susceptible to outliers [48]. Wilcoxon Signed-Rank test offers a non-parametric way to compare paired samples and the corresponding p-value indicating whether those differences are statistically significant [49].

3. Results and Discussion

3.1. K-Fold Cross Validation Outcomes

Table 2 show 5-fold cross-validation outcomes without hyperparameter tuning, which provides a trustworthy assessment of the model's performance across several data splits. Based on the results the ElasticNet model demonstrate the consistency based on R^2 scores between 0.9304-0.9387 in the folds, indicating its robustness and strong generalization ability.

Table 2. 5-fold cross-validation outcomes for R^2 without models hyperparameter tuning

Fold	ElasticNet	RF	XGBoost	LightGBM	CatBoost	HGB
1	0.9387	0.8607	0.8008	0.5065	0.5433	0.4083
2	0.9364	0.8662	0.7529	0.4815	0.6750	0.4758
3	0.9381	0.8681	0.6689	0.4204	0.5947	0.4726
4	0.9382	0.8611	0.7956	0.4623	0.5915	0.4524
5	0.9304	0.8658	0.6956	0.4200	0.5923	0.4688

3.2. Optimization Outcomes

Model hyperparameter tuning provides insightful information, promoting enhanced performance and generalization. Table 3 summarizes the tuning findings and provides the optimal settings for the different models. These are the settings that yielded the optimal experimentation results. Elastic Net demonstrates the importance of achieving a balance between L1 and L2 regularization by indicating suitable values for alpha and l1_ratio. Given the RF tuning findings, the min_ samples_ split and n_ estimators are 2 and 50, respectively, while the ideal max_depth is unlimited. Gradient boosting models, including XGBoost, LightGBM, CatBoost, and HGB, show consistency in depth, learning rate, and the number of iterations, suggesting a balanced approach to learning complexity. The observed results reveal how each model responds to parameter changes, which aids in the refinement of performance for certain tasks.

Table 3. Hyperparameters tuning outcomes

Model	Hyperparameters	Optimum value
ElasticNet	alpha	0.1
	l1_ratio	1.0
RF	max_depth	None
	min_samples_split	2
	n_estimators	50
XGBoost	learning_rate	0.1
	max_depth	7
	n_estimators	200
LightGBM	learning_rate	0.2

	max_depth	5
	n_estimators	200
CatBoost	depth	7
	iterations	200
	learning_rate	0.2
HGB	depth	7
	iterations	200
	learning_rate	0.2

3.3. Models Comparison

It is significant to evaluate and compare the predictive performance of models in terms of efficacy in processing complex structures of the dataset. In this section, a comparative analysis of six optimized models, i.e., RFOA-ElasticNet, RFOA-RF, RFOA-XGBoost RFOA-LightGBM, RFOA-CatBoost, and RFOA-HGB, is conducted. As demonstrated in Fig. 4, each of these models performs variously during the training and testing sets. Notably, some models have great generalization capabilities despite having lower training accuracy, and other models with high training performance underperform when presented with new, unexplored data. The highest-performing model, RFOA-ElasticNet, as shown in the test set, achieves the best balance and robustness, stressing its real-world utility. This finding emphasizes model selection in terms of practical generalizability rather than training efficiency, supporting the desire for validation techniques that prioritize flexible, expandable controls in the energy supply chain's service network.

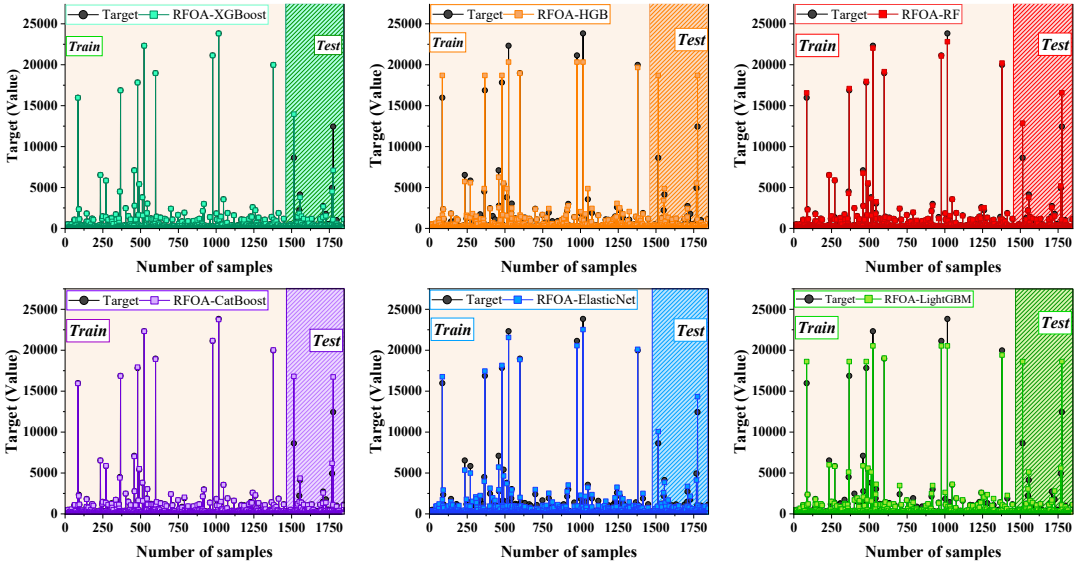


Fig. 4. Value plots of several optimized models.

