

# Hybrid Metaheuristic-Optimized Support Vector Regression for Electricity Consumption Prediction: A Comparative Performance Analysis

Hong JIANG\*, Chen LI\*

\*Department of Basic Education, Shangqiu Institute of Technology, Shangqiu, Henan 476000, China

([ge\\_jianghong2024@163.com](mailto:ge_jianghong2024@163.com), [lichen20190202@163.com](mailto:lichen20190202@163.com).)

‡Corresponding Author; Hong JIANG, [ge\\_jianghong2024@163.com](mailto:ge_jianghong2024@163.com)

*Received: 22.07.2026 Accepted: 04.08.2026*

**Abstract-** The particular applications of metaheuristic algorithms like biogeography-based optimization for predicting electricity usage have always been considered in the literature. As buildings consume a significant part of the world's energy now, accurate models for predicting electricity usage are essential. Hence, this exploration aims to project the electricity usage based on the experimental data and training data using four optimizers: PSO, shuffled frog-leaping algorithm (SFLA), ant colony optimization (ACO), and BBO. Notably, the problem is solved considering three Support Vector Regression (SVR) models, including linear, Gaussian, and polynomial methods. The performance criteria, including Standard Error, MSE, RMSE, MAPE, MAE, Correlation Coefficient, Relative Absolute Error, Normalized Mean Square Error, and Coefficient of Determination, are assessed for the selected optimizers and networks. The results obtained illustrate that BBO achieved the highest prediction accuracy and the best final convergence quality, while ACO required the fewest iterations to reach convergence. Also, the statistical results in the training and experimental data emphasize the acceptable performance of Gaussian and linear networks, respectively.

**Keywords** electricity usage prediction; optimization methods; BBO; Gaussian network; regression method

## 1. Introduction

Nowadays, much attention has been devoted to the issue of the energy sector since it has been the main driver of industry growth during the last century [1], [2], [3], [4]. Energy is the source of all the fuel used to drive the rest of the economy. Energy is mainly sourced in two ways: non-renewables [5], [6] like fossil fuels and renewables such as solar [7], [8]. The rapidly industrializing countries particularly require energy efficiency as a promising approach for managing dramatic demand growth. The system reliability is improved, and supply constraints are facilitated. Thus, the need for energy conservation is substantial to provide the industries' necessities. Conversely, the energy usage projection in residential and commercial entities can be significantly helpful in conserving energy properly [9], [10], [11]. Thus, energy consumption prediction is a very relevant topic and has been widely investigated in published work [12], [13]. In 2017, Li et al. aimed to improve the building energy usage accuracy using an extreme DL method. The extreme learning machine (ELM) was integrated with the stacked autoencoders (SAEs) in this method to make the most of their

characteristics [2]. In the cutting-edge paper of 2019, a new hybrid prediction method through the deep evolutionary learning (DL) approach, in which the genetic algorithm was integrated with long short-term memory, was proposed [14], [15]. The objective function was optimized based on time window lags and the network's hidden neurons in this study. Besides, Luo et al. proposed an AI-oriented method including feature retrieval, evolutionary optimization, and an adaptive deep neural network (DNN) scheme for anticipating energy usage weekly [16].

In fact, with the pace of emerging algorithms, several prediction methodologies already appear in the related works. Stochastic models are trustworthy in forecasting building energy usage and identifying potential domains of energy waste [1]. As mentioned above, genetic programming methods and ML schemes like ANN, SVM, and time series techniques have often been successful in energy consumption prediction with an acceptable level of accuracy [17], [18]. DL schemes (Mocanu et al., 2016) and artificial neural networks [19] were significantly employed for prediction. As an example, Pinto et al. applied three ensemble learning

algorithms (i.e., gradient-boosted regression trees (GBRT), RF, and an adaptation of Adaboost) in 2021 [20]. The results obtained point out the better performance of the adapted Adaboost model compared to the reference schemes for the hour-ahead load forecast. Also, the noisy features of building datasets were eliminated through principal component analysis (PCA) integrated with five anticipation schemes to enhance forecasting performance [21]. These schemes were linear regression, support vector regression, RF, KNN, and regression tree. The feature reduction and prediction effectiveness were then examined using the suggested technique for four case studies in a range of temperature zones, including warm-dry, hot-humid, cold, and mild situations. The findings showed that the intended method could extract an intelligent dataset for energy consumption prediction issues from any large dataset. In a significant advance in 2021, recurrent neural network (RNN) and two variants were examined for enterprise electricity usage prediction [19]. Considering different RNN schemes, including long short-term memory-based RNN (LSTM), standard RNN, and gated recurrent unit-based RNN (GRU), a comparison was made for daily enterprise electricity usage prediction.

The previous decades have witnessed various applications of the regression method for energy usage prediction [22]. In the groundbreaking paper of 2007, Tso and Yau introduced three modeling methods, namely the traditional regression analysis, neural networks, and decision trees (DT), to forecast electrical energy usage [23]. The experimental results indicated that the DT and neural network (NN) schemes successfully understood energy usage patterns and predicted energy usage levels. In 2014, Asadi et al. conducted multiple linear regression analysis to evaluate energy consumption in the early design stages of a building and obtained an acceptable result [24]. Furthermore, to increase the reliability of predicting energy consumption, the gradient boosting regression tree method was employed [25]. The various prediction schemes were compared regarding RMSE. It was finally revealed that the recommended scheme has the lowest values of RMSE compared to others. Despite such remarkable interest, many shortcomings and gaps are still associated with the prediction schemes based on regression methods that need to be overcome.

This exploration aims to project electricity power usage, considering the experimental data presented in the study of Chang et al. in 2018 [26]. The numerical analysis is conducted based on the regression method, and the electrical power consumption is specified numerically as the output. At the same time, Chang et al. obtained the output with a class. In this work, three support vector regression (SVR) NN methods, i.e., linear, Gaussian, and polynomial methods, have been used for analysis. The outcomes of the chosen algorithms are highlighted in descending order of efficiency, according to which the biogeography-based optimization (BBO) algorithm is the best one. Based on a review of the existing literature, these results have not been documented previously. Accordingly, the BBO algorithm is introduced as optimal for forecasting electricity power consumption with the highest accuracy.

Although numerous studies have combined support vector regression with optimization algorithms for energy consumption prediction, most existing research evaluates a single optimization strategy or focuses on improving one specific prediction model, making it difficult to understand the relative effectiveness of different metaheuristic optimizers under identical experimental conditions. Furthermore, limited attention has been given to systematically investigating how different SVR kernel functions respond to diverse optimization mechanisms using the same benchmark dataset and evaluation criteria. To address this gap, the present study establishes a unified comparative framework integrating Linear, Polynomial, and Gaussian SVR models with four representative population-based optimization algorithms, namely PSO, SFLA, ACO, and BBO. These optimizers were selected because they employ different exploration–exploitation strategies and have demonstrated effectiveness across a broad range of engineering optimization problems. Their comparative evaluation provides a comprehensive understanding of optimization behavior, prediction accuracy, convergence characteristics, and model generalization for electricity consumption forecasting.

The rest of this article is outlined below:

## 2. Preliminaries

Due to the recent developments of the meta-heuristic algorithms in previous studies, many researchers have exploited them for the optimization process. The aim is to propose the most efficient algorithm for optimizing the electrical consumption prediction. Meta-heuristic algorithms are applied to heuristic algorithms, making them free of local optimization. The applications of such algorithms for energy or electricity usage prediction have been widely reported in the literature [27]. Accordingly, PSO, SFLA, ACO, and BBO algorithms are considered for solving the problem.

### 2.1. Particle swarm optimization (PSO)

As stated in computational science, PSO is regarded as a computational approach mainly employed to optimize an issue through iteratively attempting to enhance a candidate solution according to a specific measure of quality [28], [29]. Notably, the related works emphasize the remarkable capabilities of PSO in optimizing energy usage projection.

### 2.2. Shuffled frog-leaping algorithm (SFLA)

One of the most effective and practical algorithms for solving optimization problems is SFLA, proposed by Eusuff and Lansey in 2006. It should be noted that SFLA is known as a population-based metaheuristic scheme in which the advantages of memetics are integrated with PSO [30].

