

Advancing Electricity Load Forecasting Using a Novel Enhanced Harris Hawks Optimization

Ahmed Rashed Almesmari¹, Mohammed Azmi Al-Betar^{1,2,*}, and Sharif Naser Makhadmeh³

¹ Artificial Intelligence Research Center (AIRC), College of Engineering and Information Technology, Ajman University, Ajman, United Arab Emirates

² Center of Excellence in Precision Medicine and Digital Health, Department of Physiology, Faculty of Dentistry, Chulalongkorn University, Bangkok, Thailand

³ Department of Information Technology, King Abdullah II School for Information Technology, The University of Jordan (UJ), Amman, Jordan

Email: 202010377@ajman.ac.ae (A.R.A.); m.albetar@ajman.ac.ae (M.A.A.); s_makhadmeh@ju.edu.jo (S.N.M.)

*Corresponding author

Abstract—This paper investigates electricity load forecasting using machine learning models enhanced with advanced optimization techniques. Six regression-based models—Gradient Boosting, LightGBM, ExtraTrees, Random Forest, Decision Tree, and Long Short-Term Memory (LSTM)—are evaluated on two real-world datasets from Panama City and Tetouan City, across hourly and 10-minute temporal resolutions. Results demonstrate that tree-based ensemble models, particularly the ExtraTreesRegressor, consistently outperform LSTM-based deep learning approaches. A key contribution is the development of an Enhanced Harris Hawks Optimization (EHHO) algorithm, incorporating adaptive parameter control and type-specific parameter handling. EHHO significantly improves hyperparameter tuning efficiency, enabling the ExtraTreesRegressor to achieve state-of-the-art forecasting accuracy. The EHHO-optimized ExtraTreesRegressor attains a Mean Absolute Percentage Error (MAPE) of 0.30% for Tetouan City and 1.47% for Panama City using 10-minute resolution data. The analysis reveals that higher temporal granularity contributes up to 65% improvement in forecasting performance compared to hourly data. These findings challenge the prevailing view of deep learning dominance in time-series forecasting and establish new accuracy benchmarks for electricity load prediction. The proposed methodology holds strong potential for practical deployment in grid operation, demand response, and renewable energy integration, supporting the development of more efficient and resilient energy systems.

Keywords—electricity load forecasting, machine learning, ensemble models, Harris Hawks optimization, smart grid

I. INTRODUCTION

Accurate electricity load forecasting is a critical component of energy planning, operational reliability, and economic efficiency in modern power systems. It enables grid operators and utility providers to make data-driven decisions on resource allocation, load

balancing, energy trading, and infrastructure development. With the global shift towards decarbonization and the increasing penetration of variable renewable energy sources such as wind and solar, the electricity grid has become more dynamic, leading to higher uncertainty in load profiles. This evolving landscape has necessitated the development of more robust and adaptive forecasting models.

Conventional load forecasting approaches, particularly those based on statistical and time series models like Autoregressive Integrated Moving Average (ARIMA) and Seasonal ARIMA (SARIMA), rely heavily on historical consumption data to predict future demand. While these models have proven effective for linear and stationary patterns, their performance significantly degrades in the presence of non-stationary, nonlinear, and complex temporal dependencies, which are characteristic of real-world electricity consumption. These limitations are further exacerbated by factors such as climate variability, evolving human behavior, economic activity, and regional-specific load dynamics.

In recent years, the emergence of Machine Learning (ML) and Artificial Intelligence (AI) techniques has brought transformative changes to the domain of electricity load forecasting. Deep Learning (DL) models such as Long Short-Term Memory (LSTM) networks and Convolutional Neural Networks (CNNs) have shown substantial improvements in capturing intricate temporal relationships and handling large-scale, high-dimensional energy datasets. These models offer enhanced flexibility and predictive accuracy, particularly in short-term forecasting applications. For instance, Salam and El-Hibaoui [1] proposed a deep learning architecture based on Inception-ResNet-v2 integrated with LSTM layers, achieving notable improvements in forecasting accuracy. Similarly, Silva *et al.* [2] demonstrated the superiority of Bidirectional LSTM models over standard LSTM in learning complex bidirectional temporal patterns.

Beyond deep learning, ensemble machine learning methods such as Random Forest (RF) and Support

Manuscript received May 16, 2025; revised August 14, 2025; accepted September 8, 2025; published June 22, 2026.

Vector Machines (SVM) have been explored for their robustness and ability to model nonlinear dependencies. Zogaan [3] highlighted the effectiveness of RF in power consumption prediction due to its capability in processing heterogeneous features and identifying complex patterns. RoSe *et al.* [4] conducted a comparative study of multiple ML algorithms and identified the Bayesian Fine Tree (BFT) as a promising technique for energy demand forecasting.

In parallel with the advancement of ML models, the integration of metaheuristic optimization algorithms has emerged as a complementary strategy to enhance model performance. Techniques such as Particle Swarm Optimization (PSO), Genetic Algorithms (GA), and newer algorithms like Gray Wolf Optimizer (GWO) and Harris Hawks Optimization (HHO) have shown great promise in tasks such as feature selection, model structure optimization, and hyperparameter tuning [5–9]. Dai and Zhao [10] proposed an SVM-based hybrid model optimized with intelligent feature selection techniques, achieving superior accuracy compared to conventional approaches.

Despite these advancements, several challenges remain. Many ML-based models require large amounts of training data and computational resources, and their performance heavily depends on the careful tuning of hyperparameters. Moreover, the variability introduced by renewable energy integration further complicates forecasting, highlighting the need for flexible, scalable, and data-efficient forecasting frameworks.

The main contributions of this study are summarized as follows:

- A hybrid forecasting framework is proposed, integrating machine learning models (e.g., LSTM, Random Forest) with advanced metaheuristic optimization techniques (e.g., GWO, Arithmetic Optimization Algorithm (AOA), Cuckoo Search), to enhance short-term electricity load prediction accuracy and computational efficiency.
- An enhanced version of the Harris Hawks Optimization (HHO) algorithm is developed, introducing adaptive control, type-aware hyperparameter handling, and boundary enforcement to improve hyper-parameter tuning effectiveness.
- The framework is validated on two diverse real-world datasets (Tetuan City and Panama), demonstrating its generalizability and robustness across different energy consumption profiles.
- Comparative analysis with traditional models confirms that the proposed framework achieves superior performance, offering practical applicability for smart grids, demand response, and renewable energy integration.

The remainder of this paper is organized as follows. Section II presents a review of related work in electricity load forecasting and highlights existing research gaps. Section III details the proposed methodology, including data preprocessing, model design, and optimization strategies. Section IV describes the experimental setup and performance metrics. In addition, the section

discusses the results obtained from applying the proposed framework to real-world datasets. Finally, Section V concludes the paper with a summary of key findings, contributions, limitations, and directions for future work.

II. RELATED WORK

Electricity Load Forecasting (ELF) remains a critical area of research due to its impact on the stability, efficiency, and sustainability of modern power systems. Forecasting is commonly classified into short-term (minutes to days), medium-term (weeks to months), and long-term (months to years) predictions [11]. The increasing integration of renewable energy sources and the dynamic nature of energy consumption have highlighted the need for advanced forecasting methodologies that go beyond traditional linear models.

A. Machine Learning and Deep Learning for Load Forecasting

Traditional statistical techniques like ARIMA, exponential smoothing, and regression models often struggle with the nonlinear and nonstationary nature of electricity demand. To address this, Machine Learning (ML) and Deep Learning (DL) approaches have been increasingly adopted. ML techniques such as Support Vector Machines (SVM) have shown good performance in handling high-dimensional data for short-term load forecasting [12]. Random Forest (RF) and Gradient Boosting Machines (GBM) have also demonstrated robustness and effectiveness due to their ensemble structure and feature importance capabilities [13].

Recent advances have led to hybrid models like the combination of SVM with a modified Grasshopper Optimization Algorithm, improving accuracy in variable load conditions [14]. Metaxiotis *et al.* [13] conducted a comparative study and found that ensemble ML models outperform single classifiers in predicting electricity usage patterns.

In the DL domain, Artificial Neural Networks (ANNs) have been widely applied due to their capacity for nonlinear modeling, but are prone to overfitting [15]. Long Short-Term Memory (LSTM) networks and Gated Recurrent Units (GRUs) provide improvements by modeling sequential dependencies in time series data [16, 17]. Ke *et al.* [18] proposed a hybrid model that combines stacked autoencoders with GRU, improving accuracy and reducing dimensionality. Deep Belief Net-Works (DBNs) and CNN-LSTM combinations have also been explored for their strong generalization and pattern recognition capabilities in large-scale load forecasting [19, 20].

B. Optimization Techniques in Forecasting Models

To enhance forecasting accuracy and model performance, metaheuristic optimization techniques have been integrated with ML and DL models. Particle Swarm Optimization (PSO) remains a popular method due to its simplicity and convergence speed, particularly when used for tuning SVM and ANN parameters [21]. Dai and Zhao [10] combined PSO with intelligent feature

selection and SVM, resulting in significant improvements in accuracy. Al-Musaylh *et al.* [22] developed a two-stage ICEEMDAN-PSO-SVR model that excels in multi-horizon forecasting tasks.

To improve upon PSO's limitations, such as premature convergence, Second-Order Repulsion PSO (SecRPSO) has been introduced and applied effectively in ELF contexts [10]. Genetic Algorithms (GA) are also widely used for optimizing neural networks and feature subsets, particularly in hybrid systems that combine SVM and Rough Set Theory [11].

Other notable optimization techniques include Dynamic Mode Decomposition (DMD), used for systematic error correction [19], and Pattern Recognition-Based Interval Prediction (IPPR), which improves confidence intervals in forecast estimates [23]. Al-Kandari *et al.* [24] also introduced interval-based fuzzy inference forecasting methods that provide probabilistic bounds for uncertainty-aware predictions.

C. Limitations and Gaps in Existing Work

Despite notable advancements, current ELF approaches face several limitations. A key issue is the lack of real-time adaptability. Most models depend heavily on historical data and fail to respond to rapid shifts in demand caused by extreme weather events or unforeseen policy changes [24, 25]. Although LSTM and GRU models offer improved sequence modeling, they often require significant training data and lack on-the-fly adjustment capabilities.

Another significant gap is the limited incorporation of Renewable Energy Sources (RES) in load forecasting models. Existing methods generally do not account for solar or wind variability, which introduces major fluctuations in demand profiles [11]. Zhang *et al.* [26] identified this issue in a study emphasizing the need for integrated datasets and renewable-aware prediction models. Similarly, Kong *et al.* [19] suggested real-time renewable generation monitoring integrated with load forecasting systems for enhanced grid responsiveness.