Figure 5 shows the performance comparison of the ML models for forecasting renewable energy. Most models have high R^2 train values (~0.99), suggesting a proper fitting during training. In particular, RFOA-XGBoost (0.9999), RFOA-CatBoost (0.9998), and RFOA-RF (0.9993) obtain the highest training scores. However, these models demonstrate a considerable loss of generalization in their test performance. RFOA-ElasticNet outperforms other models in terms of generalization, with R^2 test scores of 0.9481, and relatively low error value. On the other hand, test results for RFOA-LightGBM and RFOA-HGB are the poorest, which may indicate overfitting.

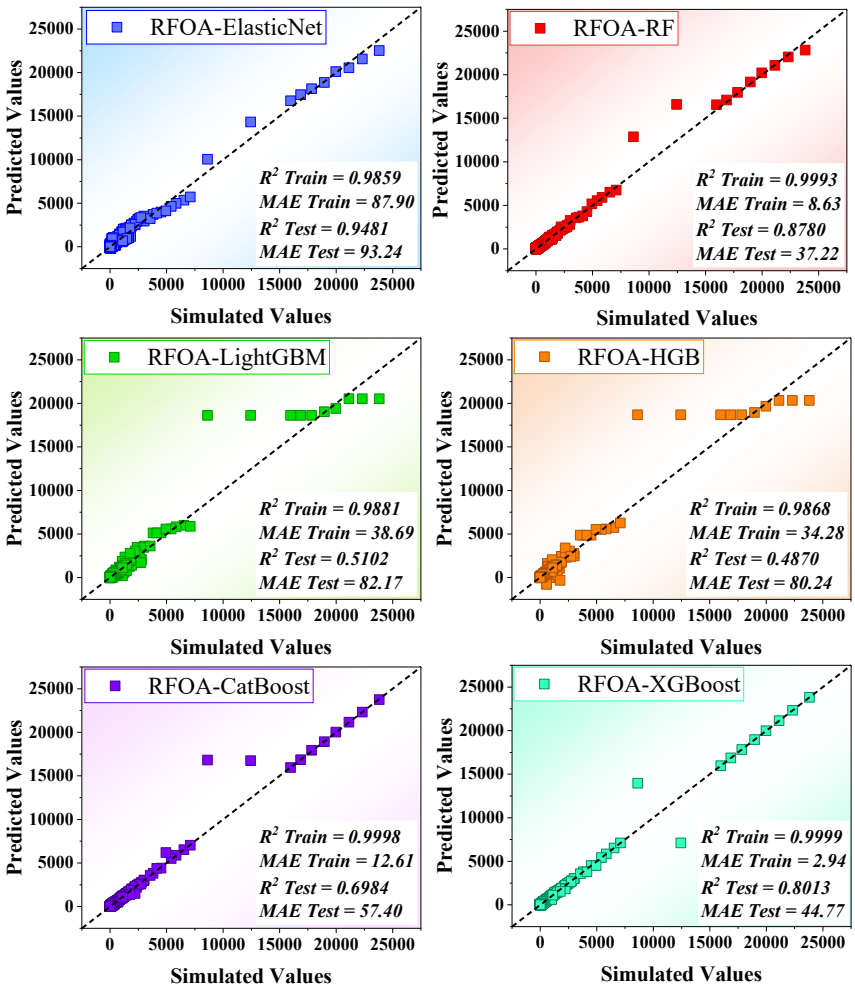


Fig. 5. Predicted-simulated plots of several optimized models.

Figure 6 compares the performance of the various models, considering their error distribution plots during the test phase, providing valuable insights into the reliability of each optimized model. In general, it can be noticed that the RFPA-ElasticNet model outperforms since it possesses a narrow error distribution with a zero mean. The comparatively small distributions of RFOA-XGBoost, RFOA-CatBoost, and RFOA-RF further suggest high generalization. RFOA-LightGBM and RFOA-HGB, on the other hand, have broader distributions, which most likely reflect their high levels of instability and prediction.