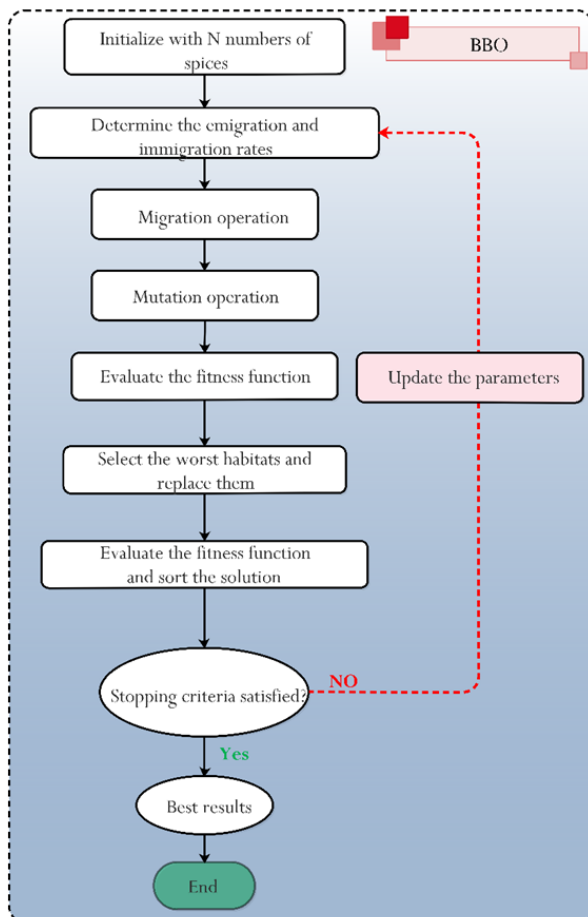
### 2.3. Ant colony optimization algorithm (ACO)

A probabilistic technique, known as ACO, is utilized for problem-solving in computational computer science and operations research [31]. Considering ACO, it is possible to

reduce the computational costs of finding correct paths through graphs. The foundation of this method is the actual behavior of real ants. Based on this, artificial ants are a multi-agent technique. The main concept that is mostly used in this technique is the pheromone-based communication of biological ants. The artificial ants and local search frameworks are combined to create a selection method for a large number of optimization tasks, including some graph-like vehicle routing and internet routing.

#### 2.4. Bio geography-based optimization (BBO)

One of the most well-known evolutionary algorithms (EA) is the BBO algorithm, which improves candidate solutions stochastically and repeatedly under a particular quality or fitness function to maximize a function. A large number of existing explorations in the broader literature have exploited the BBO algorithm to accurately project energy and electricity usage [32]. The general flowchart of this algorithm is shown in Fig. 1, based on which seven steps are required to obtain the best solution.



**Fig. 1.** Flow chart of the BBO algorithm with details

According to the information in the following sections, the BBO algorithm is proposed as the selected algorithm, which is considered for optimization.

### 3. Methodology

#### 3.1. Dataset Description

This study is based on a full-year dataset that was first presented in [31] and is used as a point of reference for the test and analysis. This dataset consists of hourly energy use readings from a smart metering system in a big hypermarket, along with hourly temperature readings (in degrees Celsius) from meteorological sensors of a professional weather institute. The original database, as mentioned in [15], was preprocessed by the authors to fix problems like missing or anomalous values and to ensure clean and reliable input data.

The benchmark dataset employed in this study is the publicly available 12-month hourly electricity consumption dataset released in [33], which originates from the smart-meter measurements reported in [15]. The categorized energy levels introduced in [24] are adopted only to describe different energy consumption categories and are not used as prediction targets in the proposed regression framework. Instead, the continuous electricity consumption values constitute the regression output, while the categorized labels provide supplementary information regarding consumption ranges. Prior to model development, the dataset was preprocessed to remove missing or anomalous records following the benchmark preprocessing protocol reported by the original data providers. Input variables include historical electricity consumption, ambient temperature, and temporal features such as hour of day, day of week, day of month, and month of year. All predictor variables were normalized before model training to ensure numerical stability and fair comparison among the evaluated optimization algorithms. The same training and testing partition was consistently applied to all compared models to ensure an unbiased performance evaluation.

#### 3.2. Energy Level Categorization

The version of the dataset used in this study [31] deviates from its original form [15] through the inclusion of categorized energy levels. These levels were generated by dividing the energy usage data into percentile-based bins. For example, a five-level classification scheme was established using evenly distributed percentiles, resulting in the following energy usage ranges:

Very Low: [0.174, 0.366)

Low: [0.366, 0.634)

Medium: [0.634, 0.782)

High: [0.782, 0.874)

Very High: [0.874, 1.360]

Each energy level was then mapped to an ordinal value ranging from 1 to 5 for use in predictive modeling. The dataset can be classified employing three distinct classification cases, namely, three, five, and seven energy level classes. These variations were used to conduct separate prediction tasks using ML techniques.

It should be emphasized that the proposed methodology performs continuous electricity consumption prediction using support vector regression. The three-, five-, and seven-level energy categories are presented only to describe the alternative categorization schemes available in the benchmark dataset and to provide additional information regarding consumption ranges. These categorical labels are not employed as prediction targets in the proposed regression model. The categorized energy-level labels were not included among the predictor variables during model development and therefore did not introduce target leakage into the regression framework.

To clarify the associations between the database's features, a correlation matrix heatmap was generated, as shown in Fig. 2. The matrix entries are the Pearson correlation coefficient for a given pair of variables, and the values can only range from -1 to 1.

A value of 1 showcases a perfect positive linear relationship.

A value of -1 showcases a perfect negative linear relationship.

A value of 0 showcases no linear correlation.

The heatmap employs a color gradient to improve interpretability:

Dark red signifies high positive correlations.

Dark blue signifies negative or very low correlations.

Intermediate shades indicate moderate correlations, transitioning through cyan and yellow.

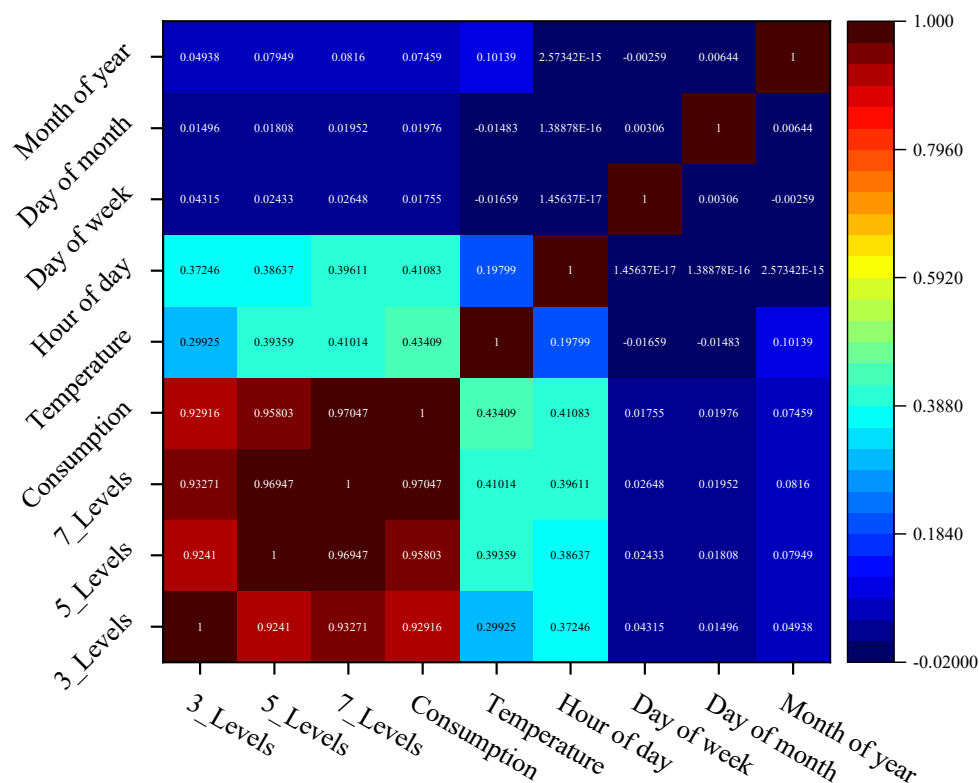
Key Observations:

Strong correlations (above 0.9) are observed among a cluster of variables related to energy usage at different times or possibly lagged versions of energy usage. For instance, values such as 0.97047 and 0.95903 indicate near-perfect correlations, implying redundancy or temporal dependence in those variables.

Temperature seems to show weak to moderate correlations (0.2–0.4 range) with the variables of energy usage, indicating the influence of temperature on energy use, but it is not the only factor.

The diagonal values are 1, as the variables are perfectly correlated with themselves.

This also shows an indication of possible multicollinearity, which is an important feature for consideration when doing predictive modeling, as a very high correlation of independent variables may affect the stability of certain ML schemes, such as linear regression, and dimensionality reduction or feature selection techniques may be required.



**Fig. 2.** Correlation matrix heatmap to visualize the linear relationships between numerical variables in the dataset

Fig. 3 presents a Descriptive Statistical Feature Analysis of the main features in the dataset, summarizing their statistical properties, including minimum, mean, median, maximum, and coefficient of variation (CV). This analysis is

important to understand the distribution and variability of each feature and how it might impact the actual predictive modeling process.

### Description of Metrics:

**Minimum and Maximum:** Represent the range of each feature, identifying the spread of values.

**Mean and Median:** Provide central tendency measures; comparing them helps to determine skewness in the data.

**Coefficient of Variation (CV):** A normalized measure of dispersion (standard deviation divided by the mean), useful for comparing variability across different features regardless of their scale.

### Key Observations:

Features with high CV values (e.g., top bars in yellow) demonstrate significant variability, indicating that they may be more sensitive or influential in prediction tasks. These variables should be analyzed carefully during model training and normalization steps.