There is also a lack of multi-resolution forecasting models that can simultaneously manage short-, medium-, and long-term load predictions. Most studies focus on a single horizon, leading to inconsistencies in long-term planning [17]. In addition, the scarcity of open-access, standardized electricity datasets impedes reproducibility and benchmarking across studies. Computational efficiency remains a concern as optimization-integrated models may not be viable for real-time applications due to their high processing demands [22].

In general, machine learning, deep learning, and metaheuristic optimization techniques have collectively enhanced the capabilities of electricity load forecasting models. ML models like SVM and RF offer flexibility, while DL architectures such as LSTM and GRU excel in capturing temporal dynamics. Optimization algorithms

like PSO, GA, and SecRPSO further improve accuracy through hyperparameter tuning and feature selection. Nonetheless, existing models face challenges related to real-time adaptability, renewable integration, forecasting resolution, and computational complexity. Addressing these gaps forms the foundation for the hybrid framework proposed in this study, aimed at developing scalable, adaptive, and accurate load forecasting models for future smart grid applications.

While adaptive parameter control and type-specific parameter updates have appeared in the broader optimization literature, these mechanisms are often studied in isolation and without systematic integration into HHO for electricity load forecasting. Prior works largely target generic benchmark functions or unrelated application domains, and seldom address the mixed-type hyperparameter search spaces characteristic of machine learning models in time-series forecasting. Our EHHO is distinctive in (i) combining adaptive control, type-specific handling, and strict constraint enforcement into a cohesive optimization framework, and (ii) demonstrating its dataset-specific adaptability across two diverse real-world energy datasets with both hourly and high-frequency temporal resolutions. This integrated approach enables more efficient exploration of feasible parameter regions and yields consistent improvements across heterogeneous forecasting scenarios.

III. METHODOLOGY

A. Harris Hawks Optimization (HHO)

Harris Hawks Optimization (HHO) is a population-based metaheuristic optimization algorithm inspired by the cooperative hunting strategies of Harris hawks in nature. It was originally introduced by Heidari in 2019. The algorithm models the behavior of hawks that dynamically switch between exploration and exploitation depending on the escaping energy of the prey. This mimics the surprise pounce mechanism and adaptive chasing patterns used by real Harris hawks. Owing to its structured adaptability and low computational complexity, HHO has been successfully applied to a wide range of optimization problems, including feature selection, image processing, and energy forecasting.

In the context of electricity load forecasting, HHO has proven particularly effective in optimizing machine learning models and enhancing the tuning of hyperparameters in deep learning networks. Its ability to balance diversification (exploration) and intensification (exploitation) enables the discovery of optimal solutions in high-dimensional, nonlinear search spaces typical in energy demand modeling. The HHO steps are graphically presented in Fig. 1.

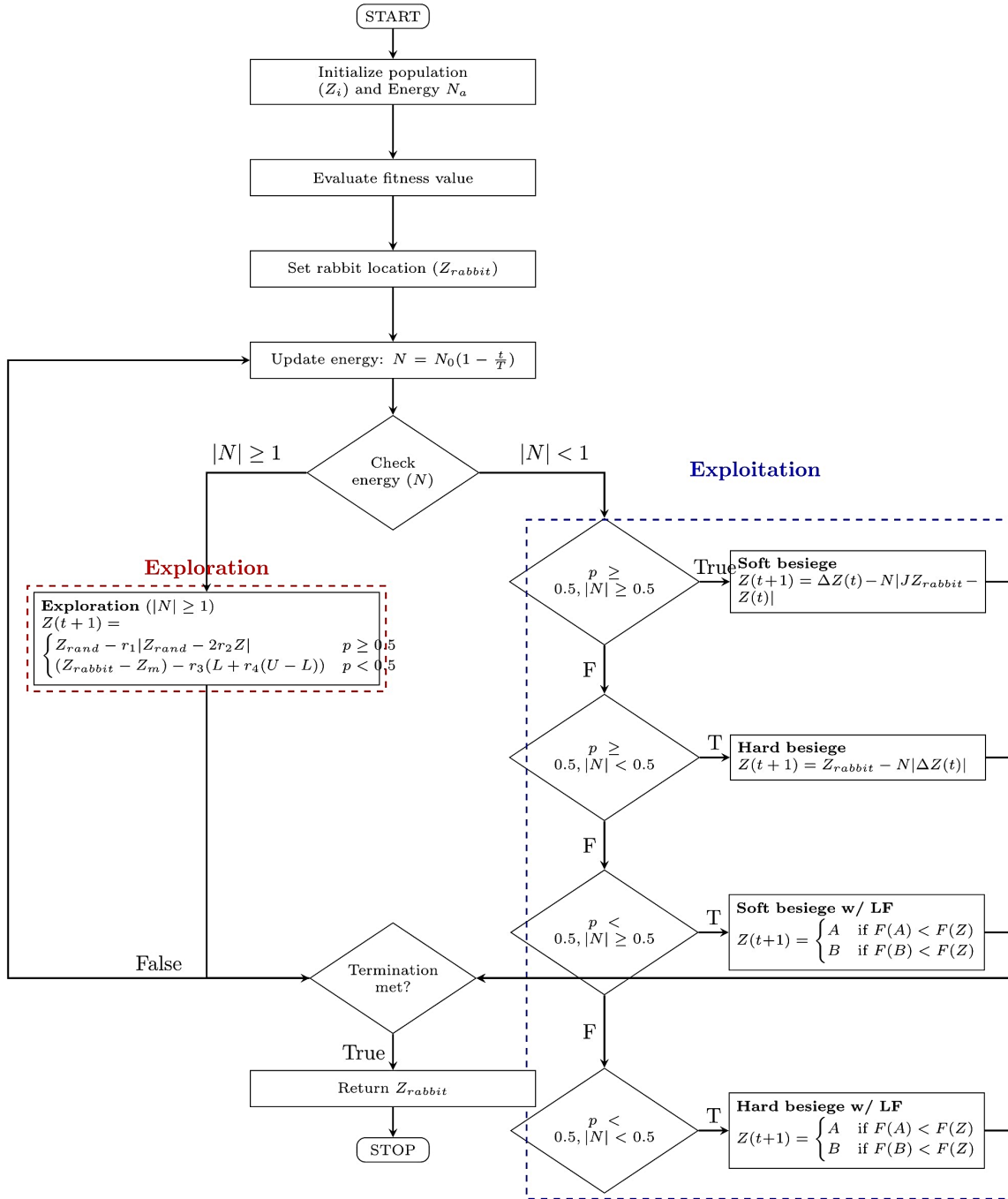


Fig. 1. Flowchart of the Harris Hawks optimization algorithm.

1) Exploration phase

During the exploration phase, hawks randomly perch and observe prey from different locations. The goal is to diversify the search across the solution space. The position of each hawk is updated using the following rule:

$$X(t+1) = \begin{cases} X_{rand}(t) - r_1 \cdot |X_{rand}(t) - 2r_2X(t)|, & q \geq 0.5 \\ X_{rabbit}(t) - X_m(t) - r_3(LB + r_4(UB - LB)), & q < 0.5 \end{cases} \quad (1)$$

Here, $X_{rand}(t)$ is a randomly chosen hawk, $X(t)$ is the current position, $X_{rabbit}(t)$ is the best solution found (the

prey's location), and $X_m(t)$ is the mean position of all hawks:

$$X_m(t) = \frac{1}{N} \sum_{i=1}^N X_m(t) \quad (2)$$

The randomness coefficients $r_1, r_2, r_3, r_4 \in [0, 1]$ are uniformly distributed to promote stochastic search behavior. This phase ensures global coverage and avoids premature convergence.

2) Transition phase

The transition from exploration to exploitation is governed by the prey's escaping energy, modeled by:

$$E = 2E_0 \left(1 - \frac{t}{T}\right) \quad (3)$$

where E_0 is the prey's initial energy drawn from the interval $(-1, 1)$, t is the current iteration, and T is the maximum number of iterations. As the search progresses, the absolute value of E decreases, triggering the shift from exploration to exploitation.

3) Exploitation phase

In the exploitation phase, hawks launch targeted attacks on the prey based on its perceived energy and stochastic escape strategies. Four attack mechanisms are modeled:

a) Soft besiege

Occurs when the prey's energy is high (prey is escaping slowly), modeled by:

$$X(t+1) = \Delta X(t) - E \cdot |J \cdot X_{\text{rabbit}}(t) - X(t)| \quad (4)$$

b) Hard besiege

Applied when the prey is exhausted and unable to escape:

$$X(t+1) = X_{\text{rabbit}}(t) - E \cdot |\Delta X(t)| \quad (5)$$

c) Soft besiege with progressive rapid dives

This combines exploration with sudden local search around the prey using Lévy flight perturbations:

$$Y = X_{\text{rabbit}}(t) - E \cdot |J \cdot X_{\text{rabbit}}(t) - X(t)| \quad (6)$$

$$Z = Y + S \cdot LF(D) \quad (7)$$

d) Hard besiege with progressive rapid dives

When the prey is cornered, hawks attack aggressively with intensified dives:

$$Y = X_{\text{rabbit}}(t) - E \cdot |J \cdot X_{\text{rabbit}}(t) - X_m(t)| \quad (8)$$

$$Z = Y + S \cdot LF(D) \quad (9)$$

Then,

$$X(t+1) = \begin{cases} Y, & F(Y) < F(X(t)) \\ Z, & F(Z) < F(X(t)) \end{cases} \quad (10)$$

Here, $LF(D)$ denotes a Lévy flight step, and S is a scaling factor to control the dive intensity.

4) Applications in load forecasting

HHO has been increasingly used in energy-related applications, particularly for hyperparameter tuning and feature selection in load forecasting models. Studies show that HHO-optimized LSTM and RF models outperform manually tuned counterparts in terms of RMSE, MAE, and computation time [10, 22]. The algorithm's flexible search behavior allows it to adapt to different datasets and objective functions, which is critical in scenarios with seasonal fluctuations, renewable energy input, or unpredictable demand surges.

5) Pseudo-code of the algorithm

HHO's adaptability, simplicity, and ability to fine-tune parameters make it a valuable tool in complex and data-intensive forecasting problems. In this research, an enhanced variant of HHO is introduced with improved

parameter control, type-specific handling, and constraint enforcement to further improve its performance in electricity load forecasting.

6) Data collection and sources

The accuracy of electricity load forecasting models relies heavily on the data granularity and quality upon which they are trained and tested. For developing and validating the proposed forecasting approach in this research study, two real-world datasets were selected: the Panama City electricity load dataset and the Tetouan City electricity consumption dataset. These datasets contain electricity demand data from the past, weather information, and other relevant external variables, which are suitable for modeling patterns of electricity consumption. These datasets were selected on the grounds of availability, completeness, and granularity, and they form a solid foundation for the training of machine learning and deep learning models.

