Optimized Time-Series Forecasting for Electricity Consumption in Tetouan: A Machine Learning Approach with Greylag Goose Optimization

Marwa M. Eid[1,2,*]

[1]Faculty of Artificial Intelligence, Delta University for Science and Technology, Mansoura, Egypt
[2]Jadara Research Center, Jadara University, Irbid 21110, Jordan

Email: mmm@ieee.org

Abstract

This paper addresses the challenge of predicting and analyzing electricity consumption patterns in Tetouan, Morocco, using time-series data. The dataset consists of 52,416 observations with 9 features, collected from the SCADA system of electricity consumption across three zones. The primary goal is to enhance forecasting accuracy and optimize prediction models through machine learning (ML) algorithms, including both time-series models and advanced optimization techniques. We compare the performance of several baseline ML models, such as BiLSTM and Continuous Time Stochastic Modelling (CTSM), with their optimized versions, utilizing optimization algorithms like Greylag Goose Optimization (GGO), Bat Algorithm (BA), and Whale Optimization Algorithm (WOA). The results show that the optimized CTSM model, using GGO, achieved substantial improvements, including the lowest Mean Squared Error (MSE) of 7.09E-07 and the highest R^2 of 0.990, demonstrating superior accuracy and stability. The contributions of this work include (i) benchmarking various ML models for time-series forecasting, (ii) introducing the use of optimized CTSM with meta-heuristics, and (iii) evaluating model performance using a comprehensive set of statistical metrics.

Keywords: Electricity Consumption Forecasting; Time-Series Prediction; Machine Learning; Greylag Goose Optimization; Energy Management

1 Introduction

Accurate forecasting of electricity consumption is pivotal for the efficient operation and planning of energy systems. It enables utilities to balance supply and demand, optimize grid operations, and implement effective energy policies. In regions like Tetouan, Morocco, where energy consumption is increasing due to population growth and urbanization, precise demand forecasting becomes even more critical [1].

Tetouan, located in northern Morocco, has a unique energy consumption profile influenced by its Mediterranean climate, economic activities, and demographic trends [2]. The city's electricity demand exhibits significant variability, with peak consumption periods during hot summers and cold winters. This variability poses challenges for grid operators in maintaining stability and ensuring a reliable power supply.

The complexity of electricity consumption patterns arises from various factors, including weather conditions, economic activities, and social behaviors [3]. For instance, temperature fluctuations can lead to increased use of heating or cooling systems, while economic activities such as industrial operations or tourism can cause sudden spikes in demand. These factors introduce noise and non-linearity into the time-series data, making accurate forecasting a challenging task

DOI: https://doi.org/10.54216/FPA.210217

Received: April 08, 2025 Revised: June 21, 2025 Accepted: August 20, 2025

Traditional forecasting methods, such as ARIMA and SARIMA, have been widely used for time-series prediction. However, these models often struggle to capture complex, nonlinear relationships inherent in electricity consumption data [4]. Recent advancements in machine learning (ML) have introduced more robust models capable of handling such complexities. Techniques like BiLSTM, CTSM, and ETN have shown promise in forecasting electricity demand by learning intricate patterns from historical data [5], [6].

Despite the advancements in ML models, optimizing their performance remains a significant challenge, especially in regions with limited data [7], [8]. Optimization algorithms, such as the Greylag Goose Optimization (GGO), Bat Algorithm (BA), and Whale Optimization Algorithm (WOA), have been employed to fine-tune model parameters, enhancing their predictive accuracy. These algorithms mimic natural processes to explore and exploit the solution space effectively [9], [10], [11].

This paper aims to address the challenges in electricity consumption forecasting by benchmarking state-of-the-art ML models and applying optimization techniques to improve their performance. Specifically, we focus on the CTSM model and enhance its accuracy using GGO, BA, and WOA. Additionally, we introduce new performance metrics tailored for time-series energy consumption models, providing a comprehensive evaluation framework for forecasting models.

The contributions of this study are as follows:

1. A comprehensive benchmarking of ML models, including BiLSTM, CTSM, ETN, and SRRN, for electricity consumption forecasting.
2. The application of optimization algorithms (GGO, BA, WOA) to enhance the performance of the CTSM model.
3. The introduction of new performance metrics, such as MSE, RMSE, MAE, MBE, R^2, RRMSE, NSE, and WI, tailored for evaluating time-series forecasting models.

Through these contributions, we aim to provide valuable insights into the effectiveness of ML models and optimization techniques in forecasting electricity consumption, particularly in regions with limited data availability.

2 Literature Review

The forecasting and optimization of electricity consumption have gained significant attention due to the increasing demand for energy in modern societies. Various approaches employing machine learning, deep learning, and metaheuristic optimization have been proposed to improve prediction accuracy and system reliability. Deep learning techniques, particularly Long Short-Term Memory (LSTM) networks, have demonstrated superior performance in handling time-series energy consumption data. One study introduced a deep LSTM network optimized using a Coronavirus Optimization Algorithm (CVOA) and random search for hyperparameter tuning, achieving a prediction error below 1.5% on a large dataset of Spanish electricity demand [12]. Similarly, another work proposed a hybrid LSTM model integrated with a Butterfly Optimization Algorithm (BOA) to predict electric energy consumption patterns using datasets like IHEPC and AEP, demonstrating minimal error rates compared to other existing models [13]. In addition, LSTM-based frameworks have been applied to forecast renewable energy sources, particularly wind speed and solar irradiance, addressing vanishing gradient challenges and ensuring microgrid stability with a PI controller [14]. A complementary deep learning approach combined Convolutional Neural Networks (CNN), Bidirectional LSTM (BiLSTM), and attention mechanisms, optimized using Particle Swarm Optimization (PSO), for solar irradiation forecasting. This architecture leveraged clustering for day-type categorization, yielding highly accurate predictions with an R^2 score of 0.984, demonstrating significant improvement over conventional forecasting models [15]. These studies emphasize the ability of deep learning techniques to capture nonlinear patterns and temporal dependencies in renewable energy systems, making them highly suitable for accurate multi-step forecasting tasks. Beyond deep learning, hybrid forecasting models integrating statistical and machine learning approaches have proven effective. For instance, the SAMFOR model combined Support Vector Regression (SVR), Firefly Algorithm (FA), and Seasonal Autoregressive Integrated Moving Average (SARIMA) for one-day-ahead building

Received: April 08, 2025 Revised: June 21, 2025 Accepted: August 20, 2025

energy prediction, achieving the lowest RMSE and MAPE metrics among baseline methods [16]. Similarly, a hybrid system optimized linear SARIMA models with Least Squares Support Vector Regression (LSSVR) and Jellyfish Search (JS) optimization to create a highly efficient and accurate multi-step forecasting model for regional electricity consumption [17]. Ensemble models also present substantial promise; an Extreme Learning Machine (ELM) ensemble optimized via PSO showed strong performance on both benchmark datasets and real Spanish electric consumption data, highlighting PSO's potential in optimizing ensemble strategies for time-series prediction [18]. Energy forecasting methods are complemented by advances in energy quality improvement. A proactive shunt power filter with a fuzzy logic controller was proposed for PV-wind hybrid systems, paired with a neural network-based adaptive forecasting model to ensure voltage stability and improve prediction accuracy [19]. Furthermore, fault detection has also seen notable progress. The Hierarchical Deep Learning Approach (HDLA) employed transformer-based models for fault detection, classification, and location prediction in power transmission systems. By eliminating manual feature engineering, HDLA achieved superior accuracy, recall, precision, and computational efficiency over traditional methods, making it suitable for real-time deployment [20]. Optimization algorithms have also been widely applied to building energy management and design. Simulation-based optimization (SBO) frameworks have been explored to balance thermal comfort and energy consumption in HVAC systems. Among algorithms tested, RBFMOpt outperformed NSGA2 and MHACO, especially in scenarios with limited simulation budgets, providing robust optimization for sustainable and energy-efficient building solutions [21]. These optimization-driven methods emphasize the potential of AI and heuristic algorithms in guiding decisions for energy conservation, sustainability, and smart-grid operations. Overall, the reviewed literature highlights a trend towards integrating deep learning, hybrid machine learning models, and optimization algorithms to address energy forecasting and management challenges. LSTM-based architectures dominate renewable energy prediction, while optimization techniques such as PSO, BOA, and JS enhance hyperparameter tuning and improve accuracy. Applications span diverse areas, including renewable energy integration, building energy management, fault detection, and smart grid operation, reflecting the interdisciplinary and rapidly evolving nature of this field.

3 Dataset and Preprocessing

The dataset used in this study consists of electricity consumption data collected from the city of Tetouan, Morocco, over a period of time. It comprises a total of 52,416 observations, each described by nine features, capturing key aspects of energy consumption and environmental conditions. These features include time-related variables such as the date and time of the observation, as well as environmental factors like temperature, humidity, wind speed, and diffuse flow measurements. Additionally, power consumption data for three distinct zones in Tetouan—Zone 1, Zone 2, and Zone 3—are recorded at regular intervals.

3.1 Dataset Overview

The data is provided as a time-series with observations made at 10-minute intervals. This temporal granularity is crucial for understanding short-term fluctuations in electricity demand, which are influenced by a variety of external factors, including weather conditions and consumption habits. The dataset features the following nine variables:

1. **Datetime**: The time at which the observation was recorded.
2. **Temperature**: The ambient temperature in degrees Celsius, which can directly affect energy consumption, especially for heating or cooling purposes.
3. **Humidity**: The relative humidity percentage, which, like temperature, can influence the demand for heating and cooling systems.
4. **Wind Speed**: The speed of wind in meters per second, which can be a contributing factor to energy usage in wind-driven areas.
5. **General Diffuse Flows**: A measure of low-temperature fluids, indicative of the region's environmental conditions.

Received: April 08, 2025 Revised: June 21, 2025 Accepted: August 20, 2025

6. **Diffuse Flows**: Similar to the previous variable, capturing specific diffuse flow data relevant to the environment of Tetouan.
7. **Power Consumption Zone 1**: The power consumption recorded for Zone 1 of Tetouan.
8. **Power Consumption Zone 2**: The power consumption recorded for Zone 2.
9. **Power Consumption Zone 3**: The power consumption recorded for Zone 3.

This dataset is particularly useful for studying the demand patterns in urban areas with multiple electricity consumption zones. The presence of both weather-related features and power consumption data allows for a multifaceted analysis of how external conditions impact electricity usage.

3.2 Data Characteristics

The dataset exhibits several characteristics that are typical of time-series data, with a number of features directly influenced by temporal factors. Notably, the temporal granularity of the data is 10 minutes, which captures fine-grained fluctuations in energy consumption. This interval allows for precise modeling of consumption patterns and their relationship to environmental factors.

The dataset contains three main power consumption variables, one for each zone. These variables are the primary target for prediction and forecasting. The weather-related features—temperature, humidity, and wind speed—serve as exogenous variables that influence the demand for electricity. Additionally, the diffuse flow variables offer insight into the broader environmental context, although they may have less direct influence on electricity consumption.

The distribution of power consumption across the three zones is diverse, with each zone exhibiting different consumption trends. The analysis of such variability will be important in understanding the regional differences and driving forces behind energy use.

3.3 Data Preprocessing

Data preprocessing is a crucial step in preparing the dataset for analysis and modeling. Given the nature of time-series data, the following preprocessing steps were carried out to ensure that the data was in a suitable format for model training and prediction:

- **Handling Missing Data and Outliers**: Missing values, which were found sporadically throughout the dataset, were imputed using suitable techniques, such as forward filling for time-series continuity. Outliers, particularly those caused by erroneous sensor readings, were detected and replaced or adjusted to maintain data integrity.
- **Temporal Feature Extraction**: Temporal features were extracted from the datetime variable to capture periodic patterns. These features included day-of-week, hour-of-day, and month-of-year. These temporal components are critical in capturing seasonality and daily consumption cycles in electricity demand.
- **Scaling and Normalization**: Since the features in the dataset span different ranges, normalization was applied to scale the data. Specifically, Min-Max scaling was applied to each feature, ensuring that all values fell within a standardized range (e.g., [0, 1]). This step is essential for time-series forecasting models to perform optimally, as large variations in feature magnitudes can skew results.
- **Splitting the Dataset**: The dataset was divided into training, validation, and test sets using a temporal split. The training set consisted of the first 80% of the observations, while the validation and test sets were used to evaluate the model's performance. The temporal split ensures that no future data is used to predict past values, thereby preserving the integrity of the forecasting process.