For several features, the mean and median values are close, suggesting a relatively symmetric distribution. In contrast, noticeable gaps between mean and median imply skewness, which may necessitate transformation (e.g., log-scaling) for certain schemes.

The range (max-min) of some variables is quite large, which indicates the existence of outliers/ extreme values and might have an impact on the model performance if not handled properly during the data pre-processing.

### Use in Modeling:

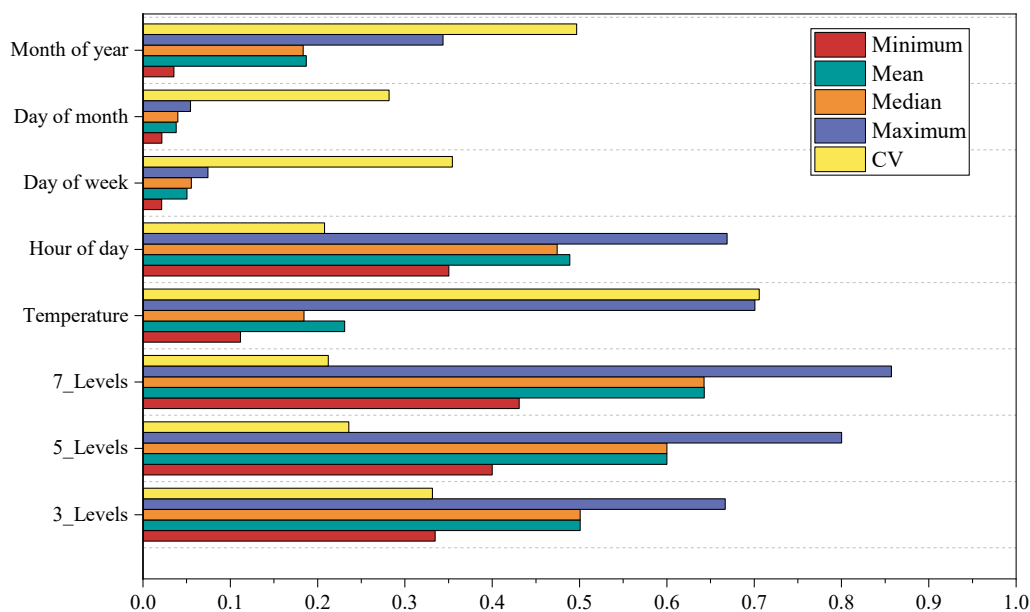
This Descriptive Statistical Feature Analysis suggests:

**Feature selection:** Features with low variation compared to their mean (low CV) may add no real value to the modeling activity and should be dropped.

**Scaling needs:** Features with large ranges or variability may need standardization or normalization.

**Outlier detection:** A significant departure of the median from the distribution bounds can indicate the occurrence of outliers that need adjusting.

The Descriptive Statistical Feature Analysis in Fig. 3 reveals distinct characteristics across the dataset's features. Temperature exhibits relatively low variability, with closely aligned mean and median values, showcasing a stable and symmetric distribution over time. In contrast, energy usage-related features, such as hourly and daily loads, display high maximum values and large coefficients of variation, highlighting significant fluctuations and a tendency toward skewed distributions due to peak demand periods. Lagged or time-shifted consumption features show similar statistical behavior, reflecting strong temporal dependence and reinforcing their importance in predictive modeling. Categorical time-based features, when encoded numerically (e.g., time of day or day type), tend to have lower variability but remain essential for capturing periodic consumption patterns. In general, these features' variability and distribution are important to the performance of the model, the feature selection process, and preprocessing choices. It should be noted that the purpose of this analysis is to summarize the statistical characteristics and variability of the predictor variables rather than to perform a formal global sensitivity or feature-importance analysis. The observed strong correlations among several energy-related variables mainly reflect temporal dependencies within electricity consumption patterns. Since the same predictor set is consistently employed across all evaluated models, these correlations do not affect the fairness of the comparative evaluation. Furthermore, SVR models generally exhibit greater robustness to correlated input variables than conventional linear regression approaches, although future work may investigate feature selection or dimensionality reduction techniques to further reduce potential multicollinearity.



**Fig. 3.** Descriptive Statistical Feature Analysis of dataset features

### 3.3. Models

The numerical analysis is conducted in this research to project electrical power consumption. The dataset used for solving the optimization problem is given in Piedad and Kuo's study in 2018 [24]. In the beginning, three methods of Support Vector Regression (SVR) NN, including linear, Gaussian, and polynomial methods, are trained. Then, the networks' test outputs are specified based on Eq. (1).

$$\text{Forecast} = a_1 * \text{Linear} + a_2 * \text{Gaussian} + a_3 * \text{polynomial} + a_4 \quad (1)$$

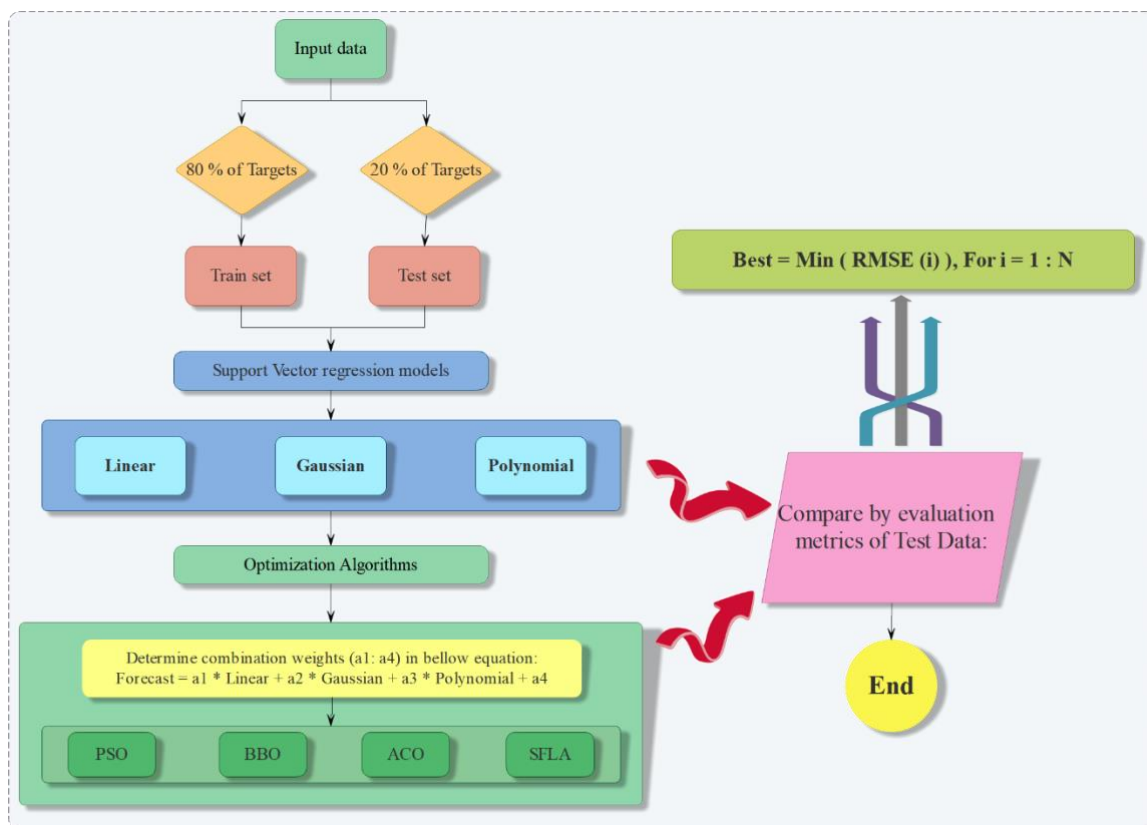
where  $a_1$ – $a_4$  denote the optimized regression coefficients determined by the selected metaheuristic algorithm. During optimization, each candidate solution represents a coefficient vector, and the objective is to minimize the prediction error between the estimated and observed electricity consumption values using the predefined regression performance criterion. The optimized coefficients are subsequently used to generate the final prediction model for each optimization algorithm. With the algorithms selected, i.e., PSO, BBO, SFLA, and ACO, the coefficients  $a_1$  -  $a_4$  are found as shown in Table 1.

**Table 1.** The values of the coefficients  $a_1$  -  $a_4$  based on the selected algorithms

|      | $a_1$       | $a_2$       | $a_3$       | $a_4$        |
|------|-------------|-------------|-------------|--------------|
| PSO  | 9.25465E-06 | 0.9550656   | 0.999969627 | -0.488640886 |
| BBO  | 0.567704591 | 0.358491623 | 0.978428542 | 0.214134671  |
| SFLA | 0.387215377 | 0.261266641 | 0.924953339 | 0.338972788  |
| ACO  | 0.217072977 | 0.561741494 | 1           | 1            |

The procedure of solving the problem and the analysis conducted in the present research are outlined in Fig. 4. To understand the solution process of the considered problem, which leads to the accurate prediction of the electricity usage, paying attention to Fig. 4 is of great importance.