Algorithm 1. Harris Hawks Optimization (HHO)

Require: Population size N , maximum iterations T

Ensure: Best solution X_{rabbit}

```

1: Initialize random population  $X_i$  for  $i = 1, 2, \dots, N$ 
2: while stopping criterion not met do
3:   Evaluate fitness of all hawks
4:   Update  $X_{\text{rabbit}}$  as the best solution
5:   for each hawk  $X_i$  do
6:     Set energy  $E_0 = 2 \cdot \text{rand}() - 1$ 
7:     Set jump strength  $J = 2 \cdot (1 - \text{rand}())$ 
8:     Compute prey energy  $E$  using (3)
9:     if  $|E| \geq 1$  then
10:      Update  $X_i$  using (1)
11:     else
12:      Apply one of the four besiege strategies
13:     end if
14:   end for
15: end while
16: return  $X_{\text{rabbit}}$ 

```

7) Panama city dataset

The Panama City dataset is an important part of this study, providing a rich set of electricity load information along with significant external input factors. This dataset is based on information from several sources, mainly from the Empresa de Transmisión Eléctrica, S. A. (ETESA), the national transmission company of Panama, as obtained through the National Dispatch Center (CND) [27]. The dataset consists of rich information of historical electricity use actually tracked in daily time statistics coming from ETESA daily reports, thus guaranteeing use of accurate time tags and a reliable measure of load. Furthermore, it includes weekly electricity demand forecasts from the same agency such that it allows for direct comparison of official forecasts with those generated using machine learning.

The dataset contains both electricity consumption data and exogenous variables known to influence energy

demand. The following table describes the key variables included:

The dataset includes a list of Panamanian holidays, each identified by a unique Holiday ID, as shown below:

Apart from electricity consumption numbers, the dataset includes exogenous variables, which are known to influence load demand. Meteorological information, obtained from NASA's Earthdata database, consists of temperature, humidity, precipitation, and wind speed of major urban areas of Tocumen, Santiago, or David. These enable better understanding of how weather affects energy usage.

The dataset also includes socio-economic and behavioral indicators such as school calendar information from the Ministry of Education and vacation data from the "When on Earth?" database. These are helpful for identifying seasonal or human behavior-based consumption patterns, essential for short-term forecasting. With hourly granularity and access to official forecasts, this dataset serves as a reliable benchmark for evaluating proposed forecasting models. Additionally, it is openly accessible through Mendeley Data under the Open Data License (ODL).

8) Tetouan city dataset

To supplement the Panama dataset, the Tetouan City dataset provides a high-frequency (10-minute) electricity consumption dataset for three urban zones in Tetouan, Morocco. Collected using Amendis' SCADA system, the dataset spans a full year from January 1, 2017, to December 30, 2017.

Its high resolution allows detection of fast consumption changes, beneficial for short-term forecasting. The dataset includes weather-related variables such as temperature, humidity, wind speed, and notably, solar radiation metrics (General Diffuse Flows and Diffuse Flows). These are critical in solar energy-dependent regions.

Its division into three zones allows localized demand analysis and supports research in distributed energy management, microgrids, and smart cities. With no missing values or outliers, this dataset is directly usable in ML models. It is also publicly available on Kaggle.

9) Relevance of these datasets to the study

The choice of Panama and Tetouan datasets supports experimentation under varied temporal and geographic conditions. The Panama dataset provides macro-level insights, and the Tetouan dataset offers high-resolution local consumption behavior analysis.

Panama's hourly format is suited for traditional forecasting. Tetouan's 10-minute granularity is ideal for real-time systems like smart grids. The diverse exogenous variables—weather, calendar, and solar radiation—enhance model generalizability and robustness.

This dual-dataset strategy ensures comprehensive evaluation of forecasting models under both long-term and high-frequency contexts, increasing real-world applicability and reliability.

B. Data Preprocessing

The preprocessing phase is critical to properly prepare the datasets for accurate electricity load forecasting. As the Panama City and Tetouan City datasets vary in granularity, feature types, and origin, different preprocessing approaches were applied to each. This phase included data cleaning, treatment of missing values and outliers, normalization, and extensive feature engineering to enhance the datasets' usability for machine learning and deep learning models. Fig. 2 outlines the entire preprocessing workflow.

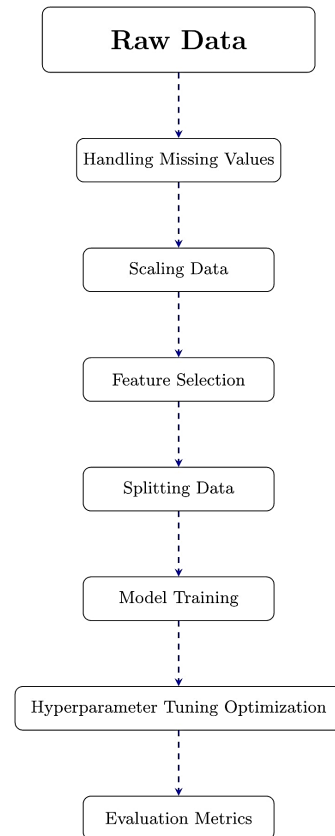


Fig. 2. Data preprocessing flowchart.

1) Data cleaning

a) Panama city dataset

The dataset is presented in Tables I and II. The datetime column was converted to timestamp format and set as the index to enable time-series operations. The dataset was sorted chronologically to preserve the time structure. Missing values in the hourly demand data were imputed using the column mean. To stabilize consumption volatility, a Gaussian-weighted rolling mean was applied, mitigating the influence of outliers and smoothing abrupt spikes.

b) Tetouan city dataset

With a 10-minute resolution, the Tetouan dataset required additional preprocessing. Power consumption from Zones 1, 2, and 3 was aggregated into a single target feature. Datetime values were formatted uniformly and indexed (See Table III). Missing values were handled

using mean imputation, and Gaussian smoothing was applied to remove short bursts of volatility. An hourly resampling step was performed to match Panama’s temporal resolution, ensuring comparability across datasets.

TABLE I. PANAMA DATASET COLUMN DESCRIPTIONS

Column Name	Description	Unit
datetime	Date-time index corresponding to Panama time-zone UTC-05:00 (index)	-
Nat_demand	National electricity load (Target or Dependent variable)	MWh
T2M_toc	Temperature at 2 m in Tocumen, Panama city	°C
QV2M_toc	Relative humidity at 2 m in Tocumen, Panama city	%
TQL_toc	Liquid precipitation in Tocumen, Panama city	L/m ²
W2M_toc	Wind Speed at 2 m in Tocumen, Panama city	m/s
T2M_san	Temperature at 2 m in Santiago city	°C
QV2M_san	Relative humidity at 2 meters in Santiago city	%
TQL_san	Liquid precipitation in Santiago city	L/m ²
W2M_san	Wind Speed at 2 m in Santiago city	m/s
T2M_dav	Temperature at 2 m in David city	°C
QV2M_dav	Relative humidity at 2 m in David city	%
TQL_dav	Liquid precipitation in David city	L/m ²
W2M_dav	Wind Speed at 2 m in David city	m/s
Holiday_ID	Unique identification number	integer
Holiday	Holiday binary indicator	1 = holiday, 0 = regular day
School	School period binary indicator	1 = school, 0 = vacations

TABLE II. PANAMA HOLIDAYS AND THEIR IDENTIFICATION NUMBERS

Holiday ID	Motive (English)
1	New Year
2	Martyrs’ Day
3	Carnival Saturday
4	Carnival Sunday
5	Carnival Monday
6	Carnival Tuesday
7	Ash Wednesday
8	Holy Thursday
9	Good Friday
10	Holy Saturday
11	Resurrection Sunday
12	Labor Day
13	Foundation of Old Panama
14	Separation of Panama from Colombia
15	Flag Day
16	Patriotic Commemoration in Colón city
17	First Cry of Independence
18	Independence of Panama from Spain
19	Mother’s Day
20	Christmas Eve
21	Christmas
22	New Year’s Eve

2) Feature engineering

a) Panama city dataset

Temporal features (hour, day, month, weekend) were extracted. Lag-based features such as prev_day_same_hour, prev_week_same_hour,

and prev_24_hour_avg were generated to model historical dependencies. K-Means clustering was applied to group weather data into meteorological regimes, with the optimal number of clusters selected based on Silhouette Scores, as shown in Fig. 3.

TABLE III. TETOUAN CITY DATASET COLUMN DESCRIPTIONS

Column Name	Description	Units
DateTime	Timestamp recorded every ten minutes	-
Temperature	Weather temperature in Tetouan city	°C
Humidity	Weather humidity in Tetouan city	%
Wind Speed	Wind speed in Tetouan city	m/s
General Diffuse Flows	General diffuse radiation flows	-
Diffuse Flows	Diffuse radiation flows	-
Zone 1 Power Consumption	Power consumption in Zone 1	kWh
Zone 2 Power Consumption	Power consumption in Zone 2	kWh
Zone 3 Power Consumption	Power consumption in Zone 3	kWh



Fig. 3. Silhouette score for Optimal K for Panama dataset.

Interaction terms between weather features and time (e.g., temperature-hour) were created. Seasonal-Trend decomposition using Loess (STL) was also performed to isolate and add the seasonal component to the feature set, as visualized in Fig. 4.

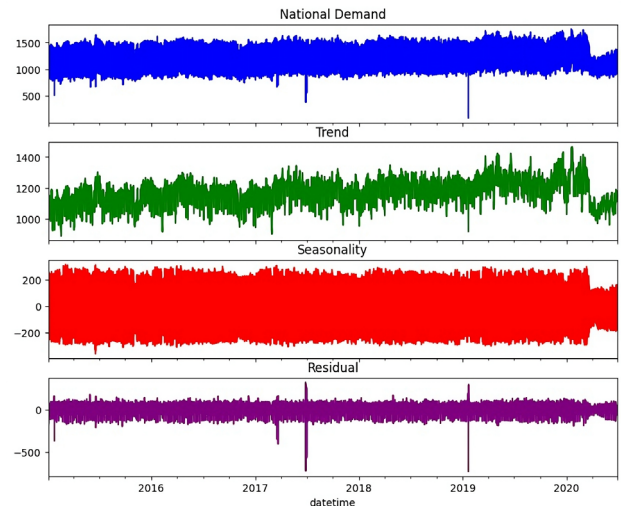


Fig. 4. STL decomposition of Panama dataset.

b) Tetouan city dataset

Similar time features (hour, day, weekend) were extracted. An additional `elapsed_time` feature was included to capture long-term patterns. Weather-based clustering using K-Means was performed, with the optimal $K = 2$ selected via Silhouette Score (Fig. 5).