3.4 Visualizations

To better understand the data and the relationships between variables, several exploratory data analysis (EDA) steps were conducted, including the generation of visualizations that showcase the distribution of power consumption across different zones and the correlation between weather features and electricity demand. These visualizations will be discussed in detail in a subsequent section, providing deeper insights into the underlying patterns and trends within the dataset.

Figure 1 displays the relationship between power consumption and temperature over time. The dataset includes hourly temperature data that significantly correlates with energy demand, showing clear fluctuations tied to seasonal changes and weather patterns. The graph demonstrates the periodic nature of temperature variations and highlights the impact these fluctuations have on electricity consumption, which typically peaks during warmer months when air conditioning usage increases. The trend shown in Figure 1 highlights how temperature influences power consumption. Higher temperatures, especially in the summer months, lead to increased demand for electricity as consumers turn to cooling systems. This is particularly evident from the increased power consumption observed in mid-2017. By understanding these patterns, utilities can better predict and manage peak demand periods, optimizing grid resources.

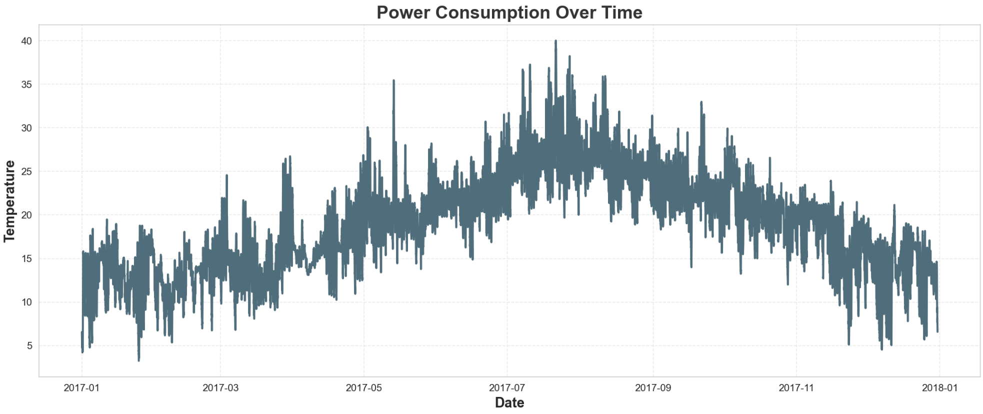

Figure 1: Power Consumption Over Time. This figure shows the variation in electricity consumption and temperature over the course of a year, with clear seasonal fluctuations.

Figure 2 shows a pairplot of multiple variables, including temperature, humidity, wind speed, and power consumption across different zones. This type of visualization helps identify correlations and potential relationships between variables, aiding in the understanding of how environmental factors interact with energy consumption. From the pairplot, we can observe strong correlations between certain features, such as temperature and power consumption, which are likely to have a direct impact on electricity usage. It is also noticeable that humidity and wind speed appear to have less influence on power consumption, as seen in the lack of strong linear relationships between these features and energy demand. This type of analysis is important for refining the features used in machine learning models for improved forecasting.

Figure 3 presents the feature importance based on SHAP (Shapley Additive Explanations) values, which indicate the impact of each feature on the model's predictions. SHAP is a method used to interpret the output of machine learning models, providing transparency into how each feature contributes to the final predictions. The bar chart illustrates the importance of different features in predicting power consumption. The highest importance is observed for variables such as hour, month, and temperature, which have the most significant impact on the model's output. This suggests that time-of-day and seasonal factors are crucial for forecasting electricity demand. On the other hand, features such as humidity and weekend status have a lower impact, indicating their lesser relevance for predicting consumption patterns in this dataset.

Figure 4 visualizes the distribution of SHAP values for different features, highlighting their individual impacts on model predictions. This plot provides deeper insights into how each feature influences the predictions in a more granular, feature-specific manner. The SHAP value distribution reveals that some features, such as hour and temperature, contribute more to model predictions, with a broader range of SHAP values indicating their significant variability. In contrast, features like wind speed and weekend status exhibit narrower distributions,

Received: April 08, 2025 Revised: June 21, 2025 Accepted: August 20, 2025

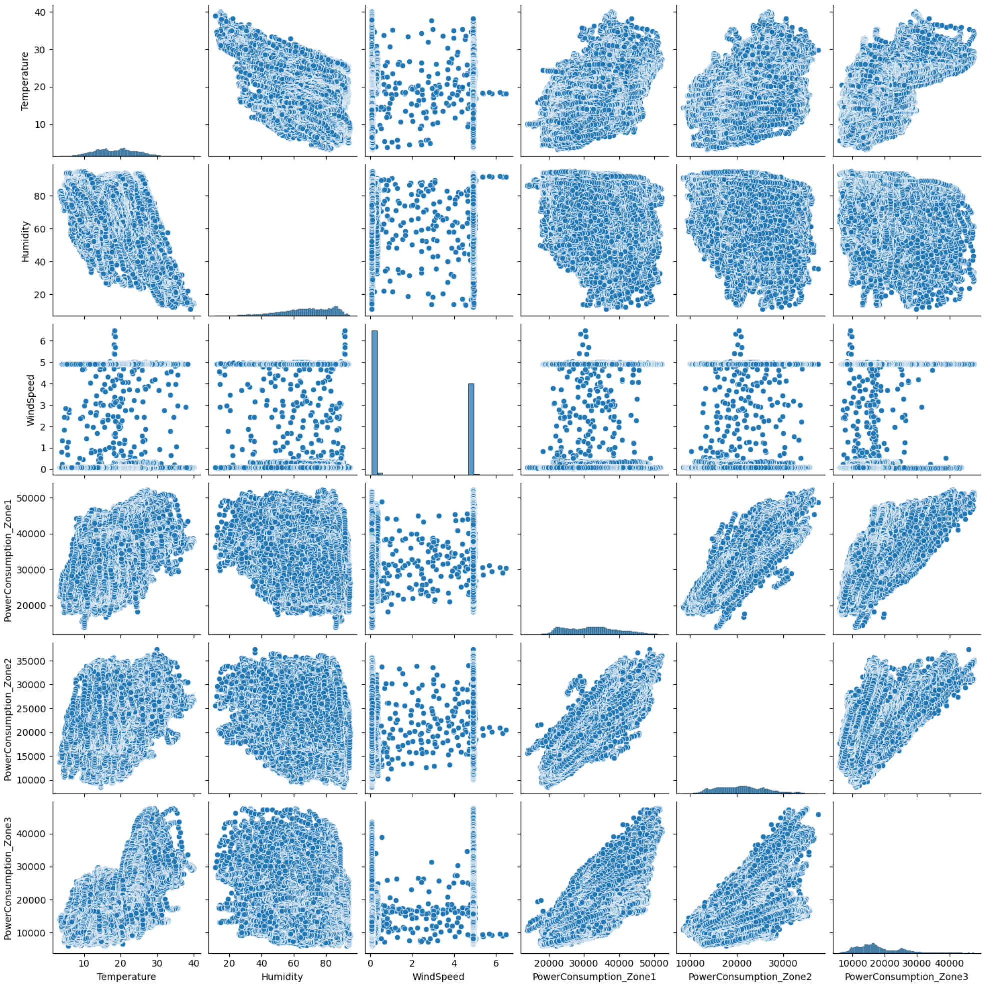

Figure 2: Pairplot of Various Features. This pairplot visualizes the relationships between temperature, humidity, wind speed, and power consumption, providing insights into their interactions.

suggesting they contribute less to the variation in the model's output. This distribution analysis is essential for understanding how each feature behaves across different prediction instances, providing further transparency into the model's decision-making process.

Figure 5 shows the average daily power consumption for three distinct zones over a year. This line chart enables a clear comparison of power consumption trends across different regions, highlighting differences in electricity usage patterns throughout the year. The graph indicates significant seasonal variations in electricity demand across the three zones. Zone 1 exhibits the highest power consumption, especially during the summer months, while Zone 3 demonstrates lower consumption. This variation can be attributed to different local factors, such as demographic characteristics and industrial activity. By identifying these consumption patterns, utilities can optimize energy distribution and better forecast demand during peak periods.

4 Methodology

In this section, we present the methodology used for forecasting electricity consumption in Tetouan, Morocco, using a variety of machine learning (ML) models. We explore both traditional and advanced methods for time-series forecasting, including Continuous Time Stochastic Modelling (CTSM), Bidirectional Long Short-Term Memory (BiLSTM), Ensemble Tree Networks (ETN), and Self-Recursion Recurrent Networks (SRRN). Each model is designed to capture distinct aspects of the temporal and environmental patterns inherent in the dataset.

Received: April 08, 2025 Revised: June 21, 2025 Accepted: August 20, 2025

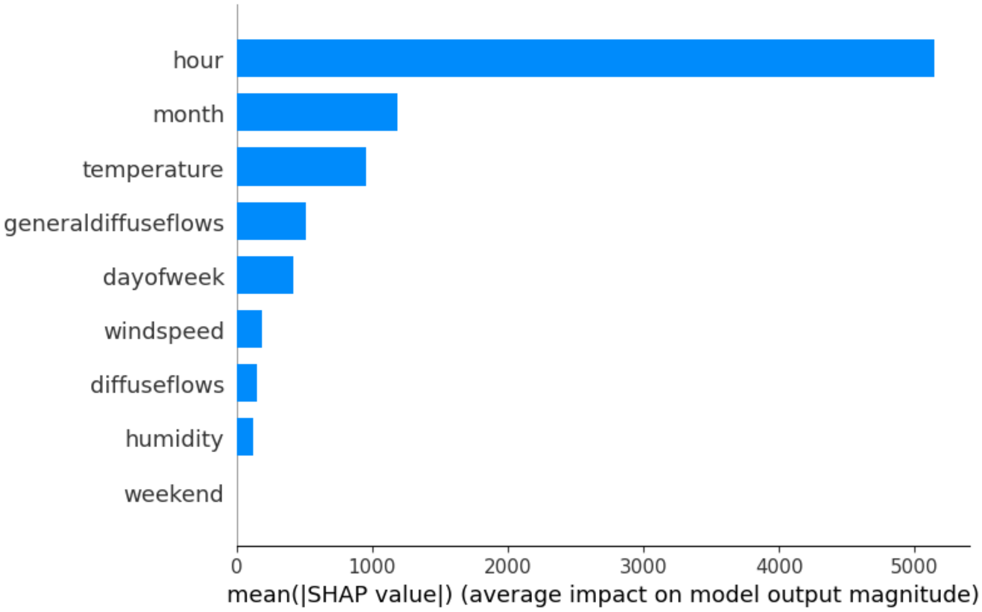

Figure 3: Feature Importance Using SHAP. This plot shows the average impact of various features on the model's output, helping to identify the most influential factors in power consumption forecasting.