To ensure a fair comparative evaluation, all optimization algorithms were executed under the same computational environment, identical training and testing datasets, consistent preprocessing procedures, and uniform model evaluation criteria. Consequently, the reported performance differences primarily reflect the characteristics of the optimization strategies rather than differences in experimental conditions.



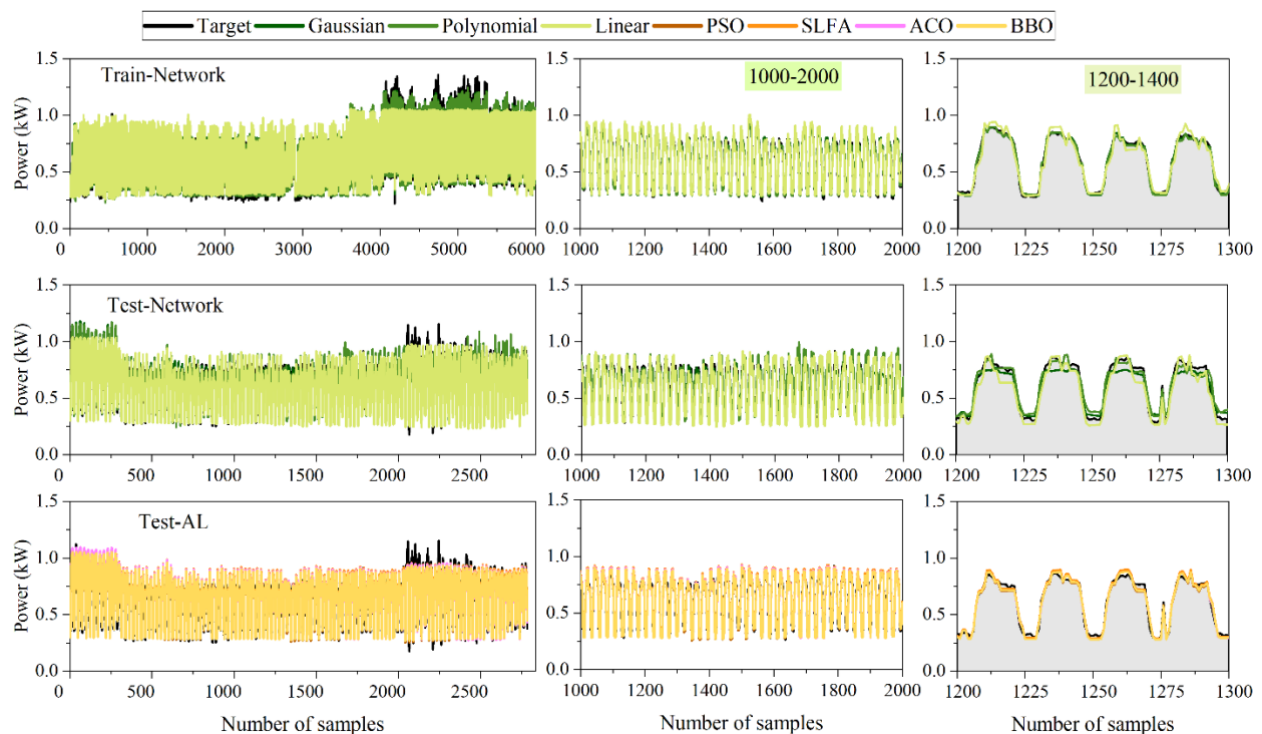
**Fig. 4.** The whole procedure considered for reaching the ultimate results of the research

To comprehensively evaluate prediction performance, multiple statistical indicators are employed. The correlation coefficient ( $R$ ) measures the strength of the linear relationship between predicted and observed values, while the coefficient of determination ( $R^2$ ) quantifies the proportion of variance explained by the prediction model. The normalized mean square error (NMSE) evaluates the prediction error after normalization, thereby facilitating comparisons across different scales. The performance index (PI) provides an overall indication of prediction quality by combining estimation accuracy into a single performance measure. Lower values of NMSE and PI indicate better predictive performance, whereas higher values of  $R$  and  $R^2$  indicate stronger agreement between predicted and observed values.

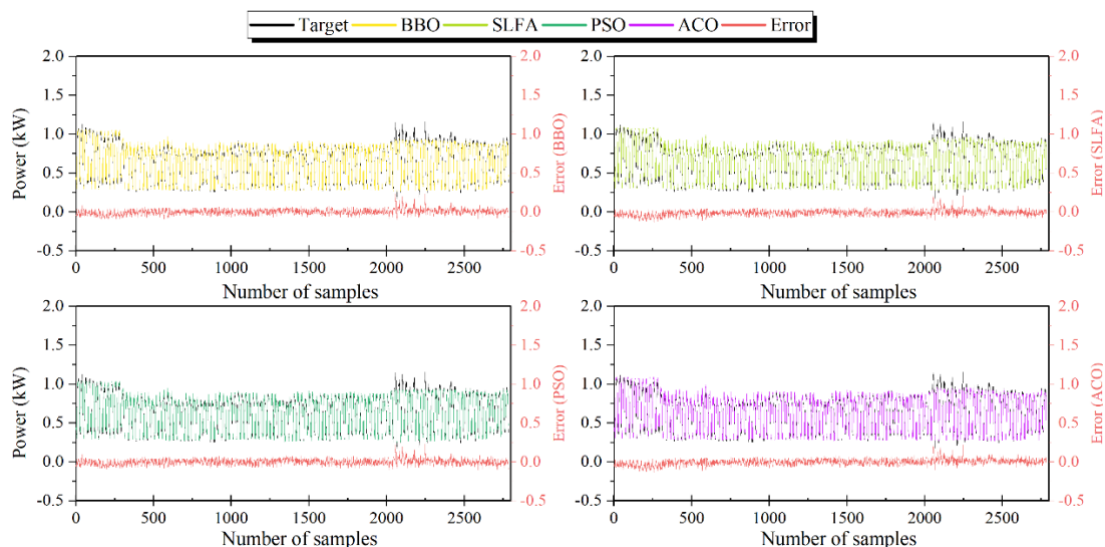
#### 4. Results and Discussion

This part illustrates the main outcomes attained from applying the selected algorithms to the considered problem of this research.

Regarding Figs. 5(a) and 5(b), careful comparison is made between the empirical and obtained data for the various networks and algorithms. Eq. (1) is considered for the selected optimizers, and its parameters are optimized. The results obtained show the best BBO with regard to other optimizers.

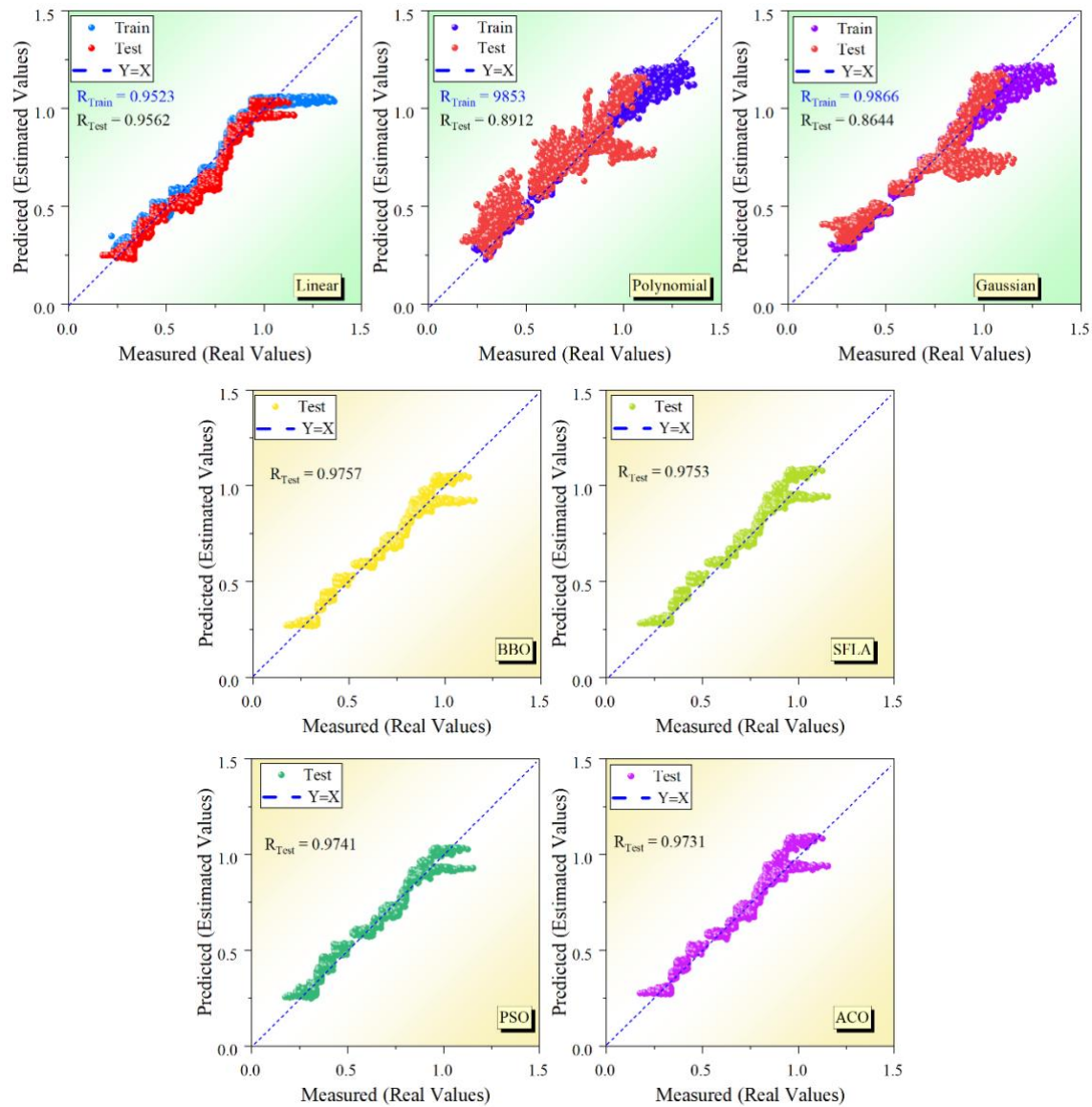


(a)