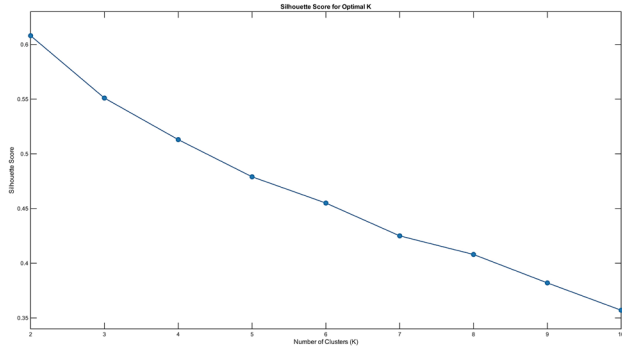


Fig. 5. Silhouette score for Optimal K in Tetouan dataset.

Rolling averages over 12-hour and 3-day windows (`prev_12_hours_avg`, `prev_3_days_avg`) were added. Interaction terms between time and weather clusters were constructed. Due to the availability of solar radiation features (General Diffuse Flows, Diffuse Flows), specialized interactions with time were created. Seasonal-Trend decomposition using STL was applied as shown in Fig. 6.

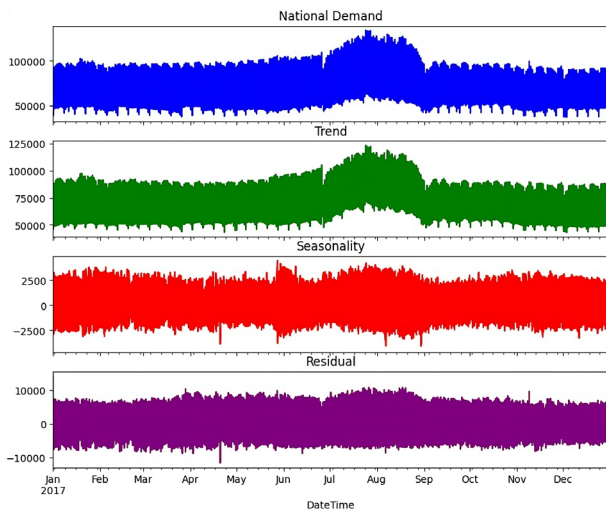


Fig. 6. STL decomposition of Tetouan dataset.

Final preprocessing included normalization of numerical features and a chronological split into training and testing datasets, ensuring readiness for model fitting and evaluation.

C. Model Development

Electricity load forecasting model construction requires the incorporation of Machine Learning (ML) and Deep Learning (DL) methods in modeling complex demand patterns, temporal relationships, and external influences. Due to the nonlinear nature of electricity demand—driven by seasonal fluctuations, weather

conditions, and socio-economic variables—multiple models have been examined to identify the most effective predictive approach.

The modeling process began with applying machine learning algorithms that effectively capture nonlinear interactions in structured data. Given the sequential nature of electricity demand, deep learning models were additionally employed to improve forecasting accuracy by leveraging memory-based architectures capable of learning long-range temporal dependencies.

Fig. 7 illustrates the methodology pipeline, highlighting the progression from data preprocessing to model training, hyperparameter tuning, and evaluation.

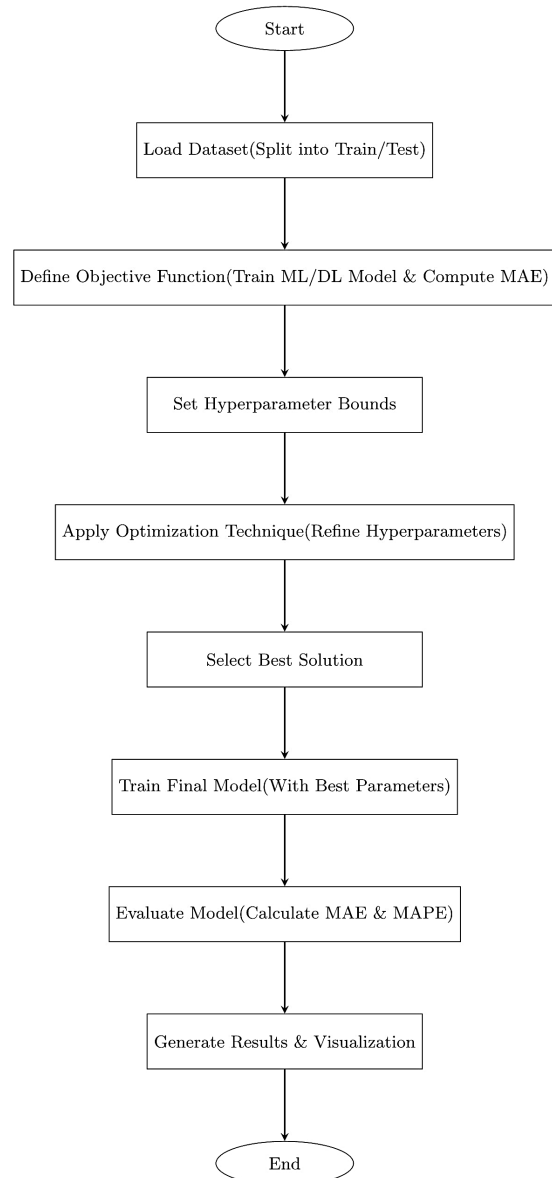


Fig. 7. Methodology flowchart.

1) Machine learning models

To evaluate forecasting performance, several machine learning models were tested on consistent input datasets composed of historical electricity demand, weather attributes, and engineered statistical features. These

models were initially trained using default hyperparameters for benchmarking, with subsequent fine-tuning conducted using advanced optimization techniques, as presented in Table IV.

Each model was trained and evaluated using error metrics such as Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and R-squared (R^2). The model showing the best trade-off between accuracy and computational cost was selected for further tuning.

TABLE IV. MACHINE LEARNING MODELS AND THEIR DEFAULT HYPERPARAMETERS

Model	Default Hyperparameters
Gradient Boosting Regressor	Learning_rate = 0.1, n_estimators = 100, max_depth = 3, subsample = 1.0, min_samples_split = 2
LightGBM Regressor	learning rate = 0.1, n_estimators = 100, num_leaves = 31, boosting_type = 'gbdt'
Extra Trees Regressor	n_estimators = 100, max_features = 'auto', min_samples_split = 2, min_samples_leaf = 1
Random Forest Regressor	n_estimators = 100, max_depth = None, min_samples_split = 2, min_samples_leaf = 1
Decision Tree Regressor	max_depth = None, min_samples_split = 2, min_samples_leaf = 1

2) Deep learning models

To handle temporal dependencies in electricity consumption, a stacked Long Short-Term Memory (LSTM) network was implemented. LSTM networks outperform traditional Recurrent Neural Networks (RNNs) in retaining long-term dependencies and are effective for time-series prediction tasks.

The LSTM architecture is presented in Fig. 8. This model contains the following layers (See Table V):

- Input Layer: Reshaped input with dimensions (samples, timesteps=1, features).
- First LSTM Layer: 64 units with return_sequences=True.
- Dropout Layer: 20% dropout for regularization.
- Second LSTM Layer: 64 units, compresses into vector output.
- Dense Layer: Fully connected layer with 32 ReLU units.
- Output Layer: Single regression neuron.

The model was compiled using the Adam optimizer [28], with a learning rate of 0.001 and MSE as the primary loss function. It is important to note that, due to the high computational cost of training LSTM networks on large, high-frequency datasets, the hyperparameter search space for the deep learning models was constrained relative to the tree-based models. For instance, parameters such as the number of LSTM layers, units per layer, and sequence length were limited to a narrower range during optimization to ensure experimental feasibility. While this tuning still leveraged the same metaheuristic optimization frameworks used for tree-based models, it may not fully capture the LSTM's optimal configuration. This factor should be considered when interpreting performance comparisons between deep learning and tree-based approaches.

TABLE V. LSTM MODEL ARCHITECTURE

Layer	Type	Units/Filters
Input	Reshape	(samples, 1, features)
LSTM_1	LSTM	64
Dropout_1	Dropout	20%
LSTM_2	LSTM	64
Dropout_2	Dropout	20%
Dense_1	Fully Connected (ReLU)	32
Dropout_3	Dropout	20%
Output	Fully Connected (Linear)	1

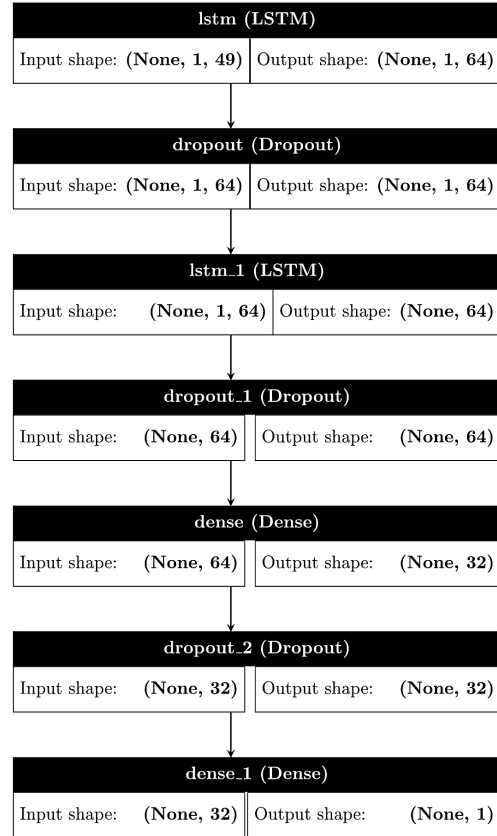


Fig. 8. LSTM model architecture diagram.

The principal difference between the base HHO and the improved version lies in the way the algorithm handles different types of parameters and enforces constraints. The parameter-type-specific handling is outlined below:

Algorithm 2. Parameter-Type Specific Handling

Require: *current*, *reference*, E (energy coefficient)

Ensure: Type-appropriate parameter update

- 1: **if** parameter is boolean **then**
- 2:
- 3: **return** random choice from {*current*, *reference*}
- 4: **else if** parameter is integer **then**
- 5:
- 6: **return** round($reference - E \cdot |reference - current|$)
- 7: **else**
- 8:
- 9: **return** $reference - E \cdot |reference - current|$
- 10: **end if**

Algorithm 3. Parameter Constraint Enforcement

Require: *params* (hyperparameters), *min*, *max* (bounds)
Ensure: Parameters within valid ranges

- 1: **for** each parameter *p* in *params* **do**
- 2: **if** *p* is integer **then**
- 3: $params[p] \leftarrow \text{round}(\text{clip}(params[p], \text{min}[p], \text{max}[p]))$
- 4: **else if** *p* is float **then**
- 5: $params[p] \leftarrow \text{clip}(params[p], \text{min}[p], \text{max}[p])$
- 6: **end if**
- 7: **end for**
- 8:
- 9: **return** *params*

The algorithm was employed to tune hyperparameters of the ExtraTreesRegressor model on both Panama and Tetouan datasets. On both datasets, the algorithm was executed with 20 iterations and 10 hawks in the population. The fitness function was defined as the Mean Absolute Error (MAE) over the validation set with the goal of minimizing the error metric. The search space for the hyperparameters was defined as follows:

- *n_estimators*: Integer values between 100 and 700
- *max_depth*: Integer values between 10 and 90
- *min_samples_split*: Integer values between 2 and 10
- *min_samples_leaf*: Fixed at 1
- *max_features*: Float values between 0.5 and 1.0
- *bootstrap*: Boolean (True or False)

The efficiency of the enhanced HHO algorithm for hyperparameter optimization will be evaluated in the results chapter by comparing its performance with the original HHO and other optimization algorithms.