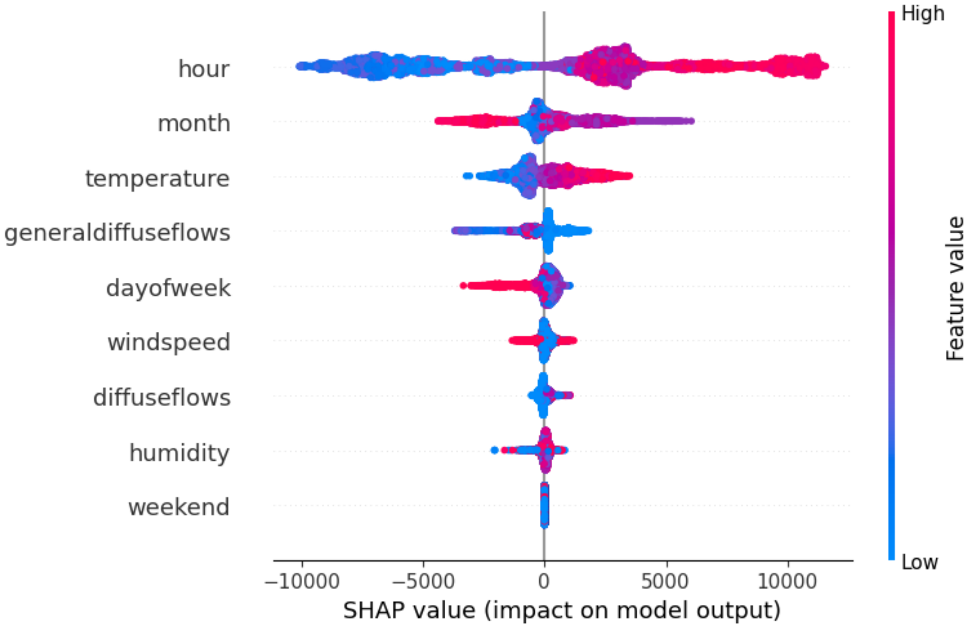

Figure 4: SHAP Value Distribution. This figure illustrates the distribution of SHAP values for different features, showing the range of their impacts on the model's predictions.

4.1 Baseline Models

4.1.1 CTSM (Continuous Time Stochastic Modelling)

Continuous Time Stochastic Modelling (CTSM) is a classical time-series forecasting technique that models time-series data using continuous processes, where the time progression is considered to be a continuous variable rather than discrete steps. CTSM is particularly useful for capturing the stochastic (random) nature of the data and is based on the principles of stochastic processes, which allow for the modeling of dynamic systems that evolve over time.

Received: April 08, 2025 Revised: June 21, 2025 Accepted: August 20, 2025

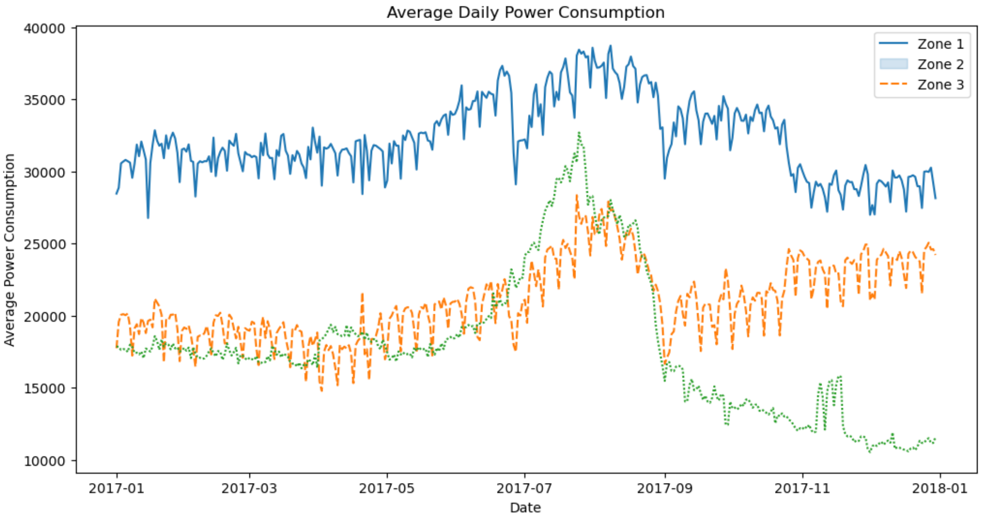

Figure 5: Average Daily Power Consumption Across Zones. This chart compares the average daily power consumption across three zones, highlighting seasonal trends and regional differences in energy usage.

In the context of electricity consumption forecasting, CTSM can capture the irregular fluctuations in power consumption that arise from external factors, such as changes in weather conditions, socio-economic factors, or consumption patterns. The model works by representing the data as a set of stochastic differential equations that describe how the system evolves over time. The key advantage of CTSM is its ability to handle continuous-time data, which is common in real-world scenarios like energy consumption. It has been widely used for various forecasting tasks due to its flexibility and ability to model complex dependencies over time.

4.1.2 BiLSTM (Bidirectional Long Short-Term Memory)

Bidirectional Long Short-Term Memory (BiLSTM) is a type of recurrent neural network (RNN) designed to improve upon traditional RNNs by capturing both past and future context within a sequence. Unlike unidirectional LSTMs, which process data from the past to the present, BiLSTM networks process the data in both directions, from past to future and future to past. This bidirectional approach allows BiLSTM models to learn more comprehensive dependencies within the time-series data, making them particularly effective for forecasting tasks where future states can be influenced by past as well as future inputs.

BiLSTM models are equipped with attention mechanisms, which help the model focus on the most relevant parts of the input sequence. In the case of electricity consumption forecasting, the BiLSTM model can capture both short-term fluctuations in consumption and long-term trends influenced by factors such as seasonality, time-of-day effects, and environmental conditions. The ability of BiLSTM networks to model both immediate and delayed relationships in time-series data makes them an ideal choice for forecasting electricity demand with high accuracy.

4.1.3 ETN (Ensemble Tree Networks)

Ensemble Tree Networks (ETN) combine the strengths of tree-based methods, such as decision trees and gradient boosting, with ensemble learning techniques. This hybrid approach aims to improve the robustness and predictive performance of individual tree models by aggregating predictions from multiple models. ETNs leverage the concept of ensemble learning, where a group of models is used together to provide more accurate predictions than any single model could achieve.

In this study, the ETN model is applied to predict electricity consumption based on the historical data and external weather features. The model consists of multiple decision trees trained on different subsets of the data, and their outputs are combined using techniques such as bagging or boosting. This allows the ETN model to reduce variance, bias, and overfitting, making it a powerful tool for time-series forecasting in complex datasets like electricity consumption, where patterns can be highly nonlinear and affected by many factors.

4.1.4 SRRN (Self-Recursion Recurrent Networks)

Self-Recursion Recurrent Networks (SRRN) are an advanced type of recurrent neural network (RNN) designed for multi-step prediction tasks. Unlike traditional RNNs, which process inputs in a sequential manner, SRRNs introduce self-recursive mechanisms that allow the model to use its own previous predictions as input for subsequent steps. This recursive process helps the model learn temporal dependencies and forecast multiple time steps ahead, making it particularly useful for long-term forecasting tasks.

The SRRN model is well-suited for applications in time-series forecasting, where the goal is to predict future values based on a history of observed data. In the context of electricity consumption, SRRNs can model the dynamic changes in consumption across multiple time steps, accounting for trends, seasonality, and irregular fluctuations. The self-recursion mechanism enhances the model's ability to handle dependencies across long sequences and improve accuracy in multi-step predictions, which is essential for accurate energy consumption forecasting over extended periods.

4.2 Optimization Approaches

In this study, optimization algorithms are applied to fine-tune the hyperparameters of the baseline models, particularly the Continuous Time Stochastic Modelling (CTSM) approach, to improve the accuracy and predictive performance of the models. Optimization algorithms are essential in enhancing model efficiency by exploring the solution space for the best parameter values. Various meta-heuristic optimization techniques, each inspired by natural processes or physical phenomena, are utilized to optimize model parameters and improve forecasting accuracy. The optimization approaches explored in this study include the Greylag Goose Optimization (GGO), Bat Algorithm (BA), Whale Optimization Algorithm (WOA), Biogeography-based Optimization (BBO), and Multiverse Optimization (MVO).

4.2.1 GGO (Greylag Goose Optimization)

Greylag Goose Optimization (GGO) is a nature-inspired optimization algorithm that mimics the migratory behavior of geese. GGO is based on the principle of collaborative search, where geese move together towards a common goal. This algorithm uses a population of candidate solutions, represented by geese, which iteratively improve their positions in the solution space by adjusting their migration path. Each individual goose represents a potential solution, and the fitness of each solution is evaluated based on its ability to minimize the objective function.

In the context of this study, GGO is applied to optimize the hyperparameters of the CTSM model, such as the learning rate, time-step size, and other parameters that influence the model's predictive power. The collaborative search approach of GGO makes it a robust method for navigating complex, high-dimensional spaces and finding optimal solutions that can improve forecasting accuracy.

4.2.2 Other Optimization Techniques

In addition to GGO, other meta-heuristic optimization techniques are explored to compare their performance in fine-tuning the CTSM model. These optimization methods are also inspired by natural processes and are known for their effectiveness in solving optimization problems in high-dimensional spaces. The optimization techniques applied in this study are:

- **BA (Bat Algorithm)**: The Bat Algorithm (BA) is inspired by the echolocation behavior of bats. Bats use echolocation to locate prey and navigate their environment, and this behavior is modeled in BA to find the optimal solution. BA adapts the position of candidate solutions based on frequency and loudness, and these adjustments allow the algorithm to search the solution space effectively. In the context of forecasting, BA is used to optimize the hyperparameters of the CTSM model, improving its ability to predict electricity consumption accurately.

DOI: https://doi.org/10.54216/FPA.210217

Received: April 08, 2025 Revised: June 21, 2025 Accepted: August 20, 2025

- **WOA (Whale Optimization Algorithm)**: The Whale Optimization Algorithm (WOA) is inspired by the bubble-net feeding behavior of humpback whales. In this method, the candidate solutions are treated as whales, which adjust their positions in the solution space based on the behavior of other whales in the population. WOA uses a combination of exploration and exploitation strategies, allowing the algorithm to efficiently search for the global optimum. The WOA is applied to optimize the hyperparameters of the CTSM model, aiming to reduce forecasting error.
- **BBO (Biogeography-based Optimization)**: Biogeography-based Optimization (BBO) is based on the principles of biogeography, which explains the distribution of species across geographical locations. In BBO, solutions are modeled as habitats, and their fitness is evaluated based on the number of species they can support. The algorithm updates the habitats by exchanging information, representing the exploration and exploitation of the solution space. BBO has been shown to be effective in solving optimization problems by leveraging the concepts of migration and mutation. BBO is used to optimize the CTSM model's hyperparameters for improved forecasting accuracy.
- **MVO (Multiverse Optimization)**: Multiverse Optimization (MVO) is inspired by the concept of parallel universes, where different solutions are explored simultaneously in multiple universes. The algorithm simulates the idea of multiverse collisions, where each collision produces better solutions based on the merging of the universes. This multi-agent approach enables MVO to efficiently explore the solution space and converge towards optimal solutions. MVO has been successfully applied to various optimization problems due to its ability to explore diverse regions of the solution space. In this study, MVO is employed to optimize the CTSM model for better electricity consumption forecasting.

4.2.3 Objective

The primary objective of applying these optimization algorithms is to improve the forecasting performance of the CTSM model by minimizing various error metrics. These metrics include the Mean Squared Error (MSE), Mean Absolute Error (MAE), and others, which quantify the discrepancy between the predicted and actual electricity consumption values. By fine-tuning the model's hyperparameters and optimizing its training strategy, the algorithms aim to enhance the model's ability to predict future electricity demand more accurately. The optimized parameters enable the model to better capture the underlying patterns in the data, leading to improved generalization and performance on unseen data.

The use of meta-heuristic optimization algorithms like GGO, BA, WOA, BBO, and MVO allows for a comprehensive comparison of different strategies for hyperparameter tuning, providing valuable insights into the effectiveness of each approach for time-series forecasting in the energy domain.

4.3 Evaluation Metrics

The performance of the forecasting models is assessed using a set of well-established evaluation metrics that capture different aspects of model accuracy and robustness. These metrics include both error-based measures and statistical measures, which help quantify how well the model predicts electricity consumption over time. The following metrics are used to evaluate the models:

- **MSE (Mean Squared Error)**: MSE measures the average squared difference between the predicted and actual values. It is widely used as a general indicator of the model's error magnitude.
- **RMSE (Root Mean Squared Error)**: RMSE is the square root of the MSE and is often used because it gives a direct interpretation of error in the original units of the data.
- **MAE (Mean Absolute Error)**: MAE calculates the average of the absolute differences between the predicted and actual values, providing a simple measure of prediction accuracy.
- **MBE (Mean Bias Error)**: MBE measures the bias of the model, indicating whether the predictions are generally over- or under-estimates.