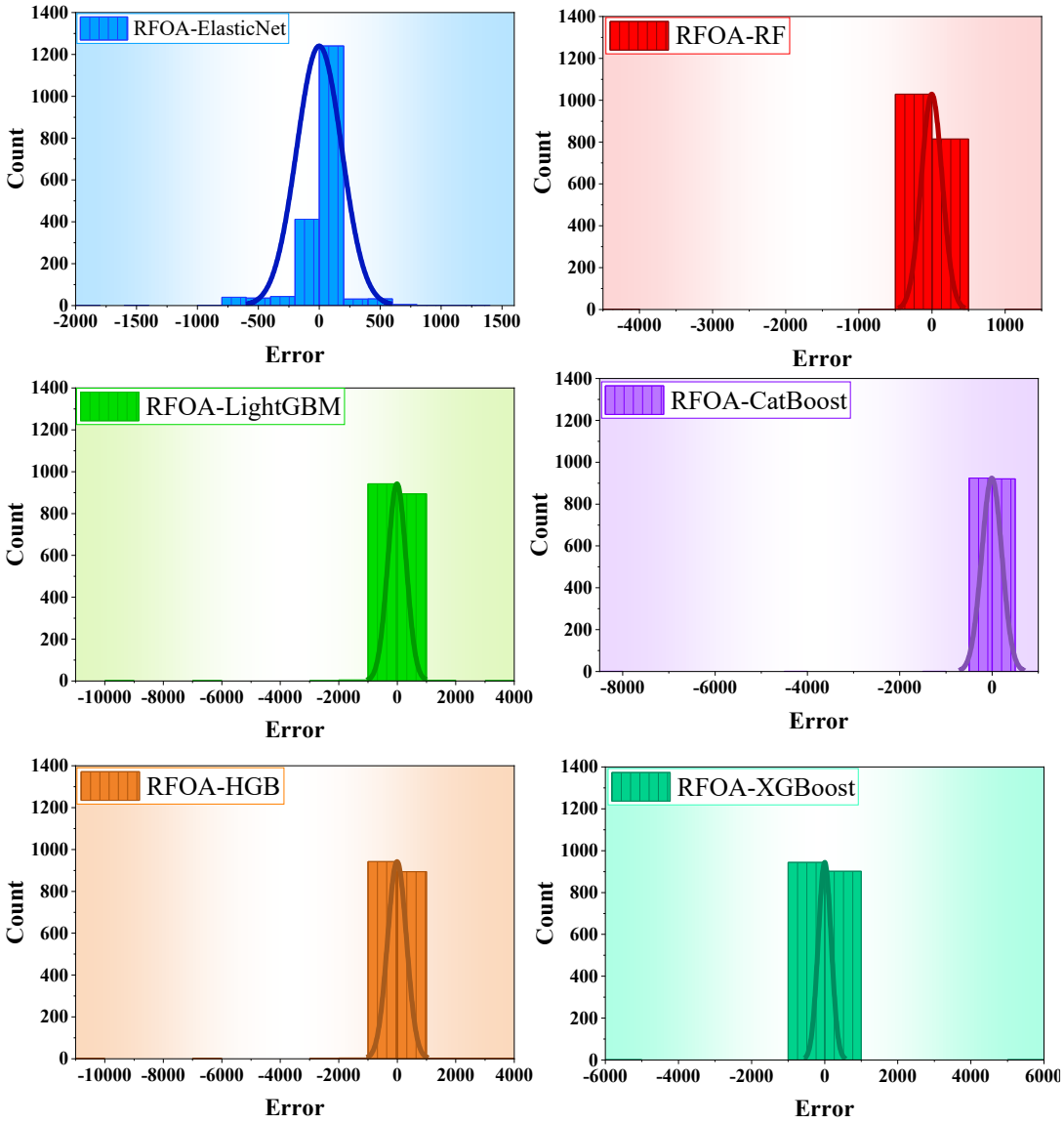


Fig. 6. Error distribution plots of several optimized models.

In examining the performance of several models for renewable energy forecasts, RFOA-ElasticNet appeared as the most effective model, as shown in Fig. 7, which provides an overview of the results of analyzing the optimized models using various metrics. Throughout the testing phase, it performed admirably, earning the greatest R^2 score and the lowest MSE value. Furthermore, the RFOA-XGBoost, RFOA-CatBoost, and RFOA-RF models functioned successfully, particularly during the training phase. However, there were greater variations during the test phase, suggesting that overfitting may have occurred. RFOA-LightGBM and RFOA-HGB had proper values for each measure during the training phase, whereas their performance dropped dramatically during the testing phase, resulting in the lowest R^2 value and the highest MAE, MSE, and MASE values.

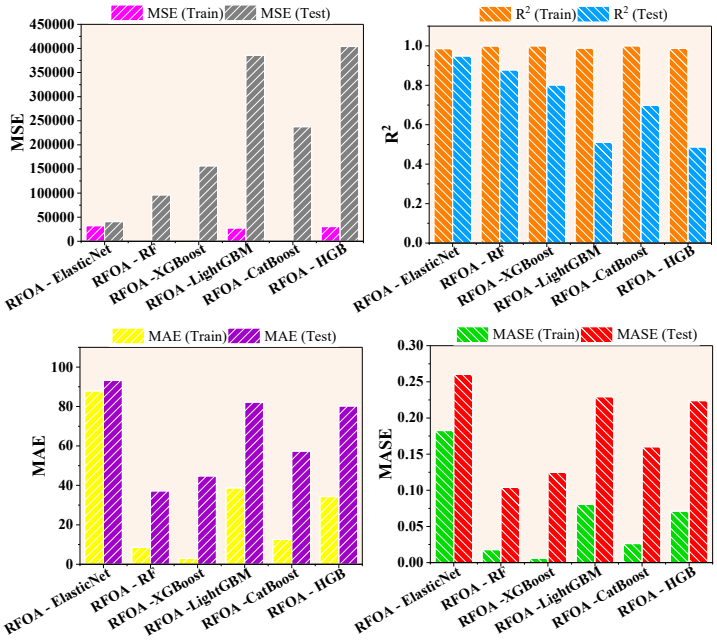


Fig. 7. Comparison of several optimized models based on metrics.