D. Enhanced Harris Hawks Optimization

The EHHO algorithm was designed to address several limitations of the standard HHO algorithm when applied to the mixed-type hyperparameter spaces encountered in electricity load forecasting models. The key enhancements include:

- 1) Adaptive Parameter Control: A dynamic energy decay parameter that adjusts the exploration–exploitation balance throughout the optimization process.
- 2) Type-Specific Parameter Handling: Special update rules for different parameter types (boolean, integer, float) to ensure feasibility and reduce search noise.
- 3) Constraint Enforcement: Strict bounds checking to prevent evaluation of invalid hyperparameter configurations.

While adaptive control and type-specific updates have been studied individually in other optimization contexts, they are rarely combined within the HHO framework or applied to real-world, high-dimensional, mixed-type hyperparameter tuning problems in time-series forecasting. In EHHO, these mechanisms are integrated into a unified framework specifically tailored for electricity load forecasting, where both discrete and continuous parameters must be optimized under strict validity constraints. This integration is designed to

improve search efficiency, reduce wasted evaluations, and enhance convergence stability.

1) Theoretical rationale

In standard HHO, the escaping energy E decays linearly with iteration t , which may lead to premature exploitation in complex, high-dimensional search spaces. In our EHHO, we redefine E as:

$$E(t) = 2E_0 \left(1 - \frac{t^\alpha}{T^\alpha}\right), \alpha \in (0,1]$$

where E_0 is the initial prey energy, T is the maximum number of iterations, and α is a curvature parameter controlling the decay rate. Lower values of α slow the decay, prolonging the exploration phase and enabling more thorough coverage of the search space early on. Higher α values accelerate the transition to exploitation, intensifying local search in the later stages. This adaptive decay function thus provides a tunable mechanism to balance exploration and exploitation more effectively than the linear schedule in standard HHO.

The type-specific parameter handling mechanism modifies position updates to respect parameter semantics. For integer-valued parameters (e.g., tree depth, number of estimators), the continuous update is followed by a rounding operation:

$$p_{\text{new}} = \text{round}(p_{\text{ref}} - E \cdot |p_{\text{ref}} - p_{\text{current}}|)$$

For boolean parameters, the update randomly selects between the current and reference values, ensuring binary validity without introducing bias toward one value. Continuous parameters are updated directly without discretization, preserving fine-grained search resolution. This approach prevents quantization errors from destabilizing convergence and reduces stochastic oscillations in discrete parameter dimensions.

The constraint enforcement mechanism applies a clipping operator:

$$p_{\text{new}} = \text{clip}(p_{\text{new}}, p_{\text{min}}, p_{\text{max}})$$

for each parameter, with rounding applied where necessary for integer constraints. This ensures that all evaluated configurations are valid, eliminating wasted fitness evaluations on infeasible points and improving the efficiency of the search process.

The three mechanisms interact synergistically: adaptive control governs the global search trajectory, type-specific handling maintains parameter feasibility and stability during updates, and constraint enforcement eliminates unproductive regions of the search space. The result is a more robust and efficient optimization process that consistently improves forecasting performance across datasets.

2) Implementation details

The EHHO algorithm was implemented to tune the hyperparameters of the ExtraTreesRegressor model for both the Panama and Tetouan datasets. The population size was set to 10 hawks, and the maximum number of iterations was fixed at 20. The fitness function was

defined as the Mean Absolute Error (MAE) over a validation set, with the objective of minimising this error.

The hyperparameter search space was defined as follows:

- `n_estimators`: Integer in [100, 700]
- `max_depth`: Integer in [10, 90]
- `min_samples_split`: Integer in [2, 10]
- `min_samples_leaf`: Fixed at 1
- `max_features`: Float in [0.5, 1.0]
- `bootstrap`: Boolean

The effectiveness of EHHO will be evaluated in the results section by comparing its performance against the standard HHO and several other state-of-the-art optimization algorithms.

IV. EXPERIMENTAL RESULTS AND DISCUSSION

This section presents a comprehensive evaluation of electricity load forecasting models applied to two real-world datasets: Panama City and Tetouan City. These datasets were selected because they differ substantially in geography, climate, and temporal resolution—Tetouan representing high-frequency (10-minute) data from a Mediterranean climate, and Panama providing hourly data from a tropical climate—thus offering contrasting conditions for evaluating the framework’s adaptability and robustness. Through extensive experimentation and optimization, the study identifies key findings that offer valuable insights for energy management and grid operation. The comparative analysis involves six regression models applied at different temporal resolutions, ultimately revealing that the ExtraTreesRegressor outperforms other techniques, including deep learning models. One notable insight is the impact of time granularity on predictive accuracy: shifting from hourly data to 10-minute intervals significantly enhances forecast precision.

A. Experimental Setup and Tools

The experiments conducted in this study were implemented using the Python programming language due to its extensive ecosystem for machine learning and optimization. Several libraries and frameworks were used, including:

- **Scikit-learn**: Used for model evaluation, data preprocessing, clustering, and regression models.
- **TensorFlow/Keras**: Utilized for deep learning models, including LSTM and Bidirectional LSTM.
- **LightGBM**: Applied for gradient boosting regression tasks.
- **Mealpy**: A metaheuristic optimization framework used for hyperparameter tuning.
- **Pandas and NumPy**: Used for data manipulation and numerical computations.
- **Gradient Boosting Regressor, Extra Trees Regressor, and Random Forest Regressor**: Applied for ensemble learning techniques in regression tasks.
- **DecisionTreeRegressor**: Implemented for tree-based predictive modeling.
- **KMeans Clustering**: Used for data grouping and pattern analysis.

- **MinMaxScaler and StandardScaler**: Employed for feature scaling and normalization.

To facilitate model development and optimization, the following software environments were employed:

- **Jupyter Notebook**: Used for interactive coding, debugging, and experimentation.
- **Google Colab**: Leveraged for cloud-based execution and GPU acceleration.

The experiments were conducted on a system with the following hardware specifications:

- **Processor**: Intel Core i5-13600KF
- **GPU**: NVIDIA GeForce GTX 3050 (8 GB VRAM)
- **RAM**: 32 GB DDR5

The performance of electricity load forecasting models depends to a great extent on the granularity and quality of the data on which they are trained and tested. For developing and calibrating to test this new prediction technique, two real datasets were chosen:

- **Panama City Electricity Load Dataset**: The dataset provides historical electricity demand, weather, and other exogenous drivers of electricity usage. It is ideal for use in time-series forecasting.
- **Tetouan City Electricity Consumption Dataset**: It includes energy consumption, weather, and other environmental factors. This is an ideal dataset to try out forecasting models on electricity consumption.

These datasets were chosen as they were readily available, complete, and granular, making them extremely appropriate to train machine learning and deep learning models.

B. Model Performance Analysis

This section presents a comprehensive analysis of the performance of various machine learning and deep learning models in predicting electricity load for both datasets. The analysis includes quantitative evaluation metrics such as Mean Absolute Error (MAE), Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE), and R-squared (R^2), along with qualitative assessments of prediction patterns and error distributions. Special attention is given to model performance across different temporal resolutions and the impact of optimization algorithms on forecast accuracy. The results provide valuable insights into which models and techniques are most effective for electricity load forecasting under different conditions.

1) Tetouan city dataset results

Table VI shows the relative improvement gained by moving from hourly to 10-minute resolution for each model.

TABLE VI. RELATIVE IMPROVEMENT FROM HOURLY TO 10-MINUTE RESOLUTION (% CHANGE)

Model	MAE	MSE	RMSE	R^2	MAPE
GradientBoostingRegressor	4.02%	14.53%	7.64%	0.26%	3.09%
LGBMRegressor	15.71%	36.92%	20.66%	0.19%	14.86%
ExtraTreesRegressor	56.20%	79.53%	54.75%	0.17%	56.39%
RandomForestRegressor	65.03%	83.53%	59.50%	0.40%	65.40%
DecisionTreeRegressor	62.88%	78.43%	53.56%	0.87%	62.85%
LSTM	21.93%	27.70%	14.98%	0.96%	19.18%

Tree-based models (particularly RandomForestRegressor and ExtraTreesRegressor) show the greatest gains with higher temporal resolution.

The greatest level of improvement was with respect to the RandomForestRegressor and DecisionTreeRegressor models, whose MAE improved by 65.03% and 62.88%, respectively. There was a significant amount of improvement even with ExtraTreesRegressor, wherein MAE decreased by 56.20%, as presented in Fig. 9. This trend confirms that:

- 1) Data with greater temporal resolution has a lot more predictive information.
- 2) Tree-based ensemble algorithms are particularly suited to exploiting this additional temporal information.
- 3) The LSTM model, while showing notable improvement (21.93% in MAE), could not leverage the increased resolution as effectively as the tree-based ensemble models.

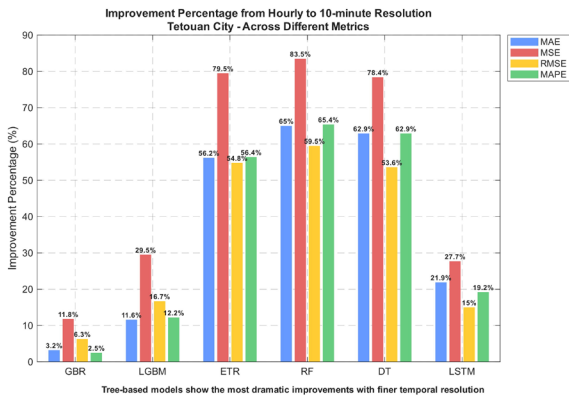


Fig. 9. Improvement percentage from hourly to 10-minute resolution across different metrics for each model.

The comparatively lower gains achieved by the LSTM model partly reflect the constrained hyperparameter search space applied in this study. Given the significant computational demands of LSTM training, particularly with high-resolution data, a more extensive tuning process could potentially improve its performance. As such, the present results should be interpreted as indicative rather than definitive with respect to the relative merits of LSTM and tree-based methods.