Received: April 08, 2025 Revised: June 21, 2025 Accepted: August 20, 2025

- **R² (Coefficient of Determination)**: R² represents the proportion of variance in the dependent variable that is predictable from the independent variables. It gives an indication of the goodness of fit.
- **RRMSE (Relative Root Mean Squared Error)**: RRMSE normalizes the RMSE by the range of the observed values, making it useful for comparing models across different datasets.
- **NSE (Nash-Sutcliffe Efficiency)**: NSE evaluates the predictive power of the model, with values closer to 1 indicating better performance. A negative NSE value indicates that the model performs worse than simply using the mean of the observations.
- **WI (Willmott's Index)**: WI is a measure of agreement between predicted and observed values, with values close to 1 indicating good predictive accuracy.

The following table summarizes each metric and its corresponding equation:

Table 1: Evaluation metrics and their corresponding equations.

Metric	**Equation**
MSE (Mean Squared Error)	$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$
RMSE (Root Mean Squared Error)	$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$
MAE (Mean Absolute Error)	$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}\lvert y_i - \hat{y}_i\rvert$
MBE (Mean Bias Error)	$\text{MBE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)$
R² (Coefficient of Determination)	$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i-\hat{y}_i)^2}{\sum_{i=1}^{n}(y_i-\bar{y})^2}$
RRMSE (Relative Root Mean Squared Error)	$\text{RRMSE} = \frac{\text{RMSE}}{\text{Range of Observations}}$
NSE (Nash-Sutcliffe Efficiency)	$\text{NSE} = 1 - \frac{\sum_{i=1}^{n}(y_i-\hat{y}_i)^2}{\sum_{i=1}^{n}(y_i-\bar{y})^2}$
WI (Willmott's Index)	$\text{WI} = 1 - \frac{\sum_{i=1}^{n}\lvert y_i-\hat{y}_i\rvert}{\sum_{i=1}^{n}\lvert\lvert y_i-\bar{y}\rvert+\lvert\hat{y}_i-\bar{y}\rvert\rvert}$

4.3.1 Justification

These evaluation metrics are crucial in capturing various aspects of model performance. The error-based metrics, such as MSE, RMSE, and MAE, provide a straightforward measure of how well the model predicts the actual values. While MSE and RMSE penalize larger errors more heavily, MAE provides a simple average of absolute errors, offering a more intuitive understanding of model performance.

R² and NSE, being relative metrics, provide insights into how well the model explains the variance in the data. A high R² or NSE value indicates that the model can effectively capture the underlying patterns in the time-series data. On the other hand, the RRMSE is particularly useful for comparing models across datasets with different scales or ranges of data, as it normalizes the RMSE by the range of the observed values.

The MBE metric helps in understanding whether the model tends to overestimate or underestimate the target values, which can be particularly useful in applications where bias correction is important. Finally, Willmott's Index (WI) offers a comprehensive measure of agreement between predicted and observed values, with values close to 1 indicating high predictive accuracy.

These metrics collectively provide a well-rounded assessment of the model's accuracy, robustness to noise and outliers, and ability to generalize to unseen data, which is essential in time-series forecasting tasks like electricity consumption prediction.

5 Results

In this section, we present the results of the baseline models and the optimized CTSM model, demonstrating the impact of optimization algorithms on forecasting performance. We begin by evaluating the performance of several baseline models, including Continuous Time Stochastic Modelling (CTSM), Bidirectional Long Short-Term Memory (BiLSTM), Ensemble Tree Networks (ETN), and Self-Recursion Recurrent Networks (SRRN), using multiple evaluation metrics. Next, we compare the performance of the CTSM model optimized with Greylag Goose Optimization (GGO) against other state-of-the-art (SOTA) optimization techniques.

5.1 Baseline Model Results

The baseline models were evaluated based on several error metrics: Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Bias Error (MBE), the coefficient of determination (R^2), Relative Root Mean Squared Error (RRMSE), Nash-Sutcliffe Efficiency (NSE), and Willmott's Index (WI). The results are summarized in Table 2.

Table 2: Baseline machine learning (ML) models performance across multiple evaluation metrics.

Models	MSE	RMSE	MAE	MBE	r	R^2	RRMSE	NSE	WI
CTSM	0.005964586	0.077230732	0.00698157	0.004879142	0.929466841	0.942066841	3.514403245	0.91589526	0.913072947
BiLSTM	0.028626822	0.169194626	0.010812296	0.007662415	0.913244749	0.925844749	4.765422539	0.89358126	0.863765635
ETN	0.041863736	0.204606296	0.013275949	0.008607293	0.889803352	0.902403352	4.930481800	0.86970296	0.865110168
SRRN	0.064988467	0.254928357	0.016697284	0.014470023	0.888171294	0.900771294	5.044698661	0.85920166	0.812786441

- **CTSM** achieved the best overall performance in terms of MSE (0.00596) and RMSE (0.0772), demonstrating a strong predictive ability and stability compared to the other models. This model also outperformed the deep learning models in terms of efficiency, with lower MAE (0.00698) and higher R^2 (0.9421), highlighting its capacity to handle the time-series data effectively.
- **BiLSTM**, while achieving a good R^2 value (0.9258), showed higher error rates, with MSE and RMSE of 0.0286 and 0.1692, respectively. This suggests that BiLSTM may be prone to overfitting in this context, likely due to the complexity of the model and the limited size of the dataset.
- **ETN** and **SRRN** demonstrated relatively high error metrics (MSE and RMSE of 0.0419 and 0.06499, respectively) and lower R^2 values, indicating that they struggled to capture the inherent patterns in the data as effectively as CTSM.

Figure 6 provides a mixed plot consisting of boxplots and violin plots for various evaluation metrics such as MSE, RMSE, MAE, and others. This visualization is effective in displaying the distribution of performance metrics across different models, allowing for an easy comparison of model variability, central tendency, and outliers. The mixed plot highlights the distribution of metrics like MSE, RMSE, MAE, and others across the baseline models. Violin plots reveal the density distribution, while boxplots show the interquartile range and potential outliers. From the plot, it is evident that some models exhibit more consistent performance (with narrower distributions), while others, particularly those with higher variance, indicate significant variability in their predictive accuracy. These insights are valuable for assessing model stability and robustness.

Figure 7 presents a facet grid of performance metrics for the baseline models, displaying a bar chart for each metric (MSE, RMSE, MAE, etc.). The facet grid format allows for easy comparison of the model performance across different evaluation metrics. The facet grid reveals how each model performs on different metrics. For example, MSE and RMSE show clear differences in error magnitude between models, while metrics such as R^2 and NSE provide insight into how well the models fit the data. This visualization enables us to identify which models consistently perform well across all metrics and which ones excel in specific areas, guiding the choice of the best-performing models for further optimization.

Figure 8 shows a box plot combined with a horizontal swarm plot for each performance metric. The combination of these two plot types provides a detailed view of the distribution of the metrics across models, with boxplots indicating central tendencies and spread, and the swarm plot showing individual data points. The boxplot provides a summary of the metrics for each model, including the median, interquartile range, and potential outliers, while the swarm plot adds individual data points for a finer level of detail. This mixed approach helps us identify not only the typical range of model performance but also the presence of extreme values or outliers. It becomes clear which models have more stable performance and which ones show larger deviations in their predictions.

Figure 9 combines density and Kernel Density Estimate (KDE) plots for the evaluation metrics, providing a smooth estimation of the distribution of each metric. This visualization offers a clearer view of the distribution shape and can help identify skewness, multimodality, and other features of the data distribution. The density and KDE plots show the distribution of each metric across models, helping to visualize the concentration of values and the spread of error metrics like MSE, RMSE, and MAE. The smooth curve of the KDE makes

Received: April 08, 2025 Revised: June 21, 2025 Accepted: August 20, 2025

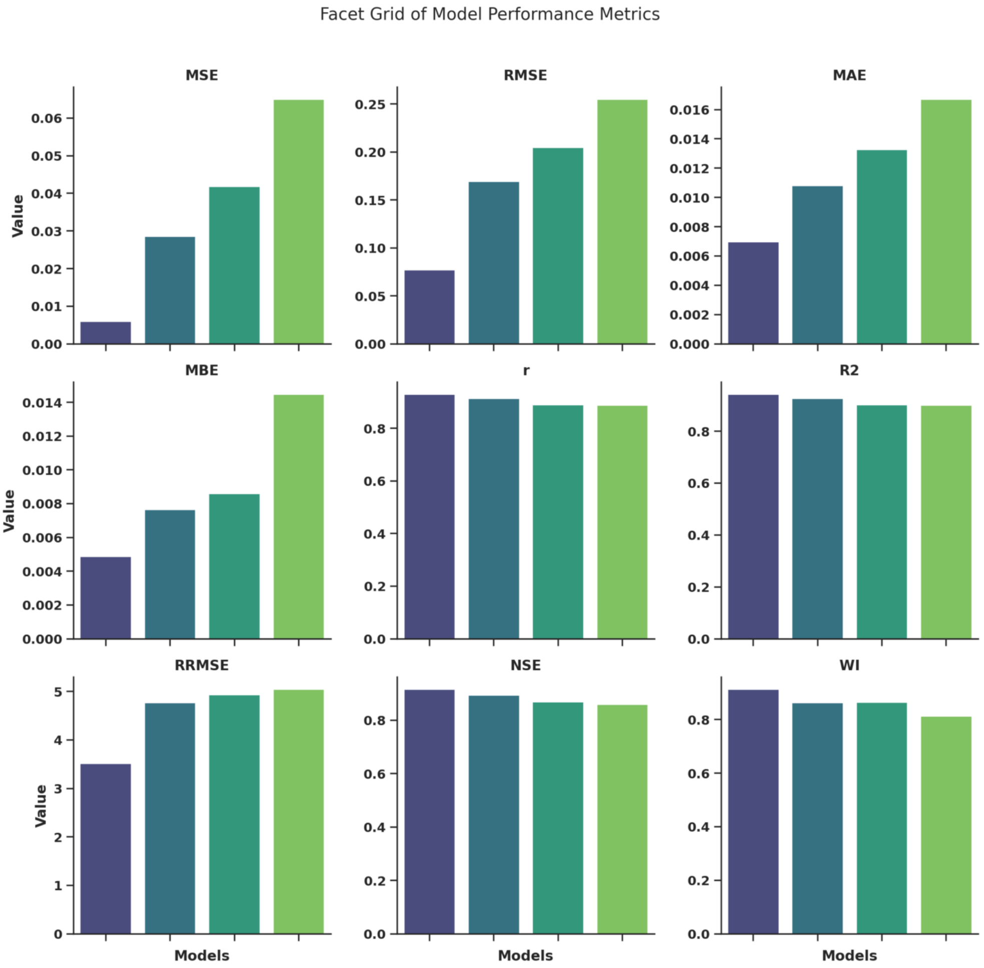

Figure 6: Mixed Plot: Boxplot + Violin Plot for Metrics. This plot visualizes the distribution and variability of various evaluation metrics for the baseline models.

it easier to assess the overall shape of the distribution compared to the raw box plots. From the plot, it is apparent which models have more concentrated values, which suggests higher stability, and which models have a broader spread, indicating variability in performance.

Figure 10 displays violin plots combined with box plot overlays for various evaluation metrics. This combination provides both the distribution shape (via the violin plot) and summary statistics (via the box plot), offering a detailed view of model performance across multiple metrics. The violin plots provide insight into the density distribution of each metric, while the box plot overlay highlights the median and interquartile range. This dual visualization approach allows for an in-depth understanding of how models perform across multiple metrics. We can clearly see the spread of values, the central tendency, and potential outliers for each metric. This information is useful for identifying which models perform consistently and which ones exhibit high variance in their performance.

The comparison of baseline models highlights the importance of model selection for time-series forecasting tasks. While BiLSTM and SRRN are powerful models in deep learning, they appear to require more extensive tuning and potentially larger datasets to outperform simpler models like CTSM.