The boxplots in Fig. 7 provide information about the model's accuracy and consistency by showing how different regression performance metrics are distributed across several model runs. The model successfully explains the data's variability and operates consistently across runs, as evidenced by metrics like R^2 that shows values near 1 with little variance. Likewise, the MASE metric exhibits tight distributions and low central values, indicating strong forecasting accuracy and resilience to outliers. Metrics like MaxError and MSE, show some spreads and outliers, emphasizing the existence of sporadic prediction errors that could be caused by unusual data points. In terms of absolute error, the Huber Loss and MAE metrics show balanced performance, maintaining moderate ranges with comparatively stable medians.

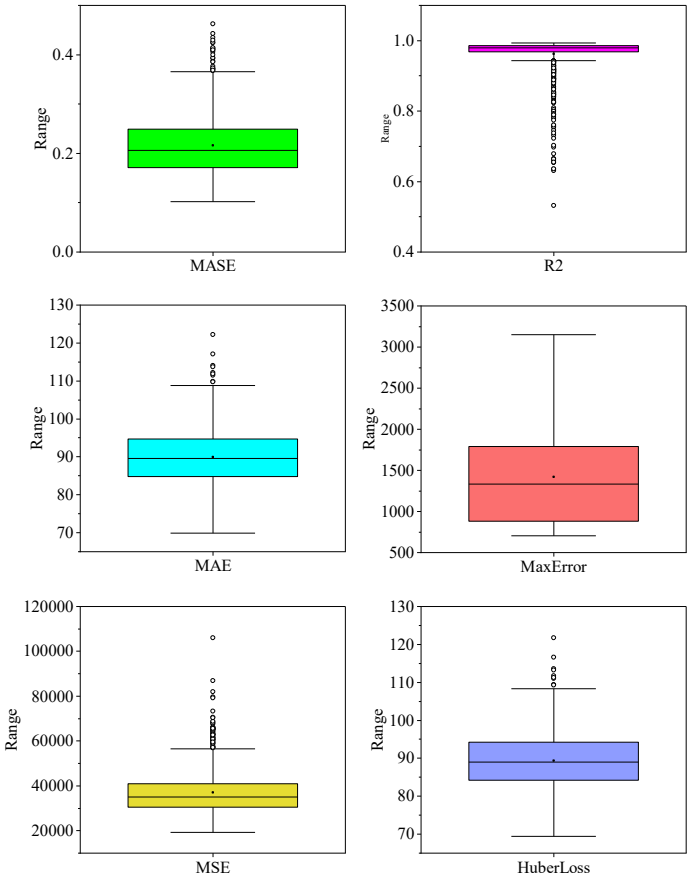


Fig. 7. The boxplot for different regression performance metrics across several runs for RFOA-ElasticNet model.

Table 4 provides a comprehensive review of the performance of all optimized models and their evaluation outcomes for both the training and testing sets. It reveals that although a few of the models did well in training, there was a significant variation in their performance when tested. The strongest outcome was obtained from RFOA-ElasticNet, with a perfect balance between learning performance and generalizability. In the training phase, it reached a great value of 0.9859 for the R^2 metric. Interestingly, RFOA-XGBoost, RFOA-CatBoost, and RFOA-RF exhibited extremely low values for error statistics and near-perfect values for R^2 (0.9999, 0.9998, and 0.9993, respectively); nevertheless, such observations did not generalize in testing. In contrast, RFOA-ElasticNet performed better throughout the test phase, with the greatest R^2 of 0.9481 and the lowest MSE of 40,865.18. This shows its great generalizability, which makes it suitable for real-world applications. Meanwhile, models including RFOA-LightGBM and RFOA-HGB were less generalizable, with R^2 values lower than 0.7 and much higher error metrics. Furthermore, the Wilcoxon p-value outcomes also, confirm the effectiveness of RFOA-ElasticNet model in prediction because obtained Wilcoxon p-value for RFOA-ElasticNet model is lower than 0.5.

Table 4. Optimized models evaluation outcomes

Model	R^2	MAE	MSE	MASE	Max Error	Huber Loss	Wilcoxon Stat	Wilcoxon p-value
	Train							
RFOA-ElasticNet	0.9859	87.90	32,414.88	0.1826	1376.823	87.400	358037	1.60E-30
RFOA-RF	0.9993	8.63	1,673.42	0.0179	865.188	7.731	503625	0.009
RFOA-XGBoost	0.9999	2.94	18.82	0.0061	20.572	2.499	542064	0.787
RFOA-LightGBM	0.9881	38.69	27,510.29	0.0804	3296.164	38.202	540032	0.693
RFOA-CatBoost	0.9998	12.61	374.68	0.0262	144.181	12.123	538092	0.608
RFOA-HGB	0.9868	34.28	30,443.98	0.0712	3492.836	33.800	542842	0.824
	Test							
RFOA-ElasticNet	0.9481	93.24	40,865.18	0.2603	1874.377	92.747	22118	3.11E-09
RFOA-RF	0.8780	37.22	96,118.55	0.1039	4298.716	35.685	33023	0.529
RFOA-XGBoost	0.8013	44.77	156,492.13	0.1250	5346.971	44.304	34040	0.892
RFOA-LightGBM	0.5102	82.17	385,800.29	0.2294	9984.915	81.681	32944	0.504
RFOA-CatBoost	0.6984	57.40	237,574.33	0.1602	8179.472	56.903	33988	0.872
RFOA-HGB	0.4870	80.24	404,066.10	0.2240	10065.110	79.754	34041	0.893

Table 5 indicates the comparison between the RFOA and BO for tuning the ElasticNet model's hyperparameters, which highlights the superior performance of RFOA in both training and testing phases. Specifically, RFOA-ElasticNet achieved a higher R^2 and lower MAE in comparion to BO-ElasticNet during testing, which recorded an R^2 of 0.9467 and MAE of 97.84. Therefore, this finding, justifies the choice of RFOA for XGBoost's hyperparameter tuning due to its steady improvement over both phases because of improved search space exploration and exploitation. RFOA is a strong contender for modelling of complex, high-dimensional dataset because of its ability to balance global and local search dynamics, which helps it avoid local optima and achieve more accurate model configurations.