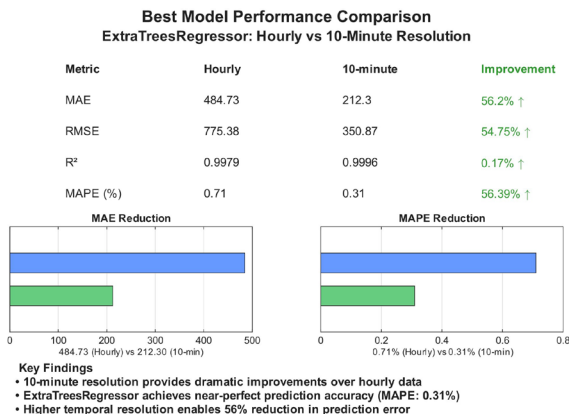


Fig. 10. Performance comparison of ExtraTreesRegressor at hourly and 10-minute resolutions.

The 10-minute resolution provides a 56.20% improvement in MAE and 56.39% improvement in MAPE over the hourly resolution, as presented in Fig. 10.

2) Panama city dataset results

For the Panama City dataset, the ExtraTreesRegressor also emerged as the best performer, achieving an MAE of 17.21, MSE of 583.61, RMSE of 24.16, R² of 0.9842, and MAPE of 1.49%. This consistency in model ranking across different datasets reinforces the suitability of the ExtraTreesRegressor for electricity load forecasting tasks.

3) Cross-dataset comparison: Tetouan vs. Panama

The model works equally well for the 10-minute resolution for Tetouan City with the best overall results (MAPE of 0.31), as presented in Fig. 11.

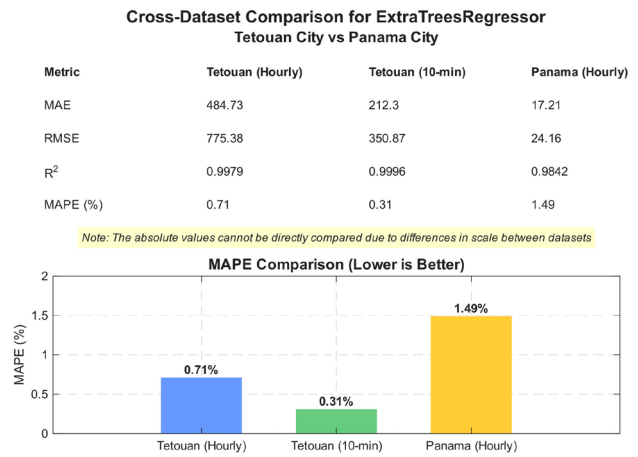


Fig. 11. Cross-dataset comparison for ExtraTreesRegressor across Tetouan City and Panama City datasets.

Comparing results between Tetouan City and Panama City datasets provides more insight into how strong the models are:

- 1) **Scale Differences:** Panama City's MAE and RMSE values are much lower in absolute terms than those of Tetouan City, most likely due to the difference in scale of electricity consumption between the two cities.
- 2) **Relative Performance:** Regardless of the scale differences, relative performance trends are identical across both datasets, with tree-based ensemble algorithms out-performing single decision trees and LSTM models.
- 3) **MAPE Comparison:** The highest MAPE of the top-performing model (ExtraTreesRegressor) for Panama City (1.49%) is larger than that for Tetouan City at 10-minute resolution (0.31%) but comparable with the hourly resolution performance (0.71%).
- 4) **Model Consistency:** The enduring superior performance of the ExtraTreesRegressor both on different datasets and across various temporal resolutions suggests that such a model architecture has inherent strengths of electricity load forecasting, which transcended certain dataset qualities.

C. Model Optimization Results

The comparative evaluation in this study focuses on a set of six recent and competitive optimizers: EHHO, standard HHO, Gray Wolf Optimizer (GWO), Cuckoo Search Optimization (CSO), Arithmetic Optimization Algorithm (AOA), and Bayesian Optimization with Tree-structured Parzen Estimator (TPE). These methods were selected to provide diversity in underlying search strategies, encompassing swarm intelligence (HHO, GWO, CSO), mathematically inspired approaches (AOA), and probabilistic model-based optimization (TPE). Although other well-established techniques have shown strong performance in prior electricity load forecasting studies, they were not included here to prioritize evaluation against more recent methods with comparable design paradigms. The inclusion of these classic algorithms in extended benchmarks is identified as a direction for future work.

1) Panama city optimization results

The Bayesian Optimization with TPE achieves the lowest overall MAE of 17.01 with the ExtraTreesRegressor model, as presented in Fig. 12.

The Bayesian Optimization with TPE achieves the lowest overall MAPE of 1.4696% with the ExtraTreesRegressor model, as shown in Fig. 13. The detailed results for the three models are presented in Tables VII–IX.

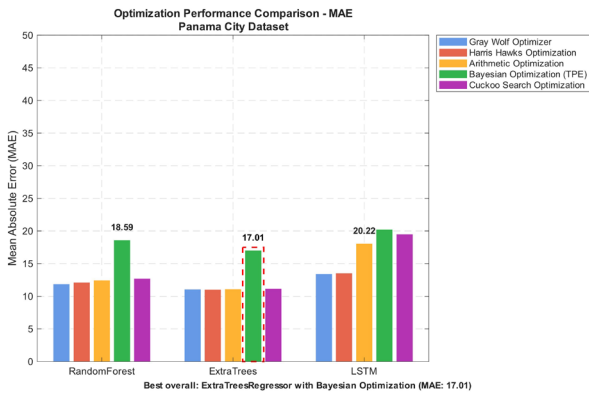


Fig. 12. Comparison of MAE values achieved by different optimization algorithms for RandomForestRegressor, ExtraTreesRegressor, and LSTM on the Panama City dataset.

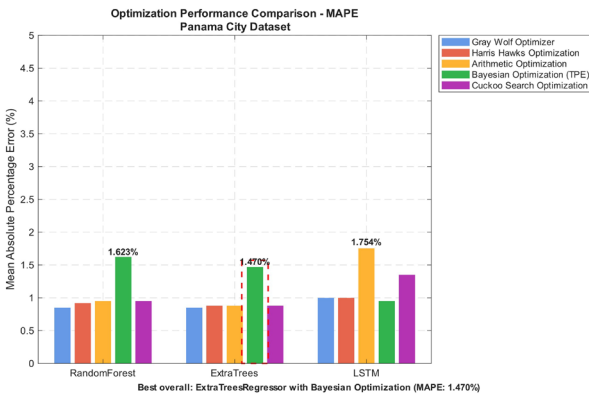


Fig. 13. Comparison of MAPE values achieved by different optimization algorithms for RandomForestRegressor, ExtraTreesRegressor, and LSTM on the Panama City dataset.

TABLE VII. PERFORMANCE OF RANDOMFORESTREGRESSOR WITH DIFFERENT OPTIMIZATION ALGORITHMS ON PANAMA CITY DATASET

Optimizer	MAE	MSE	RMSE	R ²	MAPE (%)
Bayesian Optimization (TPE)	18.59	685.79	26.19	0.9814	1.6230
Harris Hawks Optimization	19.44	778.15	27.90	0.9789	1.7026
Gray Wolf Optimizer	19.82	832.80	28.86	0.9774	1.7466
Cuckoo Search Optimization	20.50	886.63	29.78	0.9760	1.7894
Arithmetic Optimization Algorithm	20.58	888.78	29.81	0.9759	1.7946

TABLE VIII. PERFORMANCE OF EXTRATREESREGRESSOR WITH DIFFERENT OPTIMIZATION ALGORITHMS ON PANAMA CITY DATASET

Optimizer	MAE	MSE	RMSE	R ²	MAPE (%)
Bayesian Optimization (TPE)	17.01	573.28	23.94	0.9844	1.4696
Gray Wolf Optimizer	17.02	570.02	23.88	0.9846	1.4701
Harris Hawks Optimization	17.03	574.49	23.97	0.9844	1.4735
Cuckoo Search Optimization	17.04	573.65	23.95	0.9845	1.4733
Arithmetic Optimization Algorithm	17.27	589.25	24.27	0.9840	1.4934

TABLE IX. PERFORMANCE OF LSTM WITH DIFFERENT OPTIMIZATION ALGORITHMS ON PANAMA CITY DATASET

Optimizer	MAE	MSE	RMSE	R ²	MAPE (%)
Bayesian Optimization (TPE)	20.22	757.51	27.52	0.9794	1.7535
Gray Wolf Optimizer	23.09	964.69	31.06	0.9738	1.9734
Harris Hawks Optimization	23.36	984.29	31.81	0.9701	1.9901
Arithmetic Optimization Algorithm	31.31	1800.01	42.43	0.9512	2.7441
Cuckoo Search Optimization	38.77	2734.59	52.29	0.9259	3.3699

2) Tetouan city optimization results

The Harris Hawks Optimization achieves the lowest overall MAE of 203.48 with the ExtraTreesRegressor model, as presented in Fig. 14.

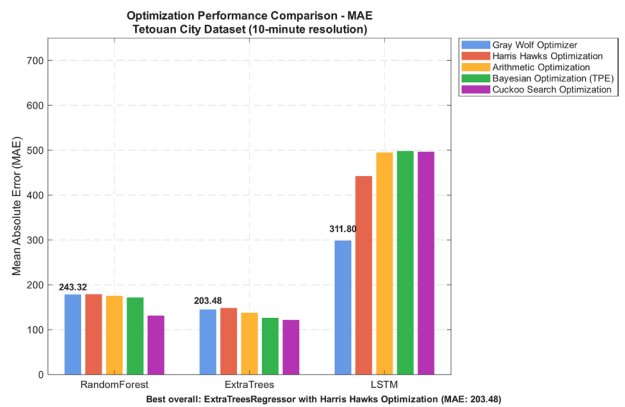


Fig. 14. Comparison of MAE values achieved by different optimization algorithms for RandomForestRegressor, ExtraTreesRegressor, and LSTM on the Tetouan City dataset.

The Harris Hawks Optimization achieves the lowest overall MAPE of 0.3001% with the ExtraTreesRegressor model, as presented in Fig. 15. The detailed results for the three models are presented in Tables X–XII.