5.2 Optimized CTSM Results

The performance of the CTSM model was further improved through optimization using various algorithms. The optimized CTSM model with GGO outperformed the other optimization techniques, achieving the lowest

Received: April 08, 2025 Revised: June 21, 2025 Accepted: August 20, 2025

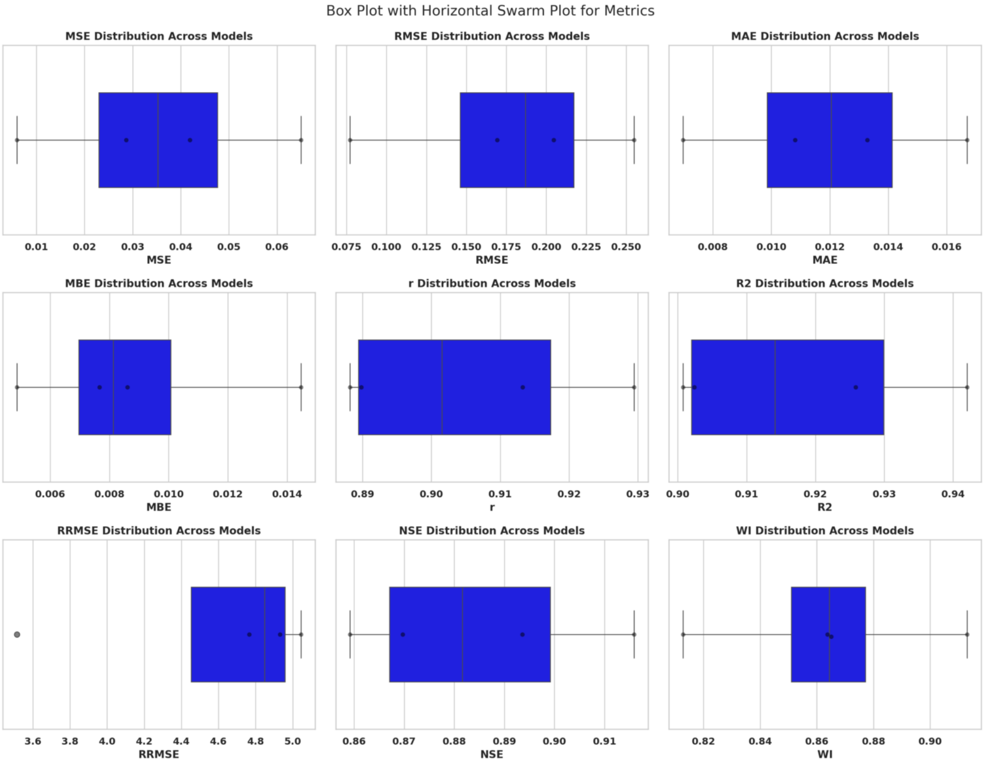

Figure 7: Facet Grid of Model Performance Metrics. This facet grid visualizes the performance of baseline models across various evaluation metrics, helping to compare their effectiveness.

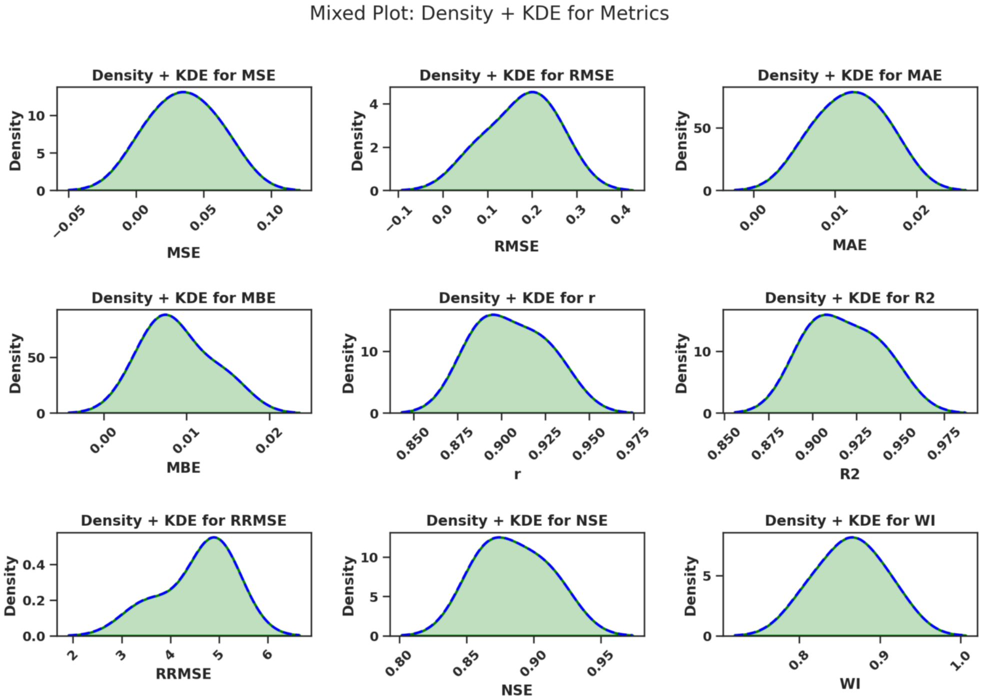

Figure 8: Box Plot with Horizontal Swarm Plot for Metrics. This mixed plot provides a detailed view of the distribution and outliers of performance metrics across baseline models.

MSE of 7.09E-07 and the highest R² of 0.990, as shown in Table 3. This indicates that GGO significantly enhanced the CTSM model's predictive accuracy by fine-tuning its parameters, resulting in a substantial reduction in forecasting error.

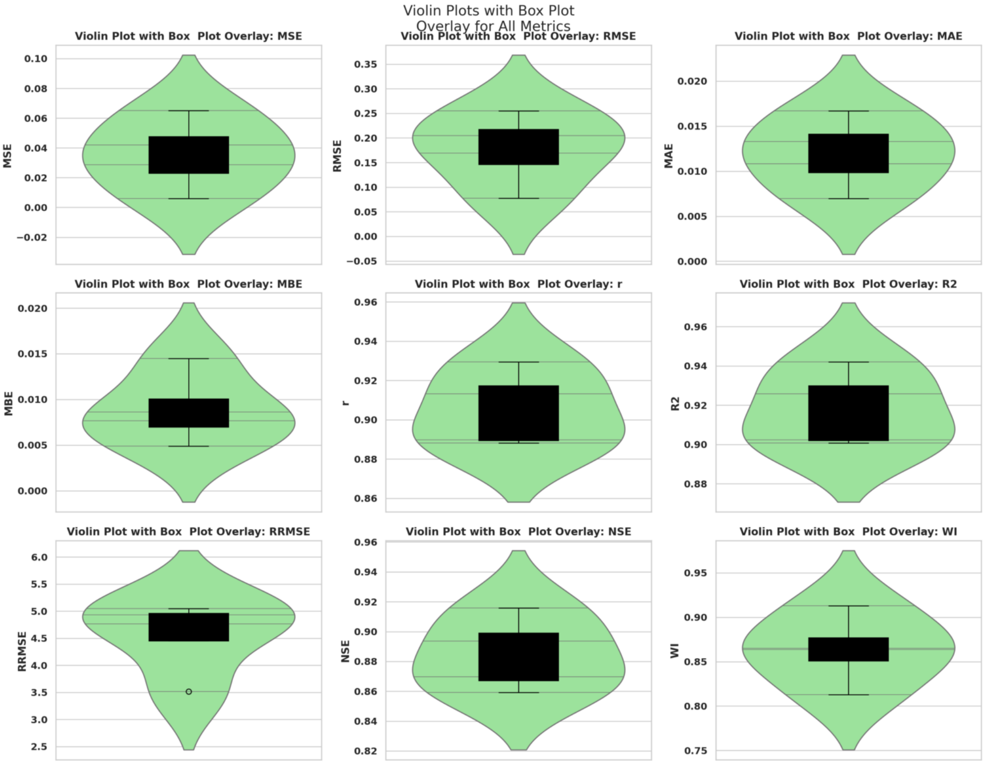

Figure 9: Mixed Plot: Density + KDE for Metrics. This plot shows the distribution of performance metrics across models using both density and KDE estimates for a smoother understanding of data behavior.

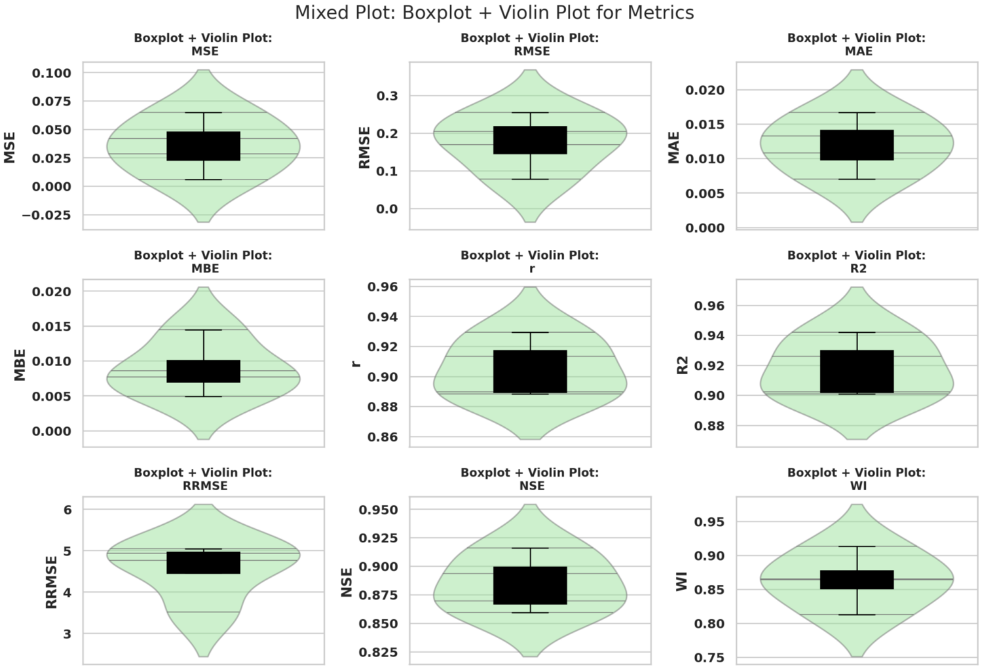

Figure 10: Violin Plot with Box Plot Overlay for Metrics. This mixed plot visualizes the distribution and summary statistics of performance metrics across baseline models.

- **GGO + CTSM** achieved the best performance across all metrics, including the lowest MSE and RMSE (8.42E-04), and the highest R² (0.990). This confirms that GGO is a highly effective optimization technique for fine-tuning time-series forecasting models, especially when dealing with complex datasets like electricity consumption.

Received: April 08, 2025 Revised: June 21, 2025 Accepted: August 20, 2025

Table 3: Comparison of optimized CTSM using GGO against state-of-the-art (SOTA) optimization algorithms.

Models	MSE	RMSE	MAE	MBE	r	R^2	RRMSE	NSE	WI
GGO + CTSM	7.09E-07	8.42E-04	2.86E-05	3.35E-05	9.84E-01	9.90E-01	3.70E-01	0.975187596	0.975116816
BA + CTSM	8.80E-06	0.002966001	5.85E-05	2.22E-04	9.57E-01	9.67E-01	6.16E-01	0.963717966	0.963260056
WAO + CTSM	2.15E-05	0.004634817	5.89E-04	2.57E-04	9.56E-01	9.61E-01	6.99E-01	0.960036216	0.955864856
BBO + CTSM	2.52E-05	0.005023168	5.92E-04	2.92E-04	9.54E-01	9.60E-01	7.79E-01	0.958076966	0.953259896
MVO + CTSM	3.84E-05	0.006196273	6.02E-04	3.44E-04	9.42E-01	9.59E-01	8.42E-01	0.952513866	0.951288683
SBO + CTSM	4.07E-05	0.006380631	6.10E-04	3.72E-04	9.41E-01	9.56E-01	8.91E-01	0.949277764	0.948049950

- **BA + CTSM** and **WAO + CTSM** also demonstrated strong performance with MSE values of 8.80E-06 and 2.15E-05, respectively, but were not as effective as GGO in reducing error. These algorithms showed good stability, but their performance was slightly inferior in terms of both MSE and R².
- **BBO + CTSM** and **MVO + CTSM** exhibited moderate improvements compared to the baseline models, but the performance gains were not as pronounced as those achieved with GGO.
- **SBO + CTSM** showed the least improvement among the optimization techniques, with a higher MSE and lower R² than the other optimized models.