Table 5. The comparison between optimized ElasticNet using two different RFOA and BO algorithms.

		R^2	MAE
RFOA-ElasticNet	Train	0.9859	87.90
BO-ElasticNet		0.9807	88.27
RFOA-ElasticNet	Test	0.9481	93.24
BO-ElasticNet		0.9467	97.84

3.4. Sensitivity Analysis

SHAP is a potent interpretability technique for elucidating ML model output in which each feature is obtained an importance value based on cooperative game theory [50,51]. SHAP offers both global and local interpretability by determining each feature's contribution to the discrepancy between the actual prediction and the average prediction, which aids in comprehending how distinct features affect model behavior. Fig. 8 displays the SHAP summary for the RFPA-ElasticNet model, which illustrates the many features that influence prediction. Carbon intensity in electricity generation is most impactful, with both positive and negative effects in relation to its value. Other important factors, albeit with differing degrees, are electricity generation, gross domestic product (GDP), gas consumption, and population. However, the impact of greenhouse gas emissions and coal consumption seems comparatively less.

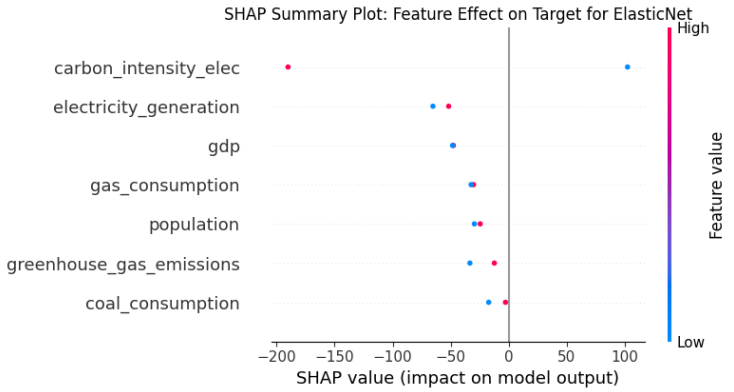


Fig. 8. SHAP summary plot of feature importance for RFOA-ElasticNet.

The Pareto plot in Fig. 9 provides an in-depth overview of the feature relevance for the RFOA-ElasticNet model, which influences renewable energy consumption forecasts. The carbon intensity of the electricity production is the most crucial element since it has a significant influence on model outcomes. This means that variations in emissions caused by power generation strongly influence energy demand trends. The connection between economic activity and energy consumption is also supported by GDP and electricity production, ranking second and third, respectively. The remaining variables, such as gas consumption, population, greenhouse gas emissions, and coal consumption, play a relatively minor but still important role. The ranking of the relative importance of the predictors showcases the model's ability to separate the primary from the secondary effects, with valuable implications for energy policy-making and strategic planning for the deployment of renewable energy.

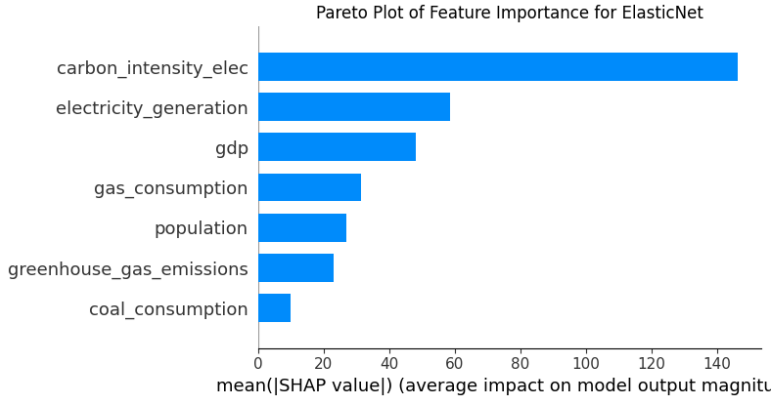


Fig. 9. Pareto plot of feature importance for RFOA-ElasticNet.

4. Conclusion

Achieving accurate and reliable forecasts is not only a matter of having models capable of representing complex dependencies but also having them maintain a level of stability under changing scenarios. Analysis of renewable consumption forecasting reveals a key consideration in selecting machine learning models capable of both adaptability and stability under a range of distributions in data. The models considered in this study were: Elastic Net, Random Forest (RF), Extreme Gradient Boosting (XGBoost) Light Gradient Boosting Machine (LightGBM), Categorical Boosting (CatBoost), and Histogram Gradient Boosting (HGB). Implementation of the Red Fox Optimization Algorithm (RFOA) helped in model improvement. Model performance was boosted, overfitting was lessened, and generalization was enhanced through hyperparameter tuning.