(b)

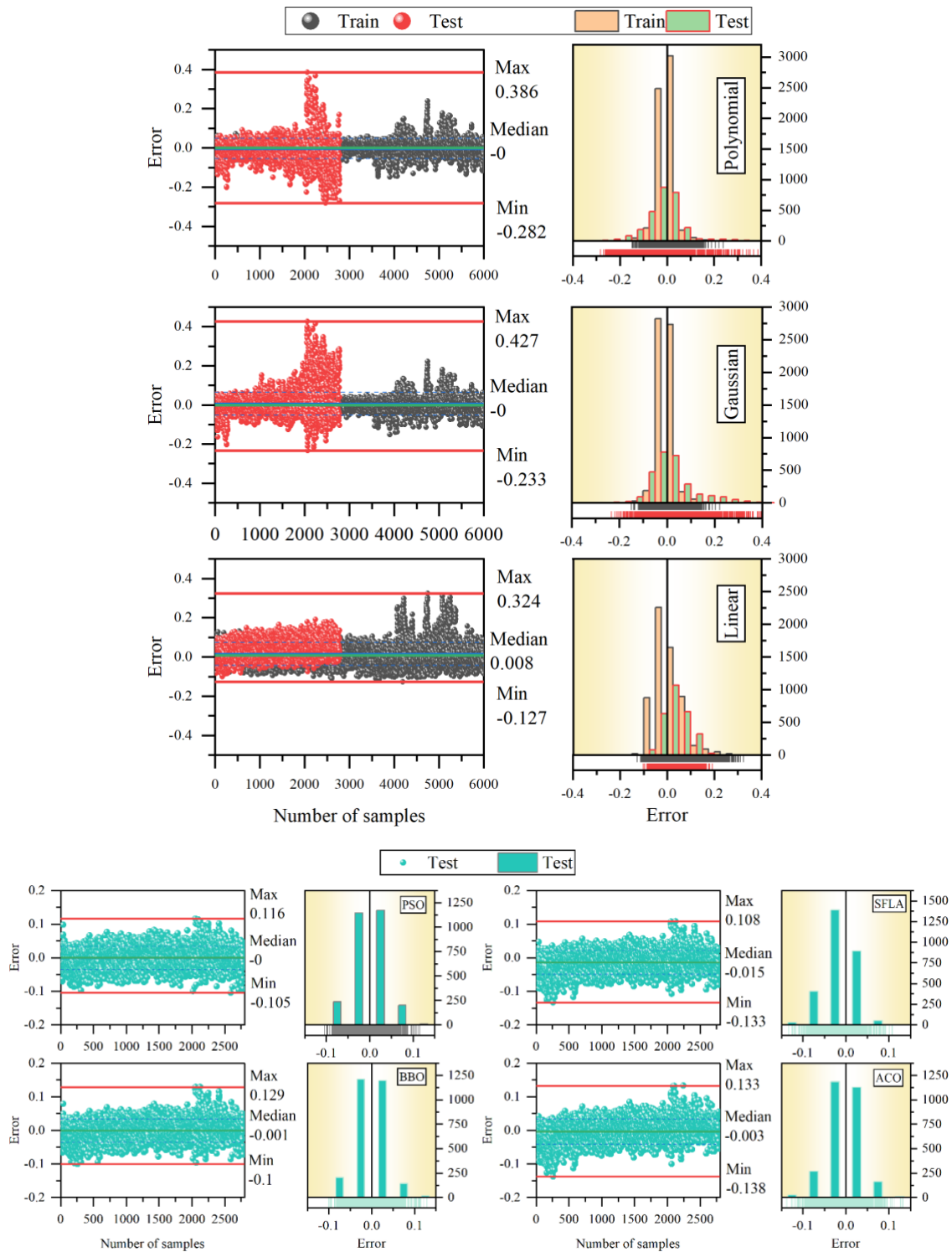
**Fig. 5.** Comparison between obtained data and experimental data for (a) different networks and (b) algorithms



**Fig. 6.** Comparison between the  $R^2$  values of the obtained data and experimental data in the various networks and algorithms

As shown in Fig. 6, the  $R^2$  values of selected optimizers are different in the experimental data. The BBO, SFLA, PSO, and ACO indicate  $R^2$  values of 0.9757, 0.9753, 0.9741, and 0.9731, respectively. Also,  $R^2$  values of linear, polynomial, and Gaussian methods for the training data are 0.9523, 0.9853, 0.9866, 0.9562, 0.8912, and 0.8644 for the experimental data,

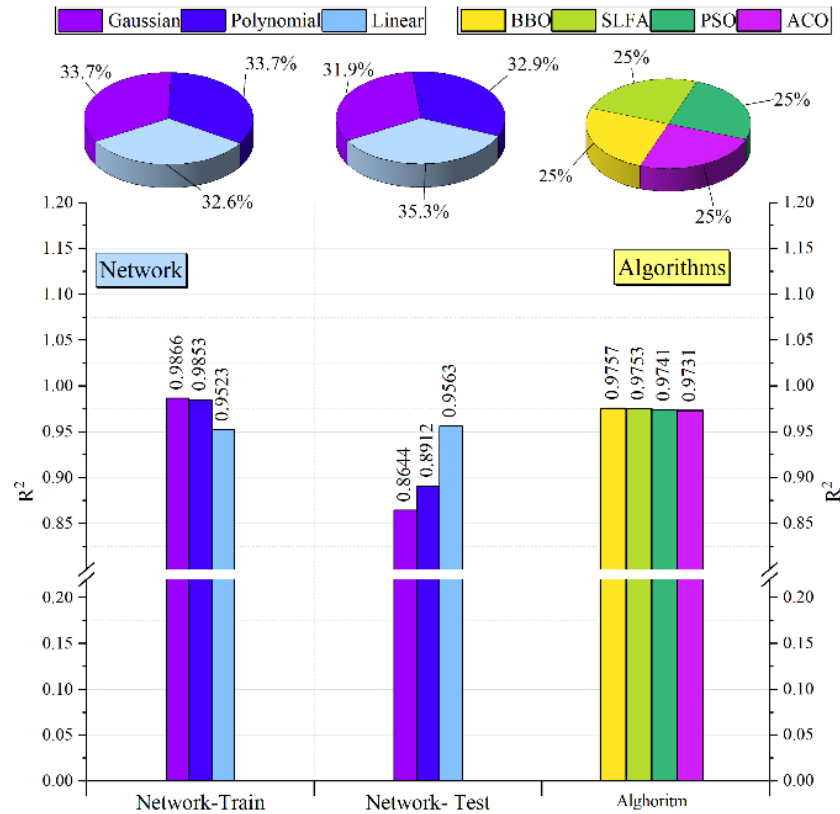
respectively. Accordingly, the BBO algorithm has the best performance regarding the  $R^2$  value. Notably, the Gaussian method is superior to linear and polynomial methods regarding the  $R^2$  value through the training data. In contrast, the linear method is the most efficient solution regarding the  $R^2$  value.