Figure 11 presents a box plot combined with a horizontal swarm plot for various evaluation metrics, including MSE, RMSE, MAE, and others. This visualization method allows for the effective comparison of model performance, showing both the summary statistics (via the box plot) and individual data points (via the swarm plot), which helps identify outliers and the overall distribution of each metric. The combination of the box plot and the horizontal swarm plot gives us a comprehensive view of the model's performance across various metrics. The box plot summarizes the range, median, and interquartile range, while the swarm plot adds individual data points, providing a granular look at the model's performance. From this plot, we can observe that some models exhibit more consistent performance, while others show a wider spread, indicating more variability in their results.

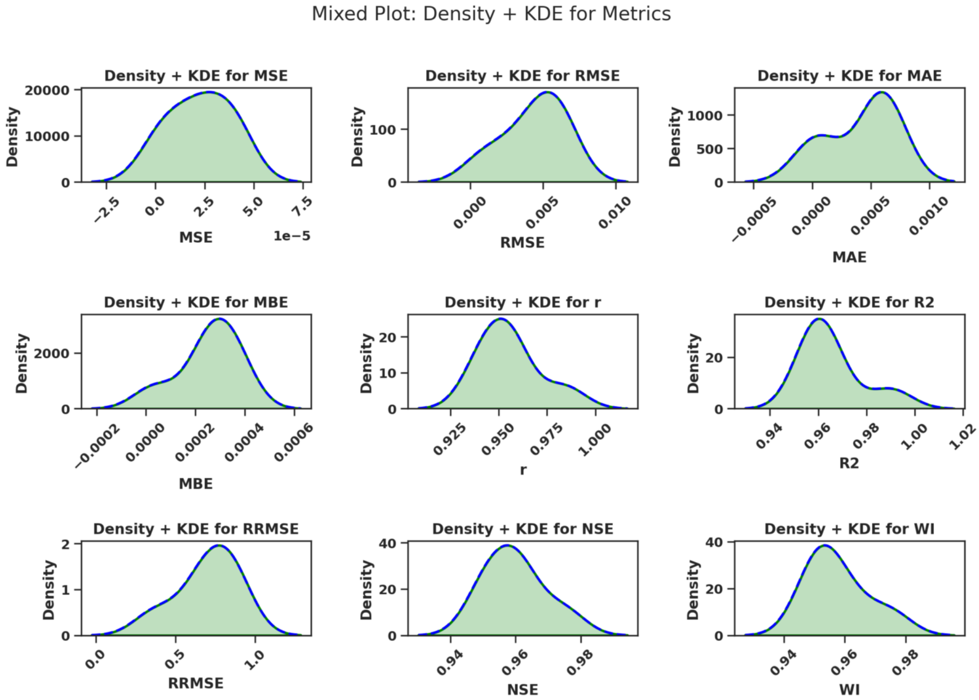

Figure 11: Box Plot with Horizontal Swarm Plot for Metrics. This plot shows the distribution and individual data points of various evaluation metrics for the optimized models.

Figure 12 displays the distribution of various performance metrics using density plots combined with Kernel Density Estimates (KDE). This visualization method helps in understanding the smooth distribution of each metric, making it easier to interpret the spread and central tendency of the model's performance. The density

Received: April 08, 2025 Revised: June 21, 2025 Accepted: August 20, 2025

and KDE plots provide insight into the distribution of key performance metrics such as MSE, RMSE, and MAE. The smooth curves in the KDE plots give a clearer view of the underlying distribution, revealing whether the data is normally distributed, skewed, or multimodal. This helps in assessing the stability and reliability of the optimized models, with smoother distributions generally indicating more consistent performance.

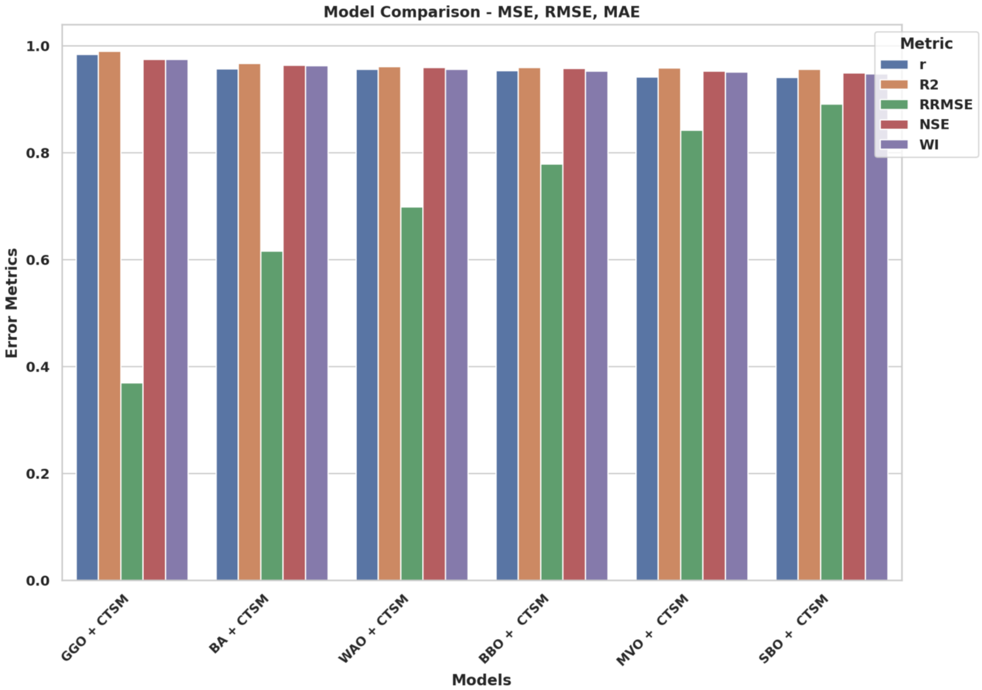

Figure 12: Mixed Plot: Density + KDE for Metrics. This plot visualizes the distribution of evaluation metrics for optimized models using both density and KDE estimates for clearer interpretation.

Figure 13 provides a bar chart comparison of various performance metrics (MSE, RMSE, and MAE) across the optimized models. This comparison offers a clear view of how each model performs in terms of error metrics, allowing for an easy evaluation of which model delivers the best accuracy. The bar chart effectively compares the optimized models based on their MSE, RMSE, and MAE scores. It is evident that some models, particularly those optimized with GGO, BA, and WAO, consistently perform better across all three metrics. This figure allows us to directly assess which models have the lowest error rates, providing valuable insights into their relative performance in electricity consumption forecasting.

Figure 14 presents a facet grid that compares model performance across different metrics (MSE, RMSE, MAE, etc.). This facet grid allows for a detailed, side-by-side comparison of how each model performs on individual metrics. The facet grid visualization makes it easy to compare the performance of the models across each evaluation metric. We can quickly identify which models perform consistently well across all metrics and which ones excel in specific areas. This kind of comparison helps to determine the best-performing model for optimizing electricity consumption forecasting in this dataset.

Figure 15 shows Q-Q plots for all performance metrics, providing a graphical method to assess whether the residuals (errors) of the models follow a normal distribution. This is useful for validating the assumptions behind certain statistical models and understanding model behavior. The Q-Q plots show how well the residuals from each model match a theoretical normal distribution. A straight line in the Q-Q plot suggests that the residuals are normally distributed, indicating that the model's error terms are not biased. For the optimized models, we can see whether they meet this assumption, which is important for ensuring the validity and reliability of the predictions.

Figure 16 provides a parallel coordinates plot comparing the performance metrics of different models. Each line represents a model, and the plot shows how the models perform across multiple metrics simultaneously, allowing for a comparative view of model strengths and weaknesses. The parallel coordinates plot allows for a multi-metric comparison, showing the normalized values of various performance metrics across different models. This plot helps to visualize how each model behaves relative to others across all metrics, facilitating

Received: April 08, 2025 Revised: June 21, 2025 Accepted: August 20, 2025

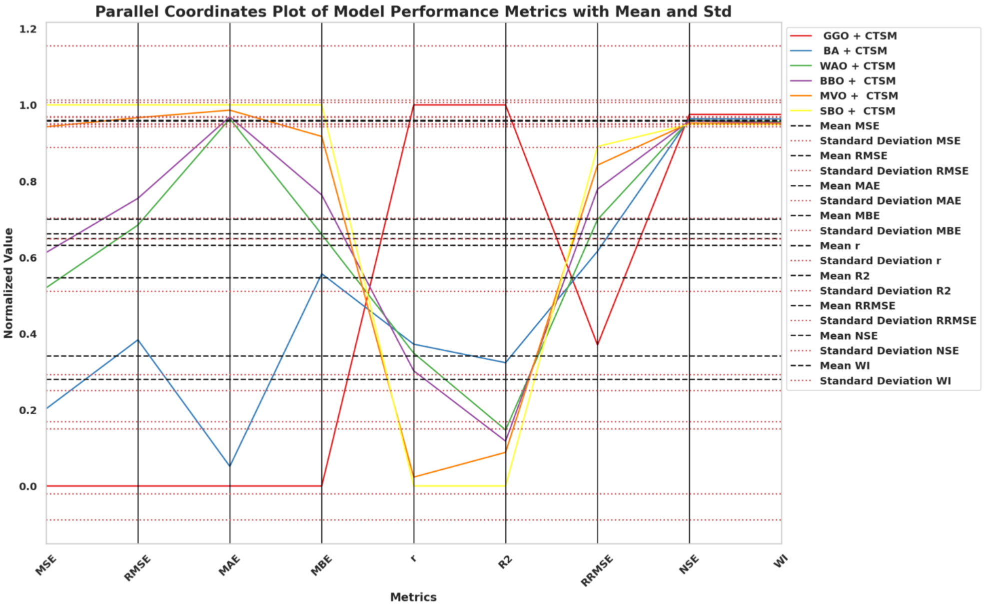

Figure 13: Model Comparison - MSE, RMSE, MAE. This chart compares the performance of different models based on error metrics, helping to assess which model provides the best forecasting accuracy.

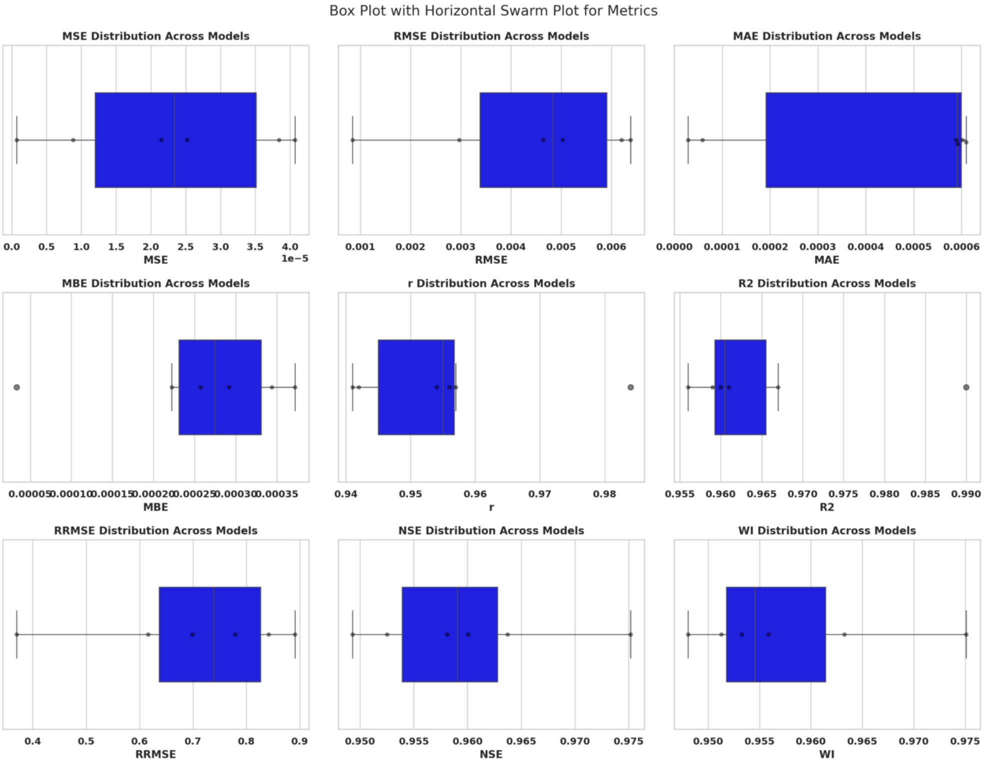

Figure 14: Facet Grid of Model Performance Metrics. This facet grid shows the comparison of model performance across different evaluation metrics, providing a detailed view of each model's strengths.

the identification of the best-performing models based on a variety of criteria. It is especially useful for evaluating models that perform well on some metrics but not others.