Analysis with regard to metrics showed deeper insights into model performance, with the best model being RFOA-ElasticNet with the high R^2 of 0.9859 on the training set and 0.9481 on the test set, indicating excellent explanatory power and generalization. Furthermore, this model with MAE and

RMSE of 93.24 and 0.2603 on the test data confirms its reliable predictive accuracy with minimal prediction errors. Besides, RFOA-RF exhibited competitive performance with high R^2 of 0.8780, and low MASE and MAE of 0.1039 and 37.22, indicative of their efficiency in handling demand fluctuations for energy. The study then examined predictive factors' contribution to RFOA-ElasticNet model performance through SHAP sensitivity and feature importance analysis, which revealed that carbon intensity in electricity generation, is dominated variable with SHPA impact factor of around 140. Conversely, the impacts of population, greenhouse gas emissions, and coal consumption were relatively minor; however, these continued to contribute towards demand fluctuations for energy.

Overall, this study highlights the significance of predictive modeling and optimization approaches in a bid to make predictions more accurate and reliable in the case of renewable energy consumption. The results of this study strengthen decision-making processes in renewable energy planning and management by offering professionals, grid operators, and regulators useful insights. Our approach enables a more effective shift to renewable energy consumption by creating forecasting frameworks that are both flexible and robust, combining generalizable techniques with important pertinent features. Better demand-supply balancing is made possible by this, which also supports the deployment of smart grid technologies and improves grid reliability. Additionally, the increased forecasting accuracy can help guide regulatory frameworks that encourage the use of sustainable energy, grid infrastructure investments, and policy decisions pertaining to the distribution of energy resources. All things considered, our study fills the gap between sophisticated computer models and realistic deployment situations in the changing energy environment.

Despite demonstrating strong predictive performance, this research has several limitations. Firstly, the proposed models when applied to different geographic regions or rapidly changing energy environments may indicates different results, which future work by incorporating comparative analysis on datasets from diverse regions would enhance model's robustness. Moreover, the study utilizes optimized individual models, which for future work, one promising direction is the utilization of hybrid machine learning models that combine the strengths of different algorithms to enhance both prediction accuracy and robustness. Furthermore, comparative research using several optimization algorithms may offer a more thorough comprehension of tuning effectiveness.

Acknowledgment

This work was supported by the following projects: 2024 Guangdong University of Science and Technology Key Research Project, titled "Optimization of Big Data Collection for Smart Agriculture Internet of Things from the Perspective of New-Form Productivity (Project No.: GKY-2024KYZDK-17)." 2024 Scientific Research Project of Guangdong University of Science and Technology, "Research on Image Denoising Based on Residual and Dense Connection Networks" (No.: GKY-2024KYYBK-11). 2023 Guangdong Province Key Field Special Project, "Research on Electromagnetic Properties of Anisotropic Media and Applications" (Protect No.: 2023ZDZX3049).

Abbreviations

AI	Artificial intelligence	MAE	Mean absolute error
CatBoost	Categorical boosting	MARS	Multivariate adaptive regression splines
DER	Distributed energy source	MASE	Mean absolute scaled error
EFB	Exclusive feature bundling	ML	Machine learning
EV	Electric vehicle	MSE	Mean squared error
GBM	Gradient boosting machine	NGR	Nonhomogeneous gaussian regression
GDP	Gross domestic product	QRF	Quantile random forest
GOSS	Gradient-based one-sided sampling	R2	Coefficient of determination
HGB	Histogram gradient boosting	RF	Random forest
LightGBM	Light gradient boosting machine	RFOA	Red fox optimization algorithm
l1 penalty	Lasso	XGBoost	Extreme gradient boosting
l2 penalty	Ridge		

References

[1] B. Yılmaz, R. Samra, "Suggesting a novel hybrid approach for predicting solar irradiance in the qinghai province of China," Advances in Engineering and Intelligence Systems 003 (2024) 104–122. https://doi.org/10.22034/aeis.2024.473891.1215.

[2] R. Lotfi, B. Kargar, S.H. Hoseini, S. Nazari, S. Safavi, G. Weber, "Resilience and sustainable supply chain network design by considering renewable energy," Int J Energy Res 45 (2021) 17749–17766.

[3] L.N. Rao, R. Somalwar, K. Joshi, P.S. Subudhi, "An Improved MPPT algorithm for performance improvement of 50 kw grid connected PV system," International Journal of Smart Grid 9 (2025).

[4] A. Hassan, "Impact of on-load tap changers and smart controllers on the distributed renewable energy hosting capacity," International Journal of Smart Grid-IjSmartGrid 6 (2022) 132–135.

[5] F. Bodur, O. Kaplan, "Fixed-time sliding mode control for BLDC motor control in UAV," in: 2025 13th International Conference on Smart Grid (IcSmartGrid), IEEE, 2025: pp. 1–9.

[6] S. Mirsaeidi, A.Z. Mendoza, K. Ghaffari, M.Z. Yousaf, K.M. Muttaqi, J. He, "Applications of machine learning in modern power systems: A comprehensive review," in: 2023 3rd International Conference on New Energy and Power Engineering (ICNEPE), IEEE, 2023: pp. 1074–1080.