**Fig. 7.** Residual and histogram for different networks and algorithms

Concerning Fig. 7, the performance of the three selected networks and algorithms is examined, derived from the residuals and histograms. Accordingly, the performance of

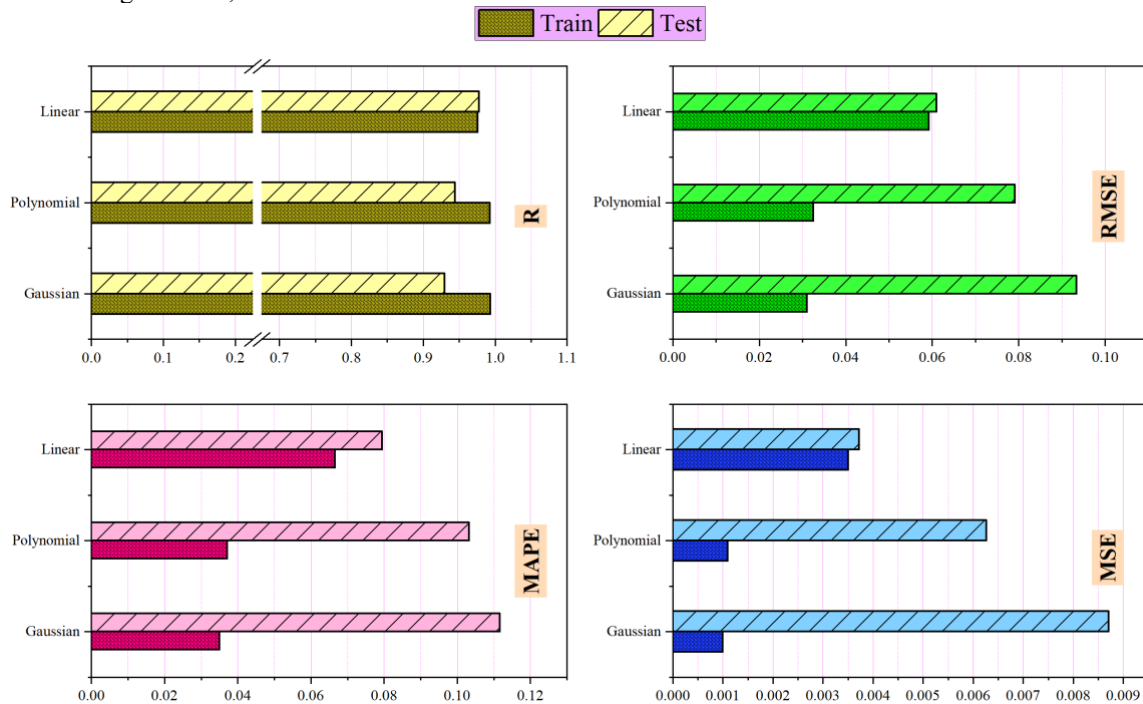
BBO is considered acceptable. It should be noted that the range of errors in the linear network is lower than the rest.

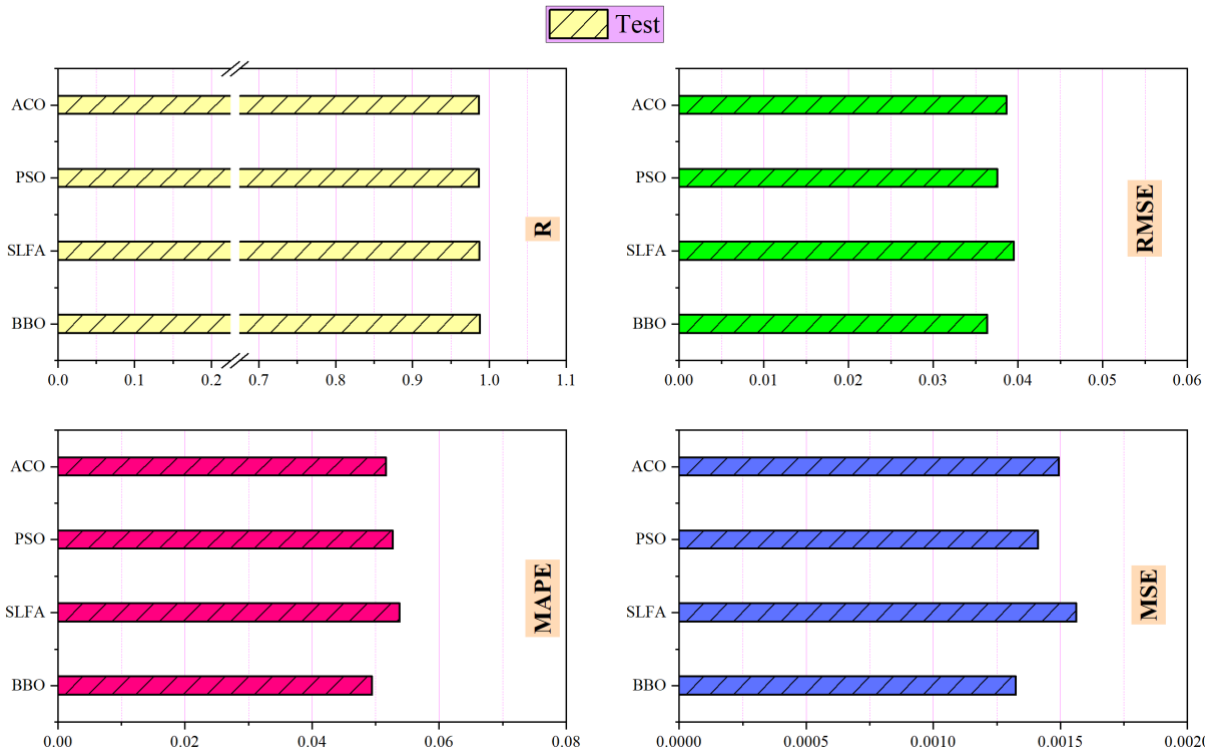


**Fig. 8.** Obtained  $R^2$  values for the recommended and other well-known optimizations.

The  $R^2$  values of the recommended method and other optimizers indicate that the BBO has the best performance. The results shown in Fig. 8 are in good agreement with the ones presented in Fig. 7. Thus, Gaussian and linear networks

have an acceptable level of performance according to the training and experimental data, respectively.

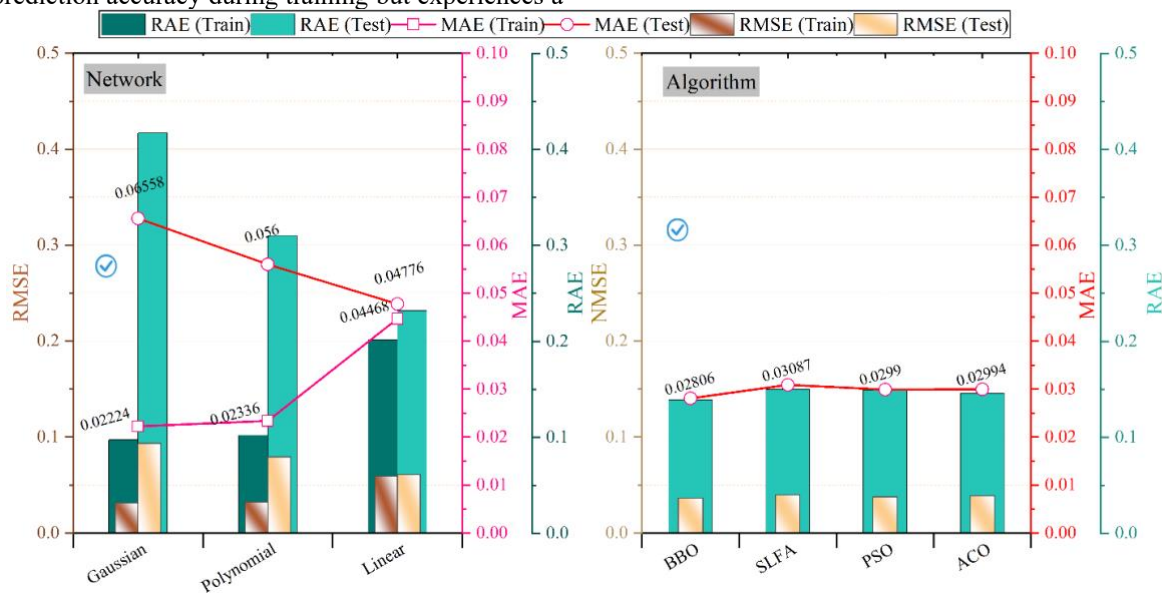




**Fig. 9.** Attained R, RMSE, MAPE, and MSE values for recommended and other well-known networks and algorithms.

In addition, the performance criteria include SE, RMSE, MSE, MAE, MAPE, R, RAE, NMSE, and  $R^2$  are assessed for each optimizer. As shown in Fig. 9, the BBO has the lowest RMSE, MSE, and MAPE and the highest R-value compared to other selected optimizers. In contrast, the SLFA with the highest RMSE, MSE, and MAPE values is considered the weakest optimizer. Furthermore, the results of RMSE, MSE, MAPE, and R criteria for the training data highlight that the best network among those selected is the Gaussian one. For experimental data, the best network based on the values of RMSE, MSE, MAPE, and R is the linear network. An interesting observation is that the Gaussian SVR achieves the highest prediction accuracy during training but experiences a

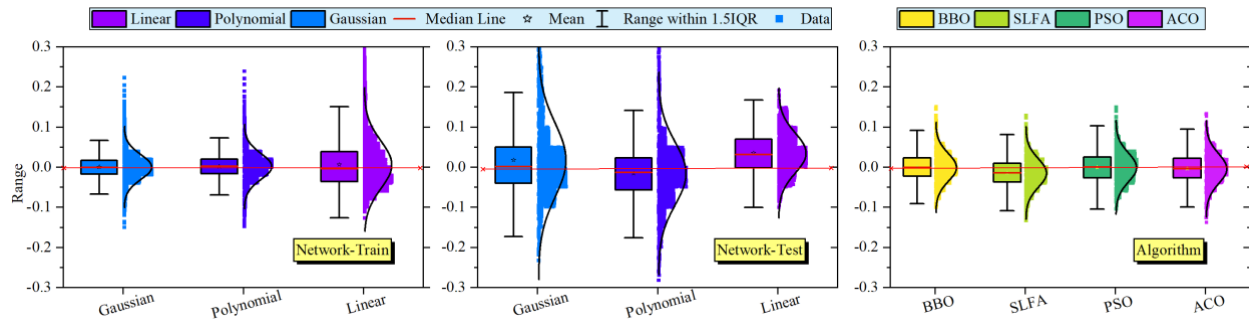
noticeable decline in testing performance compared with the Linear SVR. This behavior suggests that the greater nonlinear flexibility of the Gaussian kernel enables it to model complex relationships within the training data effectively but may also increase susceptibility to overfitting under the current dataset. In contrast, the Linear SVR employs a simpler hypothesis space that sacrifices a small amount of training accuracy while providing stronger generalization capability on previously unseen observations. This trade-off between model complexity and generalization explains why the Linear SVR achieves the highest testing performance despite the superior training accuracy of the Gaussian model.