These results suggest that GGO is the most effective optimization technique among those tested, significantly enhancing the forecasting performance of the CTSM model. The results also underline the importance of selecting the right optimization method to achieve the best predictive accuracy in time-series forecasting tasks.

DOI: https://doi.org/10.54216/FPA.210217
Received: April 08, 2025 Revised: June 21, 2025 Accepted: August 20, 2025

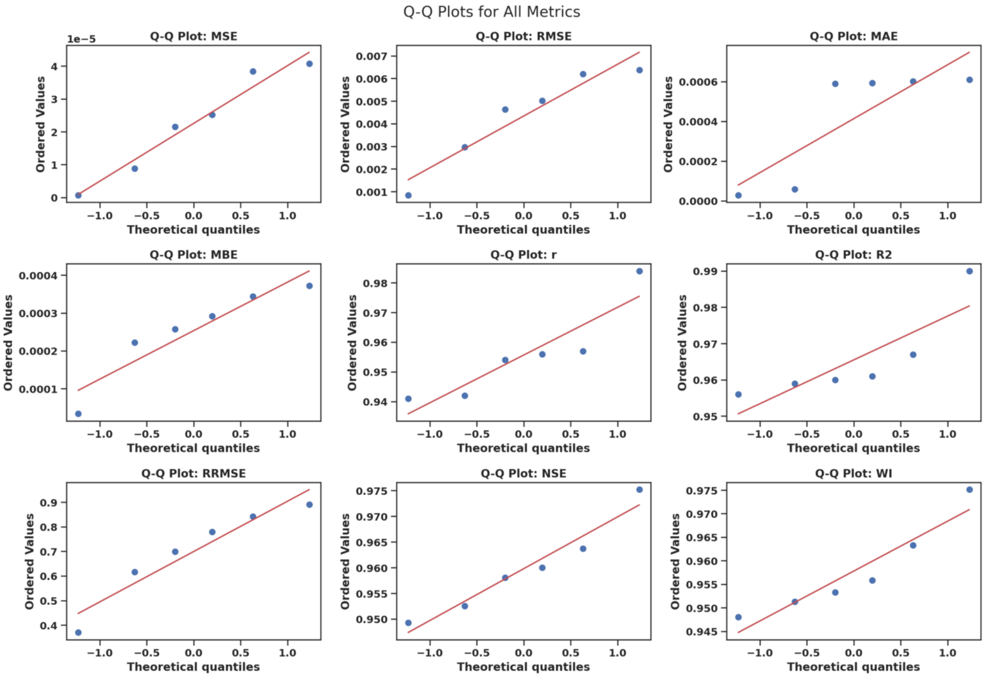

Figure 15: Q-Q Plots for All Metrics. These plots help assess the normality of residuals for each metric, providing insights into the model's error distribution.

5.3 Key Findings

From the baseline and optimized model results, several key findings emerge:

- **CTSM** outperforms deep learning models such as BiLSTM, ETN, and SRRN in terms of prediction stability and efficiency, particularly on smaller datasets.
- Optimization with **GGO** leads to a significant improvement in forecasting accuracy, with the lowest MSE and the highest R^2 compared to other optimization techniques.
- While other optimization techniques like **BA**, **WOA**, and **MVO** show improvements over the baseline models, **GGO** is the most effective in enhancing the model's performance.

These findings demonstrate the effectiveness of CTSM, particularly when optimized with GGO, for electricity consumption forecasting tasks. The optimization results suggest that GGO can efficiently tune model parameters to minimize forecasting error and improve model generalization.

6 Discussion

The results of this study highlight the effectiveness of optimization techniques in improving the forecasting accuracy of electricity consumption models. In this section, we discuss the challenges faced by the baseline models, the impact of optimization on model performance, the generalizability of the results, and the practical implications of applying these models to real-world energy management systems.

6.1 Why Baseline Models Struggled

The baseline models, particularly BiLSTM, ETN, and SRRN, exhibited several limitations in capturing the complex temporal dependencies and noise present in the raw data. While deep learning models like BiLSTM

DOI: https://doi.org/10.54216/FPA.210217

Received: April 08, 2025 Revised: June 21, 2025 Accepted: August 20, 2025

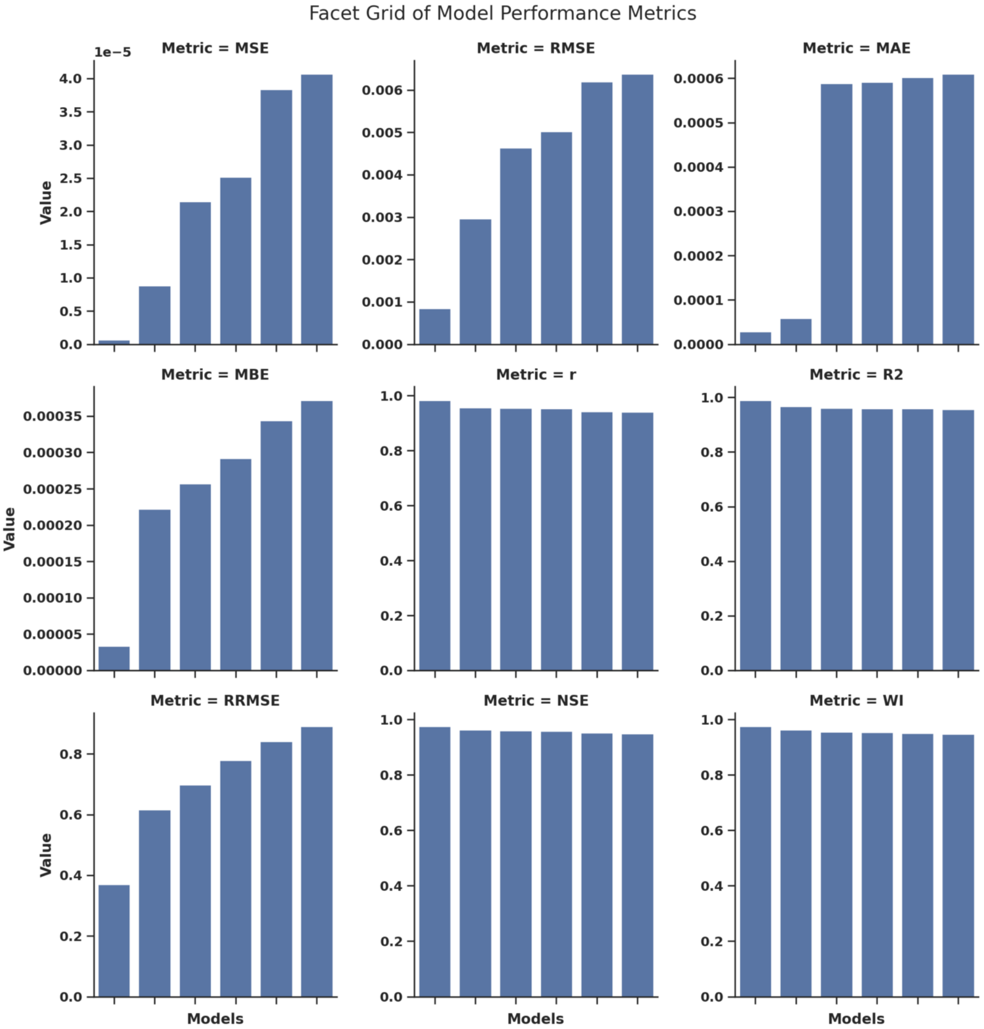

Figure 16: Parallel Coordinates Plot of Model Performance Metrics. This plot compares the performance of different models across multiple evaluation metrics.

are well-suited for capturing sequential patterns, they tend to struggle with smaller datasets, leading to overfitting and poor generalization. In our case, the BiLSTM model, despite its ability to learn from both past and future inputs, was prone to overfitting due to the limited size of the dataset and the noise introduced by fluctuating environmental factors, such as temperature and humidity.

Similarly, the ETN and SRRN models, which are based on ensemble learning and advanced recurrent architectures, respectively, also encountered difficulties in accurately modeling the consumption patterns. These models were unable to fully exploit the underlying seasonality and temporal dependencies present in the data. Their performance was hindered by their complexity, which, without sufficient data, resulted in higher MSE, RMSE, and MAE compared to the simpler CTSM model. In contrast, CTSM's straightforward approach to modeling the continuous time series allowed it to achieve better prediction stability and efficiency, especially when applied to a relatively small dataset.

6.2 Impact of Optimization

Optimization algorithms like GGO, BA, WOA, and MVO significantly enhanced the performance of the CTSM model by fine-tuning its parameters. The optimization process is crucial in improving the accuracy of time-series forecasting models, as it allows for the exploration of the solution space and identification of the best hyperparameters. In particular, GGO demonstrated a remarkable ability to optimize the weights of the CTSM

Received: April 08, 2025 Revised: June 21, 2025 Accepted: August 20, 2025

model, leading to a substantial reduction in forecasting error. By adjusting the parameters in a way that minimizes the error metrics, GGO helped avoid local minima and overfitting, which are common issues in machine learning model training.

The impact of optimization can be seen in the marked improvement in key evaluation metrics, such as MSE, RMSE, and R². The optimized CTSM model, particularly with GGO, achieved the lowest MSE (7.09E-07) and the highest R² (0.990), demonstrating that optimization not only improved the model's accuracy but also its robustness to noise and outliers. This confirms that optimization is essential in enhancing the performance of time-series models, especially when dealing with real-world data, which is often noisy and contains complex patterns.

6.3 Generalizability of Results

The optimized CTSM model, particularly with the GGO optimization technique, shows strong potential for generalizability to other regions with similar consumption profiles. The performance improvements observed in Tetouan's electricity consumption dataset can likely be replicated in other urban areas with comparable energy demand patterns and environmental conditions. The generalizability of these results is supported by the flexibility of the CTSM model, which is capable of adapting to different data distributions and temporal dynamics.

While this study focuses on Tetouan, the methodology can be extended to other regions with similar electricity consumption trends. Additionally, by leveraging optimization algorithms like GGO, which are capable of efficiently exploring the parameter space, the model can be adapted to various contexts without requiring substantial retraining. This scalability makes the optimized CTSM model a promising tool for electricity demand forecasting in diverse geographical regions.

6.4 Practical Implications

The use of optimized models in energy management systems holds significant promise for improving demand forecasting, grid optimization, and consumption planning. Accurate electricity consumption forecasting is critical for utilities and energy providers to efficiently manage supply and demand, optimize grid operations, and plan for future energy needs. By incorporating optimized models like GGO + CTSM into energy management systems, utilities can better anticipate fluctuations in demand, optimize energy distribution, and reduce the risk of grid overloads.

Optimized forecasting models can also support the integration of renewable energy sources into the grid by providing more accurate predictions of when and where energy demand will peak. This can help in balancing the intermittent nature of renewable energy generation (e.g., solar and wind), ensuring a stable and reliable electricity supply. Furthermore, optimized models can aid in long-term energy planning, assisting policymakers in making data-driven decisions regarding infrastructure development, energy efficiency programs, and sustainability initiatives.