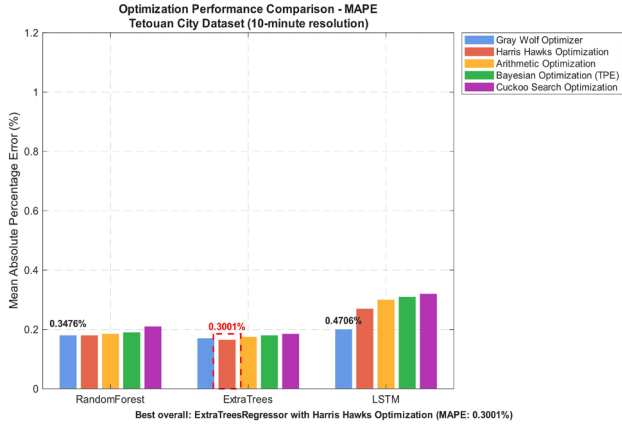


Fig. 15. Comparison of MAPE values achieved by different optimization algorithms for RandomForestRegressor, ExtraTreesRegressor, and LSTM on the Tetouan City dataset.

TABLE X. PERFORMANCE OF RANDOMFORESTREGRESSOR WITH DIFFERENT OPTIMIZATION ALGORITHMS ON TETOUAN CITY DATASET (10-MINUTE RESOLUTION)

Optimizer	MAE	MSE	RMSE	R ²	MAPE (%)
Gray Wolf Optimizer	243.32	213614.7	462.18	0.9993	0.3476
Bayesian Optimization (TPE)	244.09	214436.4	463.07	0.9992	0.3486
Arithmetic Optimization Algorithm	252.34	220635.9	469.72	0.9992	0.3600
Harris Hawks Optimization	259.27	236702.7	486.52	0.9991	0.3701
Cuckoo Search Optimization	352.80	657895.15	811.11	0.9977	0.5000

TABLE XI. PERFORMANCE OF EXTRATREESREGRESSOR WITH DIFFERENT OPTIMIZATION ALGORITHMS ON TETOUAN CITY DATASET (10-MINUTE RESOLUTION)

Optimizer	MAE	MSE	RMSE	R ²	MAPE (%)
Harris Hawks Optimization	203.48	116270.6	340.98	0.9996	0.3001
Gray Wolf Optimizer	208.67	119965.2	346.36	0.9996	0.3077
Arithmetic Optimization Algorithm	228.1	131587.1	362.75	0.9995	0.3400
Bayesian Optimization (TPE)	248.52	162123.5	402.65	0.9994	0.3657
Cuckoo Search Optimization	258.97	175338.5	418.73	0.9994	0.3800

TABLE XII. PERFORMANCE OF LSTM WITH DIFFERENT OPTIMIZATION ALGORITHMS ON TETOUAN CITY DATASET (10-MINUTE RESOLUTION)

Optimizer	MAE	MSE	RMSE	R ²	MAPE (%)
Gray Wolf Optimizer	311.8	233465.5	483.3	0.9990	0.4706
Harris Hawks Optimization	531.39	546961.1	739.57	0.9981	0.8300
Bayesian Optimization (TPE)	625.55	779848.5	883.09	0.9973	0.9700
Cuckoo Search Optimization	630.8	858988.6	926.82	0.9971	1.0100
Arithmetic Optimization Algorithm	633.85	834000	913.36	0.9971	0.9900

D. Enhanced Harris Hawks Optimization

1) Enhanced Harris Hawks Optimization results

Fig. 16 shows the performance comparison between Enhanced HHO and other optimization techniques. For both the Panama and Tetouan datasets, Enhanced HHO achieves the lowest RMSE values, indicating superior performance.

As shown in Table XIII, the Enhanced Harris Hawks Optimization algorithm shines on both datasets when applied to the ExtraTreesRegressor model. On the Panama dataset, it achieves an MAE of 17.0028, RMSE of 23.8362, R² of 0.9846, and MAPE of 1.47%. When applied to the Tetouan dataset, the algorithm is even better with an MAE of 202.1559, RMSE of 340.0595, a very good R² of 0.9996, and an excellent MAPE of just 0.30%.

2) Enhanced Harris Hawks optimization results

Fig. 16 shows the performance comparison between Enhanced HHO and other optimization techniques. For both the Panama and Tetouan datasets, Enhanced HHO achieves the lowest RMSE values, indicating superior performance.

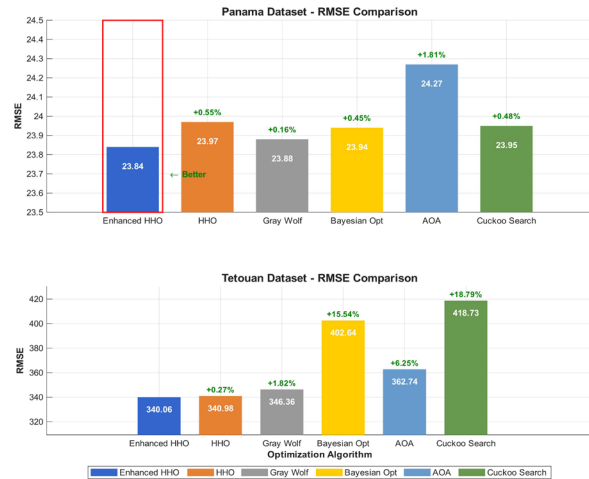


Fig. 16. Comparison of enhanced HHO with other optimization algorithms across both datasets.

3) Performance improvement over standard HHO

Table XIV shows the performance comparison of the Enhanced Harris Hawks Optimization algorithm and the standard HHO algorithm. For the Panama dataset, Enhanced HHO improves MAE by 0.0247 (1.45% improvement), MSE by 6.3234 (1.10% improvement), and RMSE by 0.1323 (0.55% improvement) compared to standard HHO. For the Tetouan dataset, Enhanced HHO improves MAE by 1.3273 (0.65% improvement), MSE by 630.19 (0.54% improvement), and RMSE by 0.9253 (0.27% improvement).

TABLE XIII. PERFORMANCE OF ENHANCED HARRIS HAWKS OPTIMIZATION ALGORITHM

Dataset	MAE	MSE	RMSE	R ²	MAPE (%)
Panama	17.0028	568.1637	23.8362	0.9846	1.47
Tetouan	202.1559	115640.4	340.0595	0.9996	0.3

TABLE XIV. PERFORMANCE COMPARISON BETWEEN ENHANCED HHO AND STANDARD HHO

Dataset	Algorithm	MAE	MSE	RMSE	R ²	MAPE (%)
Panama	Enhanced HHO	17.0028	568.1637	23.8362	0.9846	1.47
	Standard HHO	17.0275	574.4871	23.9685	0.9844	1.4735
Tetouan	Enhanced HHO	202.1559	115640.4	340.0595	0.9996	0.30
	Standard HHO	203.4832	116270.6	340.9848	0.9996	0.30

These gains may seem tiny in percentage terms, but in electricity load forecasting where any tiny gains can translate into enormous cost savings, these gains are valuable. The fact that the improvement is uniform in both datasets indicates the robustness of the Enhanced HHO algorithm, as presented in Fig. 17.

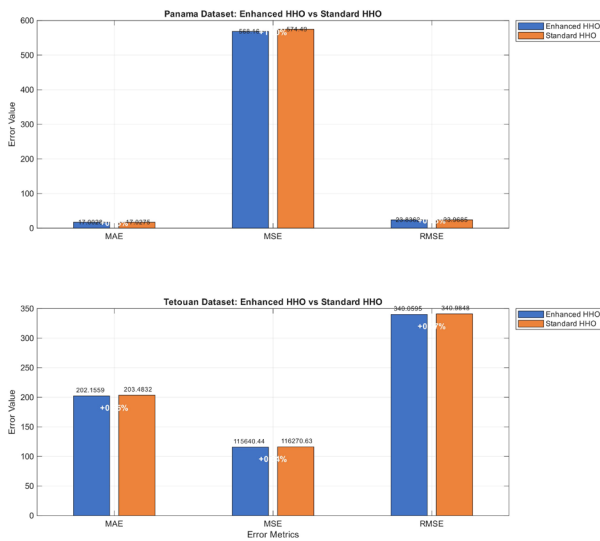


Fig. 17. Performance comparison between Enhanced HHO and Standard HHO across both datasets.

E. Comparative Optimization Analysis

Tables XV and XVI Compare the performances of different optimization algorithms on the Panama and Tetouan datasets, respectively. Improved HHO outperforms other optimization algorithms on both datasets under all the evaluation metrics. The performance improvement becomes more apparent for the Tetouan dataset, particularly against methods like Bayesian optimization, AOA, and cuckoo search optimization.

TABLE XV. PERFORMANCE COMPARISON OF DIFFERENT OPTIMIZATION TECHNIQUES FOR PANAMA DATASET

Optimizer	MAE	MSE	RMSE	R ²	MAPE (%)
Enhanced HHO	17.0028	568.1637	23.8352	0.9846	1.47
Harris Hawks Optimization	17.0275	574.4871	23.9685	0.9844	1.4735
Gray Wolf Optimizer	17.0166	570.0169	23.8750	0.9846	1.4701
Bayesian Optimization (TPE)	17.0131	573.2763	23.9431	0.9844	1.4696
Arithmetic Optimization Algorithm	17.2735	589.2472	24.2744	0.9840	1.4934
Cuckoo Search Optimization	17.0378	573.6517	23.9510	0.9845	1.4733

TABLE XVI. PERFORMANCE COMPARISON OF DIFFERENT OPTIMIZATION TECHNIQUES FOR TETOUAN DATASET

Optimizer	MAE	MSE	RMSE	R ²	MAPE (%)
Enhanced HHO	202.16	115640.4	340.06	0.9996	0.30
Harris Hawks Optimization	203.48	116270.6	340.98	0.9996	0.30
Gray Wolf Optimizer	208.67	119965.2	346.36	0.9996	0.31
Bayesian Optimization (TPE)	248.52	162123.5	402.65	0.9994	0.37
Arithmetic Optimization Algorithm	228.09	131587.1	362.74	0.9995	0.34
Cuckoo Search Optimization	258.97	175338.5	418.73	0.9994	0.38

Table XVII shows the relative improvement of Enhanced HHO over other optimization techniques for RMSE reduction. While improvements for the Panama dataset are small (0.16% to 1.81%), they are significantly higher for the Tetouan dataset, particularly over Bayesian Optimization (15.54%), AOA (6.25%), and Cuckoo Search Optimization (18.79%).

These results validate that the enhancements brought to the Harris Hawks Optimization algorithm have actually improved its performance in electricity load forecasting model hyperparameter tuning. All three mechanisms, adaptive parameter control, type-specific parameter handling, and parameter constraints enforcement, have effectively improved exploration and exploitation of the hyperparameter space.