**Fig. 10.** Obtained RMSE, RAE, and MAE values for recommended and other well-known networks and algorithms.

According to Fig. 10, another analysis is conducted that matches the obtained results demonstrated in Figs. 3-8. Hence, the selected optimizers' and networks' RMSE, RAE, and

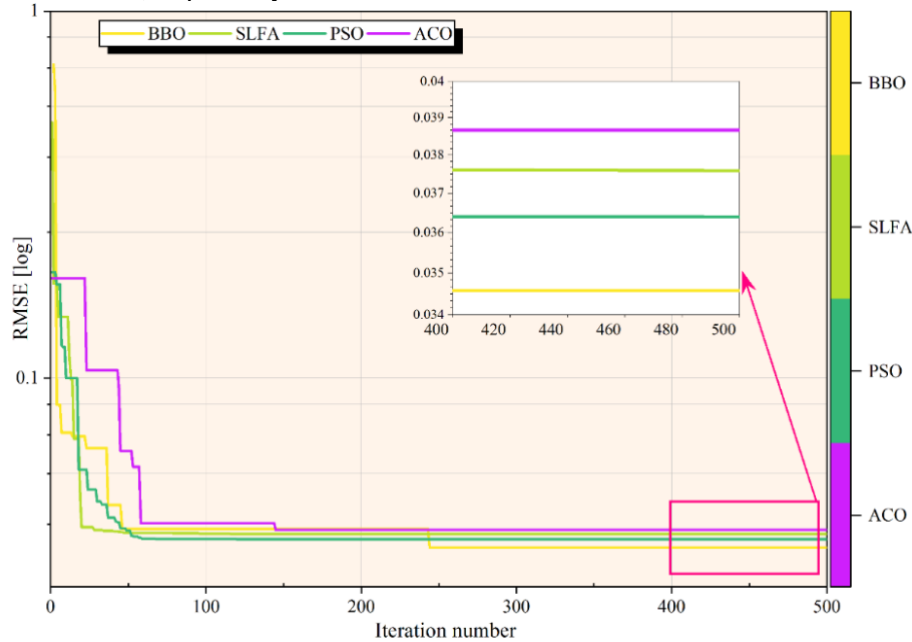
MAE are indicated. The outcomes represent the remarkable performance of the Gaussian network and BBO in predicting the amount of electricity usage.



**Fig. 11.** Obtained error values for recommended and other well-known networks and algorithms.

Fig. 11 displays that the error values of the chosen networks and optimizers are highlighted in descending order. The Gaussian and linear networks have the smallest error in the training and experimental data, respectively. In relation to

the optimizers, the least effective in terms of the error value is the one against ACO, while the BBO has provided the best accuracy.



**Fig. 12.** The recommended method's convergence curves along with other methods.

The convergence curves of the recommended approach through RMSE and iteration number are shown in Fig. 12. Accordingly, the quickest convergence belongs to ACO with the value of 0.03866 and 145 iterations. Besides, the least amount of convergence is related to BBO, with the value of 0.0345 based on 244 iterations.

Overall, the analytical results of the considered problem using the selected optimizers and networks are represented in Table 2. Accordingly, the values of MSE, RMSE, MAPE,

MAE, RAE, R,  $R^2$ , NMSE, and PI for the BBO, which is the selected optimizer, are 0.0013, 0.0364, 0.0495, 0.0281, 0.1390, 0.9878, 0.9757, 0.0043, and 0.0284, respectively. BBO has the highest values of R and  $R^2$  and the lowest values of other evaluation criteria. In contrast, the weakest optimizer is ACO, with the lowest R and  $R^2$  and the highest MSE, RMSE, MAPE, MAE, RAE, NMSE, and PI values. The same analysis has been conducted for the selected networks, revealing that Gaussian and linear networks have acceptable performance in the training and experimental data.

**Table 2.** Obtained statistical results for diverse optimizers.

| Selected optimizers |
|---------------------|
|---------------------|

| Optimizer         | MSE    | RMSE   | MAPE   | MAE    | RAE    | R      | R <sup>2</sup> | NMSE   | PI     |
|-------------------|--------|--------|--------|--------|--------|--------|----------------|--------|--------|
|                   | Test   |        |        |        |        |        |                |        |        |
| BBO               | 0.0013 | 0.0364 | 0.0495 | 0.0281 | 0.1390 | 0.9878 | 0.9757         | 0.0043 | 0.0284 |
| SFLA              | 0.0016 | 0.0395 | 0.0538 | 0.0309 | 0.1501 | 0.9875 | 0.9753         | 0.0047 | 0.0308 |
| PSO               | 0.0014 | 0.0376 | 0.0527 | 0.0299 | 0.1492 | 0.9869 | 0.9741         | 0.0047 | 0.0293 |
| ACO               | 0.0015 | 0.0387 | 0.0517 | 0.0299 | 0.1458 | 0.9865 | 0.9731         | 0.0044 | 0.0301 |
| Selected networks |        |        |        |        |        |        |                |        |        |
| Optimizer         | MSE    | RMSE   | MAPE   | MAE    | RAE    | R      | R <sup>2</sup> | NMSE   | PI     |
|                   | Train  |        |        |        |        |        |                |        |        |
| Gaussian          | 0.0010 | 0.0311 | 0.0351 | 0.0222 | 0.0974 | 0.9933 | 0.9866         | 0.0021 | 0.0228 |
| Polynomial        | 0.0011 | 0.0325 | 0.0372 | 0.0234 | 0.1021 | 0.9926 | 0.9853         | 0.0024 | 0.0239 |
| Linear            | 0.0035 | 0.0593 | 0.0666 | 0.0447 | 0.2014 | 0.9759 | 0.9523         | 0.0066 | 0.0439 |
|                   | Test   |        |        |        |        |        |                |        |        |
|                   |        |        |        |        |        |        |                |        |        |
| Gaussian          | 0.0087 | 0.0933 | 0.1117 | 0.0656 | 0.4170 | 0.9297 | 0.8644         | 0.0226 | 0.0749 |
| Polynomial        | 0.0063 | 0.0791 | 0.1032 | 0.0560 | 0.3100 | 0.9440 | 0.8912         | 0.0195 | 0.0630 |
| Linear            | 0.0037 | 0.0610 | 0.0794 | 0.0478 | 0.2321 | 0.9779 | 0.9562         | 0.0117 | 0.0477 |

Although the comparative evaluation demonstrates the favorable performance of the BBO optimizer under the adopted experimental protocol, the present study is based on a single training–testing partition applied consistently across all compared methods. Consequently, the reported results should be interpreted within the scope of this evaluation framework. Future studies may further strengthen the statistical reliability of the conclusions by incorporating rolling-window validation, time-series cross-validation, repeated experimental trials, confidence interval estimation, and statistical significance analysis to investigate the robustness of optimization algorithms under different temporal data distributions.

## 5. Conclusion

In summary, this article seeks to project electricity usage based on the regression methods using three networks, namely Gaussian, polynomial, and linear networks, and four optimizers, including the BBO, SFLA, PSO, and ACO. To assess the effectiveness of the chosen optimizers and networks, the evaluation criteria, namely, MSE, RMSE, MAPE, MAE, RAE, R, R<sup>2</sup>, NMSE, and PI, were investigated. Based on the results achieved, the one with the best R and R<sup>2</sup> is the BBO, which would suggest that this optimizer is the most precise of the remaining chosen optimizers. Conversely, the low values of MSE, RMSE, MAPE, MAE, RAE, NMSE, and PI obtained for BBO confirm the acceptable performance of this optimizer. The analytical results of the evaluation criteria for the selected networks indicate the acceptable performance of Gaussian and linear networks based on the training and experimental data, respectively. Among the evaluated optimizers, BBO achieved the highest prediction

accuracy and the best final optimization performance, whereas ACO exhibited the fastest convergence in terms of iteration count. Besides, the Gaussian networks can be introduced as a good network for solving the considered method. Future research is required to confirm the findings that may be made from this study.