[7] M.O. Qays, I. Ahmad, D. Habibi, M.A.S. Masoum, "Forecasting data-driven system strength level for inverter-based resources-integrated weak grid systems using multi-objective machine learning algorithms," Electric Power Systems Research 238 (2025) 111112.

[8] M.O. Qays, I. Ahmad, D. Habibi, M.A.S. Masoum, "System strength forecasting using hedge feedforward feedback-based gated recurrent unit algorithm with drift detection analysis," CSEE Journal of Power and Energy Systems (2025).

[9] J. Xiong, Y. Ye, Q. Wang, X. Dong, T. Lu, D. Ma, "A comprehensive review on distributed energy cooperative control and optimization method for energy interconnection system," Electric Power Systems Research 237 (2024) 111007.

[10] J. Liu, L. Ma, Q. Wang, "Energy management method of integrated energy system based on collaborative optimization of distributed flexible resources", Energy 264 (2023) 125981.

[11] M.H. Rehmani, M. Reisslein, A. Rachedi, M. Erol-Kantarci, M. Radenkovic, "Integrating renewable energy resources into the smart grid: Recent developments in information and communication technologies," IEEE Trans Industr Inform 14 (2018) 2814–2825.

[12] L. Xiang, "Energy emergency supply chain collaboration optimization with group consensus through reinforcement learning considering non-cooperative behaviours," Energy 210 (2020) 118597.

[13] S. Al-mayyahi, K. Yahya, A.E. M Yahya, R. Daw Sarreb, M. Aldababsa, "Machine learning techniques for solar power output predicting," International Journal of Smart Grid 8 (2024) 98–107. https://doi.org/https://doi.org/10.20508/ijsmartgrid.v8i2.341.g356.

[14] B. Ahmad, O. Nduka, "A comparative analysis of machine learning based power flow study with custom made open source python codes," in: 2025 13th International Conference on Smart Grid (IcSmartGrid), 2025: pp. 238–243. https://doi.org/10.1109/icSmartGrid66138.2025.11071830.

[15] Z. Chen, K.-C. Ho, Y.-H. Liu, "Enhanced parameter identification of solar photovoltaic models using a novel two-stage improved sea lion optimization algorithm," Artificial Intelligence Research and Applications 1 (2025) 79–92. www.ijaira.org (accessed September 2, 2025).

[16] B. Khan, N.S.S. Ramakrishna, S. Palakurthi, N.B. Tunga, R.R. Mandadi, "Adaptive energy management systems for smart grids: A hybrid approach using machine learning and IoT integration," in: 2025 13th International Conference on Smart Grid (IcSmartGrid), IEEE, 2025: pp. 434–439.

[17] M. Brahmankar, C.-I. Chen, "Multi-objective load scheduling optimization of cost, renewable energy utilization, and load balance using a genetic algorithm," in: 2024 13th International Conference on Renewable Energy Research and Applications (ICRERA), IEEE, 2024: pp. 463–468.

[18] M. Matos, J. Almeida, P. Gonçalves, F. Baldo, F.J. Braz, P.C. Bartolomeu, "A machine learning-based electricity consumption forecast and management system for renewable energy communities," Energies (Basel) 17 (2024) 630.

[19] W.F. Mbasso, R.J.J. Molu, A. Harrison, M. Pushkarna, F.N. Kemdoum, E.F. Donfack, P. Jangir, P. Tiako, M.B. Tuka, "Hybrid modeling approach for precise estimation of energy production and consumption based on temperature variations," Sci Rep 14 (2024) 24422.

[20] Z. De Grève, J. Bottieau, D. Vangulick, A. Wautier, P.-D. Dapoz, A. Arrigo, J.-F. Toubeau, F. Vallée, "Machine learning techniques for improving self-consumption in renewable energy communities," Energies (Basel) 13 (2020) 4892.

[21] M. Intravaia, L. Becchi, M. Bindi, L. Paolucci, F. Grasso, "Autoencoders for Hourly Load Profile Reconstruction in Renewable Energy Communities," in: IEEE EUROCON 2023-20th International Conference on Smart Technologies, IEEE, 2023: pp. 280–285.

[22] K. Bogner, F. Pappenberger, M. Zappa, "Machine learning techniques for predicting the energy

consumption/production and its uncertainties driven by meteorological observations and forecasts", Sustainability 11 (2019) 3328.

[23] N. Dimitropoulos, N. Sofias, P. Kapsalis, Z. Mylona, V. Marinakis, N. Primo, H. Doukas, "Forecasting of short-term PV production in energy communities through machine learning and deep learning algorithms," in: 2021 12th International Conference on Information, Intelligence, Systems & Applications (IISA), IEEE, 2021: pp. 1–6.

[24] R.J.J. Molu, B. Tripathi, W.F. Mbasso, S.R.D. Naoussi, M. Bajaj, P. Wira, V. Blazek, L. Prokop, S. Misak, "Advancing short-term solar irradiance forecasting accuracy through a hybrid deep learning approach with Bayesian optimization," Results in Engineering 23 (2024) 102461.

[25] A. Jozi, T. Pinto, Z. Vale, "Contextual learning for energy forecasting in buildings," International Journal of Electrical Power & Energy Systems 136 (2022) 107707.

[26] H. Musbah, G. Ali, H.H. Aly, T.A. Little, "Energy management using multi-criteria decision making and machine learning classification algorithms for intelligent system," Electric Power Systems Research 203 (2022) 107645.