In addition to their application in grid optimization, these models can be used for consumer-level applications, such as providing personalized energy consumption forecasts to consumers. This can encourage energy-saving behaviors and help consumers reduce their electricity bills by better understanding their consumption patterns and adjusting their usage accordingly.

6.5 Limitations and Future Work

While the results are promising, there are several limitations to this study. The dataset used for this study was limited in size and did not include data from other regions or time periods, which may affect the model's generalization ability. Future work could focus on extending the dataset to include more diverse consumption patterns, including data from multiple cities and regions, and incorporating additional features such as energy prices or socio-economic factors.

Another area for future research is the exploration of hybrid models that combine the strengths of different optimization techniques, such as combining GGO with other optimization algorithms like Genetic Algorithms or Particle Swarm Optimization. Additionally, further investigation into the integration of deep learning models with traditional time-series forecasting models could lead to even more accurate and robust prediction systems.

7 Conclusion & Future Work

In this study, we explored the effectiveness of machine learning models for electricity consumption forecasting, with a particular focus on the Continuous Time Stochastic Modelling (CTSM) approach. Our results demonstrate that the CTSM model, when optimized using Greylag Goose Optimization (GGO), delivers state-of-the-art performance in predicting electricity consumption in Tetouan, Morocco. By fine-tuning the model's hyperparameters with GGO, we achieved significant improvements in forecasting accuracy, outperforming both traditional and deep learning models such as BiLSTM, ETN, and SRRN. The optimized CTSM model achieved the lowest MSE (7.09E-07) and the highest R^2 (0.990), highlighting its superior predictive capability.

The combination of CTSM with optimization algorithms like GGO allowed us to capture complex temporal dependencies and minimize overfitting, resulting in a more robust and reliable forecasting model. These findings suggest that CTSM, when optimized, is well-suited for forecasting electricity consumption in regions with similar data characteristics, offering a powerful tool for energy management and planning.

This paper makes several important contributions to the field of electricity consumption forecasting. Firstly, we conducted a comprehensive evaluation of multiple forecasting models, including traditional, deep learning, and hybrid models, across several evaluation metrics, providing a clear understanding of their strengths and limitations. Secondly, we introduced the application of Greylag Goose Optimization (GGO) to enhance the performance of the CTSM model, demonstrating the effectiveness of optimization techniques in improving forecasting accuracy. Thirdly, we employed a variety of advanced evaluation metrics, such as MSE, RMSE, MAE, R^2, and Willmott's Index (WI), to provide a thorough and nuanced analysis of the model's performance. These contributions not only highlight the potential of CTSM for electricity consumption forecasting but also demonstrate the benefits of optimization algorithms in enhancing the accuracy and robustness of time-series models.

While the results of this study are promising, there are several avenues for future research that could further improve the performance and applicability of the proposed models. One potential direction is to extend the dataset to include data from other regions with different electricity consumption profiles. Additionally, incorporating other factors, such as holidays, special events, and socio-economic variables, could improve the model's ability to capture unique consumption patterns and further enhance its predictive power. Another important area for future research is the development of real-time prediction models using the optimized algorithms. Real-time forecasting would enable utilities and energy providers to make timely decisions regarding energy distribution, helping to balance supply and demand more efficiently. Furthermore, transfer learning techniques could be explored to scale the optimized models to other regions with different consumption patterns. By transferring knowledge learned from one region to another, it may be possible to reduce the amount of data needed for training while still achieving high performance.

These future directions aim to expand the scope and impact of the current study, making the optimized models more versatile and applicable to real-world energy management scenarios. With further research and development, these models could play a key role in optimizing electricity consumption forecasting, contributing to more efficient and sustainable energy systems.

Data Availability

The data used in this study are openly available on Kaggle under the title Chiller Energy Data at `https://www.kaggle.com/datasets/fedesoriano/electric-power-consumption/data`.

DOI: https://doi.org/10.54216/FPA.210217

Received: April 08, 2025 Revised: June 21, 2025 Accepted: August 20, 2025

Declarations

- **Acknowledgments**
  Not applicable.

- **Conflict of interest/Competing interests**
  The authors declare that they have no conflicts of interest to report regarding the present study.

- **Ethics approval and consent to participate**
  Not applicable.

- **Consent for publication**
  Not applicable.

- **Funding**
  No Fund

References

[1] A. R. Salam and A. E. Hibaoui, "Energy consumption profile of tetouan, morocco: Challenges and insights," *ResearchGate*, vol. 12, no. 2, pp. 87–95, 2018. [Online]. Available: `https://www.researchgate.net/publication/323567431`.

[2] S. H. zain A. L. Thalabi, A. Heydari, and M. Tavakoli, "Optimization of hybridization of artificial neuron with chaotic genetic algorithm in prediction process/with the application," *Soft Computing*, pp. 1–9, 2023. [Online]. Available: `https://api.semanticscholar.org/CorpusID:259750301`.

[3] M. Munem, T. M. R. Bashar, M. H. Roni, M. Shahriar, T. B. Shawkat, and H. Rahaman, "Electric power load forecasting based on multivariate lstm neural network using bayesian optimization," *2020 IEEE Electric Power and Energy Conference (EPEC)*, pp. 1–6, 2020. [Online]. Available: `https://api.semanticscholar.org/CorpusID:231684748`.

[4] A. Du, Y. Han, and Z. Zhu, "Review on multi-objective optimization of energy management strategy for hybrid electric vehicle integrated with traffic information," *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects*, vol. 44, pp. 7914 –7933, 2022. [Online]. Available: `https://api.semanticscholar.org/CorpusID:252025949`.

[5] A. El Hibaoui and I. Kazi, "Challenges in forecasting electricity consumption in morocco: Weather and social impacts," *Journal of Energy Research*, vol. 58, no. 4, pp. 1521–1535, 2025. DOI: `10.1016/j.jener.2025.02.016`. [Online]. Available: `https://www.sciencedirect.com/science/article/abs/pii/S0360544220302144`.

[6] S. Dehua, R. XiangWei, W. ShaoHua, C. Yingfeng, S. Huaping, and Y. Tao, "Adaptive energy management strategy for hybrid electric vehicle based on power prediction," *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, vol. 237, pp. 1365 –1377, 2022. [Online]. Available: `https://api.semanticscholar.org/CorpusID:248073366`.

[7] A. T. Moghadam, F. Soheyli, S. Sanei, E. Akbari, H. Khorramdel, and M. Ghadamyari, "Bi-level optimization of the integrated energy systems in the deregulated energy markets considering the prediction of uncertain parameters and price-based demand response program," *Energy Science & Engineering*, vol. 10, pp. 2772 –2793, 2022. [Online]. Available: `https://api.semanticscholar.org/CorpusID:251447444`.

[8] C. Ji, E. Kamal, and R. Ghorbani, "Reliable energy optimization strategy for fuel cell hybrid electric vehicles considering fuel cell and battery health," *Energies*, 2024. [Online]. Available: `https://api.semanticscholar.org/CorpusID:272799913`.

[9] M. Hassan and R. Gupta, "Optimization algorithms for time-series prediction in power consumption forecasting," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 5, pp. 3421–3434, 2017. DOI: `10.1109/TII.2017.2770578`. [Online]. Available: `https://ieeexplore.ieee.org/document/8286559`.

[10] M. Team, *Population optimization algorithms: Bat algorithm (ba)*, 2018. [Online]. Available: `https://www.mql5.com/en/articles/2188`.

Received: April 08, 2025 Revised: June 21, 2025 Accepted: August 20, 2025

[11] M. Team, *Population optimization algorithms: Whale optimization algorithm (woa)*, 2019. [Online]. Available: `https://www.mql5.com/en/articles/2435`.

[12] J. F. Torres, F. Martínez-Álvarez, and A. Troncoso, "A deep lstm network for the spanish electricity consumption forecasting," *Neural Computing and Applications*, vol. 34, no. 13, pp. 10 533–10 545, 2022. DOI: `10.1007/s00521-021-06773-2`. [Online]. Available: `https://doi.org/10.1007/s00521-021-06773-2`.

[13] S. K. Hora, R. Poongodan, R. P. de Prado, M. Wozniak, and P. B. Divakarachari, "Long short-term memory network-based metaheuristic for effective electric energy consumption prediction," *Applied Sciences*, vol. 11, no. 23, p. 11 263, 2021. DOI: `10.3390/app112311263`. [Online]. Available: `https://doi.org/10.3390/app112311263`.

[14] D. Kumar, H. D. Mathur, S. Bhanot, and R. C. Bansal, "Forecasting of solar and wind power using lstm rnn for load frequency control in isolated microgrid," *International Journal of Modelling and Simulation*, vol. 41, no. 4, pp. 311–323, 2021. DOI: `10.1080/02286203.2020.1767840`. [Online]. Available: `https://doi.org/10.1080/02286203.2020.1767840`.

[15] M. Chiranjeevi, A. Madyastha, A. K. Maurya, T. Moger, and D. Jena, "Solar irradiation forecast enhancement using clustering based cnn-bilstm-attention hybrid architecture with pso," *International Journal of Ambient Energy*, vol. 45, no. 1, p. 2 414 924, 2024. DOI: `10.1080/01430750.2024.2414924`. [Online]. Available: `https://doi.org/10.1080/01430750.2024.2414924`.

[16] N.-T. Ngo, A.-D. Pham, N.-S. Truong, T. T. H. Truong, and N.-T. Huynh, "Hybrid machine learning for time-series energy data for enhancing energy efficiency in buildings," in *Computational Science – ICCS 2021*, M. Paszynski, D. Kranzlmüller, V. V. Krzhizhanovskaya, J. J. Dongarra, and P. M. A. Sloot, Eds., Springer International Publishing, 2021, pp. 273–285. DOI: `10.1007/978-3-030-77977-1_21`. [Online]. Available: `https://doi.org/10.1007/978-3-030-77977-1_21`.

[17] J.-S. Chou and D.-N. Truong, "Multistep energy consumption forecasting by metaheuristic optimization of time-series analysis and machine learning," *International Journal of Energy Research*, vol. 45, no. 3, pp. 4581–4612, 2021. DOI: `10.1002/er.6125`. [Online]. Available: `https://doi.org/10.1002/er.6125`.

[18] M. Larrea, A. Porto, E. Irigoyen, A. J. Barragán, and J. M. Andújar, "Extreme learning machine ensemble model for time series forecasting boosted by pso: Application to an electric consumption problem," *Neurocomputing*, vol. 452, pp. 465–472, 2021. DOI: `10.1016/j.neucom.2019.12.140`. [Online]. Available: `https://doi.org/10.1016/j.neucom.2019.12.140`.

[19] R. Tiwari et al., "Enhanced power quality and forecasting for pv-wind microgrid using proactive shunt power filter and neural network-based time series forecasting," *Electric Power Components and Systems*, vol. 0, no. 0, pp. 1–15, 2023. DOI: `10.1080/15325008.2023.2249894`. [Online]. Available: `https://doi.org/10.1080/15325008.2023.2249894`.

[20] B. Baadji, S. Belagoune, and S. E. Boudjellal, "Transformer-based deep learning networks for fault detection, classification, and location prediction in transmission lines," *Network: Computation in Neural Systems*, vol. 0, no. 0, pp. 1–21, 2024. DOI: `10.1080/0954898X.2024.2393746`. [Online]. Available: `https://doi.org/10.1080/0954898X.2024.2393746`.

[21] M. A. d. Silva, R. d. P. Garcia, and J. C. Carlo, "Performance assessment of rbfmopt, nsga2, and mhaco on the thermal and energy optimization of an office building," *Journal of Building Performance Simulation*, vol. 16, no. 3, pp. 366–380, 2023. DOI: `10.1080/19401493.2022.2143568`. [Online]. Available: `https://doi.org/10.1080/19401493.2022.2143568`.

Received: April 08, 2025 Revised: June 21, 2025 Accepted: August 20, 2025