## Competing of interests

The authors declare no competing of interests.

## Authorship Contribution Statement

**Hong JIANG:** Writing-Original draft preparation, Conceptualization, Supervision, Project administration.

**Chen LI:** Methodology, Software, Validation.

## Data Availability

On Request

## Declarations

Not applicable

## Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

## Author Statement

The manuscript has been read and approved by all the authors, the requirements for authorship, as stated earlier in this document, have been met, and each author believes that the manuscript represents honest work.

### Funding

Not applicable

### Ethical approval

All authors have been personally and actively involved in substantial work leading to the paper, and will take public responsibility for its content.

### References

- [1] K. Amasyali and N. M. El-Gohary, "A review of data-driven building energy consumption prediction studies," *Renewable and Sustainable Energy Reviews*, vol. 81, pp. 1192–1205, 2018.
- [2] C. Li, Z. Ding, D. Zhao, J. Yi, and G. Zhang, "Building energy consumption prediction: An extreme deep learning approach," *Energies (Basel)*, vol. 10, no. 10, p. 1525, 2017.
- [3] H. Zhao and F. Magoulès, "A review on the prediction of building energy consumption," *Renewable and Sustainable Energy Reviews*, vol. 16, no. 6, pp. 3586–3592, 2012.
- [4] P. Lajevardy, F.-A. Parand, H. Rashidi, and H. Rahimi, "A hybrid method for load forecasting in smart grid based on neural networks and cuckoo search optimization approach," *International Journal of Renewable Energy Resources*, vol. 5, no. 1, pp. 13–20, 2015.
- [5] A. Borriello, P. F. Burke, and J. M. Rose, "If one goes up, another must come down: a latent class hybrid choice modelling approach for understanding electricity mix preferences among renewables and non-renewables," *Energy Policy*, vol. 159, p. 112611, 2021.
- [6] H. Yousaf, A. Amin, A. Baloch, and M. Akbar, "Investigating household sector's non-renewables, biomass energy consumption and carbon emissions for Pakistan," *Environmental Science and Pollution Research*, vol. 28, no. 30, pp. 40824–40834, 2021.
- [7] P. Ekins, "Step changes for decarbonising the energy system: research needs for renewables, energy efficiency and nuclear power," *Energy Policy*, vol. 32, no. 17, pp. 1891–1904, 2004.
- [8] N. A.-H. K. Hussein and O. M. Basel, "Integrating Renewable Energy Systems into Urban Planning for Sustainable Cities," *ESTIDAMAA*, vol. 2024, pp. 15–21, Feb. 2024, doi: 10.70470/ESTIDAMAA/2024/003.
- [9] L. Lei, W. Chen, B. Wu, C. Chen, and W. Liu, "A building energy consumption prediction model based on rough set theory and deep learning algorithms," *Energy Build.*, vol. 240, p. 110886, 2021.
- [10] Z. Wu and W. Chu, "Sampling strategy analysis of machine learning models for energy consumption prediction," in *2021 IEEE 9th International Conference on Smart Energy Grid Engineering (SEGE)*, IEEE, 2021, pp. 77–81.
- [11] M. H. M. N. M. L. N. H. Bakiri Hussein Abubakar, "Towards the Establishment of a Robust Load Forecasting Mechanism in Tanzania Grid: A Study to Investigate Effect of Air Temperature and Daytime on Electricity Consumption in Residential Buildings," *International Journal of Smart grid*, Mar. 2021, doi: 10.20508/ijsmartgrid.v5i1.159.g140.
- [12] I. Karijadi and S.-Y. Chou, "A hybrid RF-LSTM based on CEEMDAN for improving the accuracy of building energy consumption prediction," *Energy Build.*, vol. 259, p. 111908, 2022.
- [13] Q. Qiao, A. Yunusa-Kaltungo, and R. E. Edwards, "Towards developing a systematic knowledge trend for building energy consumption prediction," *Journal of Building Engineering*, vol. 35, p. 101967, 2021.
- [14] A. Almalaq and J. J. Zhang, "Evolutionary deep learning-based energy consumption prediction for buildings," *IEEE Access*, vol. 7, pp. 1520–1531, 2018.
- [15] M. Bairami, H. Khajavi, and A. Rastgoo, "Assessing groundwater behavior and future trends in the Ardabil Aquifer: A comparative study of groundwater modeling system and categorical gradient boosting hybrid model," *Expert Syst. Appl.*, vol. 255, p. 124728, 2024.
- [16] X. J. Luo, L. O. Oyedele, A. O. Ajayi, O. O. Akinade, H. A. Owolabi, and A. Ahmed, "Feature extraction and genetic algorithm enhanced adaptive deep neural network for energy consumption prediction in

- buildings,” *Renewable and Sustainable Energy Reviews*, vol. 131, p. 109980, 2020.
- [17] A. Pirjan, S.-V. Oprea, G. Căruțașu, D.-M. Petroșanu, A. Bâra, and C. Coculescu, “Devising hourly forecasting solutions regarding electricity consumption in the case of commercial center type consumers,” *Energies (Basel)*, vol. 10, no. 11, p. 1727, 2017.
- [18] J. Virote and R. Neves-Silva, “Stochastic models for building energy prediction based on occupant behavior assessment,” *Energy Build.*, vol. 53, pp. 183–193, 2012.
- [19] D. C. Park, M. A. El-Sharkawi, R. J. Marks, L. E. Atlas, and M. J. Damberg, “Electric load forecasting using an artificial neural network,” *IEEE transactions on Power Systems*, vol. 6, no. 2, pp. 442–449, 2002.
- [20] T. Pinto, I. Praça, Z. Vale, and J. Silva, “Ensemble learning for electricity consumption forecasting in office buildings,” *Neurocomputing*, vol. 423, pp. 747–755, 2021.
- [21] T. Parhizkar, E. Rafieipour, and A. Parhizkar, “Evaluation and improvement of energy consumption prediction models using principal component analysis based feature reduction,” *J. Clean. Prod.*, vol. 279, p. 123866, 2021.
- [22] H. Zhong, J. Wang, H. Jia, Y. Mu, and S. Lv, “Vector field-based support vector regression for building energy consumption prediction,” *Appl. Energy*, vol. 242, pp. 403–414, 2019.
- [23] G. K. F. Tso and K. K. W. Yau, “Predicting electricity energy consumption: A comparison of regression analysis, decision tree and neural networks,” *Energy*, vol. 32, no. 9, pp. 1761–1768, 2007.
- [24] S. Asadi, S. S. Amiri, and M. Mottahedi, “On the development of multi-linear regression analysis to assess energy consumption in the early stages of building design,” *Energy Build.*, vol. 85, pp. 246–255, 2014.
- [25] P. Nie, M. Roccotelli, M. P. Fanti, Z. Ming, and Z. Li, “Prediction of home energy consumption based on gradient boosting regression tree,” *Energy Reports*, vol. 7, pp. 1246–1255, 2021.
- [26] H.-C. Chang, C.-C. Kuo, Y.-T. Chen, W.-B. Wu, and E. J. Piedad, “Energy consumption level prediction based on classification approach with machine learning technique,” in *4th World Congress on New Technologies*, 2018, pp. 1–8.
- [27] J.-S. Pan, P. Hu, and S.-C. Chu, “Novel parallel heterogeneous meta-heuristic and its communication strategies for the prediction of wind power,” *Processes*, vol. 7, no. 11, p. 845, 2019.
- [28] S. Malik and D. Kim, “Prediction-learning algorithm for efficient energy consumption in smart buildings based on particle regeneration and velocity boost in particle swarm optimization neural networks,” *Energies (Basel)*, vol. 11, no. 5, p. 1289, 2018.
- [29] H. Khajavi, A. Rastgoo, and F. Masoumi, “Enhanced Streamflow Forecasting for Crisis Management Based on Hybrid Extreme Gradient Boosting Model,” *Iranian Journal of Science and Technology, Transactions of Civil Engineering*, pp. 1–22, 2025.
- [30] S. Dai, D. Niu, and Y. Li, “Forecasting of energy consumption in China based on ensemble empirical mode decomposition and least squares support vector machine optimized by improved shuffled frog leaping algorithm,” *Applied Sciences*, vol. 8, no. 5, p. 678, 2018.
- [31] K. Guleria and A. K. Verma, “Meta-heuristic ant colony optimization based unequal clustering for wireless sensor network,” *Wirel. Pers. Commun.*, vol. 105, pp. 891–911, 2019.
- [32] D. Nanjundappan, “Hybrid weighted probabilistic neural network and biogeography based optimization for dynamic economic dispatch of integrated multiple-fuel and wind power plants,” *International Journal of Electrical Power & Energy Systems*, vol. 77, pp. 385–394, 2016.
- [33] E. J. Piedad and C. Kuo, “A 12-month data of hourly energy consumption levels from a commercial-type consumer,” *Mendeley Data*, vol. 1, 2018.