[27] H. Ritchie, P. Rosado, M. Roser, "Access to energy," Our World in Data (2024).

[28] S. Altelbany, "Evaluation of ridge, elastic net and lasso regression methods in precedence of multicollinearity problem: a simulation study," Journal of Applied Economics and Business Studies 5 (2021) 131–142.

[29] P.J. García-Nieto, E. Garcia-Gonzalo, J.P. Paredes-Sánchez, "Prediction of the critical temperature of a superconductor by using the WOA/MARS, Ridge, Lasso and Elastic-net machine learning techniques," Neural Comput Appl 33 (2021) 17131–17145.

[30] G. Biau, "Analysis of a random forests model," The Journal of Machine Learning Research 13 (2012) 1063–1095.

[31] Z. El Mrabet, N. Sugunaraj, P. Ranganathan, S. Abhyankar, "Random forest regressor-based approach for detecting fault location and duration in power systems," Sensors 22 (2022) 458.

[32] M. Amjad, I. Ahmad, M. Ahmad, P. Wróblewski, P. Kamiński, U. Amjad, "Prediction of pile bearing capacity using XGBoost algorithm: modeling and performance evaluation," Applied Sciences 12 (2022) 2126.

[33] W. Li, Y. Yin, X. Quan, H. Zhang, "Gene expression value prediction based on XGBoost algorithm," Front Genet 10 (2019) 1077.

[34] T. Chen, C. Guestrin, "Xgboost: A scalable tree boosting system," in: Proceedings of the 22nd Acm Sigkdd International Conference on Knowledge Discovery and Data Mining, 2016: pp. 785–794.

[35] J. Liang, Y. Bu, K. Tan, J. Pan, Z. Yi, X. Kong, Z. Fan, "Estimation of stellar atmospheric parameters with light gradient boosting machine algorithm and principal component analysis," Astron J 163 (2022) 153.

[36] S. Deng, J. Su, Y. Zhu, Y. Yu, C. Xiao, "Forecasting carbon price trends based on an interpretable light gradient boosting machine and Bayesian optimization," Expert Syst Appl 242 (2024) 122502.

[37] T. Toharudin, R.E. Caraka, I.R. Pratiwi, Y. Kim, P.U. Gio, A.D. Sakti, M. Noh, F.A.L. Nugraha, R.S. Pontoh, T.H. Putri, "Boosting algorithm to handle unbalanced classification of PM 2.5 concentration levels by observing meteorological parameters in jakarta-indonesia using adaboost, xgboost, catboost, and lightgbm", IEEE Access 11 (2023) 35680–35696.

[38] J.T. Hancock, T.M. Khoshgoftaar, "CatBoost for big data: an interdisciplinary review," J Big Data 7 (1): 94, (2020).

[39] S. Li, "Estimating stock market prices with histogram-based gradient boosting regressor: A case study on alphabet inc., International Journal of Advanced Computer Science & Applications 15 (2024).

[40] H. Nhat-Duc, T. Van-Duc, "Comparison of histogram-based gradient boosting classification machine, random Forest, and deep convolutional neural network for pavement raveling severity classification," Autom Constr 148 (2023) 104767.

[41] D. Połap, M. Woźniak, "Red fox optimization algorithm," Expert Syst Appl 166 (2021) 114107.

[42] R. Natarajan, G. Megharaj, A. Marchewka, P.B. Divakarachari, M.R. Hans, "Energy and distance based multi-objective red fox optimization algorithm in wireless sensor network," Sensors 22 (2022) 3761.

[43] A. Jaszcz, D. Połap, R. Damaševičius, "Lung x-ray image segmentation using heuristic red fox optimization algorithm," Sci Program 2022 (2022) 4494139.

[44] D. Chicco, M.J. Warrens, G. Jurman, "The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation," PeerJ Comput Sci 7 (2021) e623.

[45] J. Terven, D.M. Cordova-Esparza, A. Ramirez-Pedraza, E.A. Chavez-Urbiola, "Loss functions and metrics in deep learning," ArXiv Preprint ArXiv:2307.02694 (2023).

[46] Y. Tang, L. Shang, R. Zhang, J. Li, H. Fu, "Hybrid divergence based on mean absolute scaled error for incipient fault detection," Eng Appl Artif Intell 129 (2024) 107662.

[47] P. Karras, N. Mamoulis, "One-pass wavelet synopses for maximum-error metrics," in: Proceedings of the 31st International Conference on Very Large Data Bases, 2005: pp. 421–432.

[48] G.P. Meyer, "An alternative probabilistic interpretation of the huber loss," in: Proceedings of the Ieee/Cvf

Conference on Computer Vision and Pattern Recognition, 2021: pp. 5261–5269.

[49] K. Okoye, S. Hosseini, "Wilcoxon statistics in R: signed-rank test and rank-sum test, in: R programming: statistical data analysis in research," Springer, 2024: pp. 279–303.

[50] S.M. Lundberg, S.-I. Lee, "A unified approach to interpreting model predictions," Adv Neural Inf Process Syst 30 (2017).

[51] Y. Nohara, K. Matsumoto, H. Soejima, N. Nakashima, "Explanation of machine learning models using shapley additive explanation and application for real data in hospital," Comput Methods Programs Biomed 214 (2022) 106584.