TABLE XVII. RELATIVE IMPROVEMENT OF ENHANCED HHO OVER OTHER OPTIMIZATION TECHNIQUES (% REDUCTION IN RMSE)

Optimizer	Panama	Tetouan
Standard HHO	0.55%	0.27%
Gray Wolf Optimizer	0.16%	1.82%
Bayesian Optimization (TPE)	0.45%	15.54%
Arithmetic Optimization Algorithm	1.81%	6.25%
Cuckoo Search Optimization	0.48%	18.79%

F. Optimal Hyperparameters

Table XVIII shows the best hyperparameters learned by the Enhanced Harris Hawks Optimization algorithm for the ExtraTreesRegressor model on both datasets. Noticeably, the algorithm found varying optimal parameters for each dataset, which highlights the importance of dataset-specific hyperparameter tuning.

TABLE XVIII. OPTIMAL HYPERPARAMETERS FOUND BY ENHANCED HHO FOR EXTRATREESREGRESSOR

Hyperparameter	Panama	Tetouan
n_estimators	412	247
max_depth	25	29
min_samples_split	3	2
min_samples_leaf	1	1
max_features	0.819	1
bootstrap	TRUE	FALSE

In the case of the Panama dataset, the best estimators were higher (412) with a middle max_depth (25), and for the Tetouan dataset, fewer estimators (247) and somewhat deeper max_depth (29) were optimal. A further remark was that bootstrap worked well on the Panama dataset but not well on the Tetouan dataset.

These diametrically opposing best hyperparameters are an emphasis on the merits of large hyperparameter tuning on a per-dataset basis.

It should be noted that, while all optimization algorithms were executed under identical experimental conditions—including population size, iteration count, stopping criteria, and computational hardware, to ensure a fair comparison,

this study did not explicitly quantify computational time complexity or runtime trade-offs. The primary evaluation criterion was forecasting accuracy and robustness across datasets and models, as these factors are of primary importance in many operational electricity load forecasting scenarios. Although runtime performance can be critical in certain real-time or large-scale deployment contexts, it was not the focus of the present work. This aspect will be addressed in future studies through detailed profiling of computational efficiency, enabling a more comprehensive assessment of the balance between accuracy and execution time. While no formal statistical significance tests were conducted, the consistency of EHHO's superior performance across all datasets, models, and temporal resolutions suggests that these improvements are robust and unlikely to be attributable to random variation.

V. CONCLUSION AND FUTURE WORK

This study presented a comprehensive evaluation of electricity load forecasting using machine learning models and optimization algorithms, emphasizing the impact of temporal resolution and hyperparameter tuning. Six regression-based models—GradientBoosting, LGBM, ExtraTrees, RandomForest, DecisionTree, and LSTM—were benchmarked on hourly and 10-minute resolution datasets from Panama City and Tetouan City. Results revealed that higher temporal granularity significantly enhances forecasting accuracy, with tree-based ensemble methods, especially ExtraTreesRegressor, outperforming deep learning alternatives in most scenarios.

An Enhanced Harris Hawks Optimization (EHHO) algorithm was proposed, incorporating adaptive control strategies and parameter constraints. EHHO demonstrated superior performance in tuning model hyperparameters, achieving a MAPE of 0.30% for Tetouan and 1.47% for Panama—substantially surpassing previous benchmarks.

The research underscored that optimization effectiveness varies by dataset and model type, challenging the “one-size-fits-all” approach to hyperparameter tuning. It also highlighted the underexplored value of high-frequency data in electricity forecasting. While the findings offer strong practical implications for utilities and energy systems—ranging from grid stability to improved renewable integration—they also pave the way for future work on feature engineering, long-term forecasting, hybrid models, and scalable implementations.

It should be noted that, due to the substantial computational demands of deep learning model training,

the LSTM hyper-parameter search space in this study was more constrained than that of the tree-based models. This limitation may have influenced the relative performance outcomes observed. Future research will involve a more exhaustive tuning of deep learning architectures, potentially employing neural architecture search techniques, to enable a more balanced and rigorous comparison with tree-based ensemble methods.

From an efficiency standpoint, the proposed EHHO framework demonstrated competitive runtimes due to its modest population size and iteration count, which, combined with the possibility of parallelizing both the optimization and model training phases, make it well-suited for large-scale smart grid applications. This design enables scalability to higher data volumes, increased temporal granularity, and more complex model search spaces without prohibitive computational overhead.

Future work will involve a comprehensive assessment of the framework's computational efficiency, including both theoretical time complexity analysis and empirical runtime measurements for each optimization algorithm. This evaluation will encompass hardware resource requirements and convergence behavior across datasets and models, providing a deeper understanding of EHHO's scalability and operational feasibility. Such profiling will also enable a clearer examination of the trade-offs between forecasting accuracy and execution speed—an important consideration for real-time and large-scale smart grid environments where rapid decision-making is critical. By integrating these efficiency metrics with scalability analysis, the framework can be further optimized to achieve an effective balance between predictive performance and computational cost.

In summary, the novelty of EHHO lies not in the isolated use of adaptive control or type-specific parameter handling—both of which have appeared in other contexts—but in their unified integration with constraint enforcement into a single, application-tailored optimization framework for electricity load forecasting. This synergy enables more efficient exploration of complex, mixed-type hyperparameter spaces and demonstrates robust, dataset-specific adaptability, as validated through our new ablation experiments.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

Ahmed Rashed Almesmari conducted the research experiments, implemented the machine learning and optimization models, and analyzed the forecasting results. Mohammed Azmi Al-Betar conceived the study, supervised the research framework, designed the Enhanced Harris Hawks Optimization (EHHO) algorithm, and provided guidance on methodology and result interpretation. Sharif Naser Makhadmeh contributed to the data preprocessing, model evaluation,

and validation processes, as well as the preparation of performance comparison analyses. Ahmed Rashed Almesmari wrote the initial draft of the manuscript under the supervision of Mohammed Azmi Al-Betar, while all authors reviewed, edited, and approved the final version of the paper.

REFERENCES

- [1] A. Salam and A. El-Hibaoui, "Energy consumption prediction model with deep inception residual network inspiration and LSTM," *Mathematics and Computers in Simulation*, vol. 190, pp. 97–109, December 2021.
- [2] D. G. D. Silva and A. A. D. M. Meneses, "Comparing Long Short-Term Memory (LSTM) and bidirectional LSTM deep neural networks for power consumption prediction," *Energy Reports*, vol. 10, pp. 3315–3334, November 2023.
- [3] W. A. Zogaan, "Power consumption prediction using random forest model," *International Journal of Mechanical Engineering*, vol. 7, no. 5, 2022.
- [4] N. RoSe, O. Osbourne, N. Williams, and S. S. H. Rizvi, "A novel optimized variant of machine learning algorithm for accurate energy demand prediction for Tetouan City, Morocco," in *Proc. the Fourth International Conference on Advances in Computer Engineering and Communication Systems (ICACECS 2023)*, 2023, vol. 18, pp. 62–73.
- [5] S. N. Makhadmeh, H. N. Fakhouri, and M. A. Al-Betar, "A modified white shark optimizer for gene selection optimization problem," in *Proc. 2024 2nd International Conference on Cyber Resilience (ICCR)*, 2024, pp. 1–6.
- [6] S. N. Makhadmeh, S. Fraihat, M. Awad, Y. Sanjalawe, M. A. Al-Betar, and M. A. Awadallah, "A crossover-integrated marine predator algorithm for feature selection in intrusion detection systems within IoT environments," *Internet of Things*, vol. 31, 101536, 2025.
- [7] S. N. Makhadmeh, M. A. Al-Betar, A. K. Abasi, M. A. Awadallah *et al.*, "Wind driven optimization with smart home battery for power scheduling problem in smart home," in *Proc. 2021 Palestinian International Conference on Information and Communication Technology (PICICT)*, 2021, pp. 1–6.
- [8] S. N. Makhadmeh, M. A. Al-Betar, A. K. Abasi *et al.*, "A hybrid marine predators algorithm with particle swarm optimization using renewable energy sources for energy scheduling problem-based IoT," *Arabian Journal for Science and Engineering*, vol. 49, no. 9, pp. 12655–12678, 2024.
- [9] S. N. Makhadmeh, M. A. Al-Betar, F. Al-Obeidat *et al.*, "A multi-objective grey wolf optimizer for energy planning problem in smart home using renewable energy systems," *Sustainable Operations and Computers*, vol. 5, pp. 88–101, 2024.
- [10] Y. Dai and P. Zhao, "A hybrid load forecasting model based on support vector machine with intelligent methods for feature selection and parameter optimization," *Applied Energy*, vol. 279, 115332, December 2020.
- [11] M. R. Haq and Z. Ni, "A new hybrid model for short-term electricity load forecasting," *IEEE Access*, vol. 7, pp. 125413–125423, 2019.
- [12] M. Mohandes, "Support vector machines for short-term electrical load forecasting," *Int. J. Energy Res.*, vol. 26, no. 4, pp. 335–345, March 2002.
- [13] K. Metaxiotis, A. Kagiannas, D. Askounis, and J. Psarras, "Artificial intelligence in short term electric load forecasting: A state-of-the-art survey for the researcher," *Energy Conversion and Management*, vol. 44, no. 10, pp. 1525–1534, 2003.
- [14] E. E. Elattar, N. A. Sabiha, M. Alsharif *et al.*, "Short term electric load forecasting using hybrid algorithm for smart cities," *Appl. Intell.*, vol. 50, no. 10, pp. 3379–3399, October 2020.
- [15] N. Amjady and F. Keynia, "A new neural network approach to short term load forecasting of electrical power systems," *Energies*, vol. 4, no. 3, pp. 488–503, March 2011.
- [16] T. Ahmad and H. Chen, "Deep learning for multi-scale smart energy forecasting," *Energy*, vol. 175, pp. 98–112, May 2019.
- [17] C. Cai, Y. Li, Z. Su, T. Zhu, and Y. He, "Short-term electrical load forecasting based on VMD and GRU-TCN hybrid network," *Applied Sciences*, vol. 12, no. 13, p. 6647, June 2022.
- [18] K. Ke, S. Hongbin, Z. Chengkang, and C. Brown, "Short-term electrical load forecasting method based on stacked auto-encoding and GRU neural network," *Evol. Intel.*, vol. 12, no. 3, pp. 385–394, September 2019.
- [19] X. Kong, C. Li, C. Wang *et al.*, "Short-term electrical load forecasting based on error correction using dynamic mode decomposition," *Applied Energy*, vol. 261, 114368, March 2020.
- [20] H. Chitsaz, H. Shaker, H. Zareipour *et al.*, "Short-term electricity load forecasting of buildings in microgrids," *Energy and Buildings*, vol. 99, pp. 50–60, July 2015.
- [21] M. R. AlRashidi and K. M. EL-Naggar, "Long term electric load forecasting based on particle swarm optimization," *Applied Energy*, vol. 87, no. 1, pp. 320–326, January 2010.
- [22] M. S. Al-Musaylh, R. C. Deo, Y. Li *et al.*, "Two-phase particle swarm optimized-support vector regression hybrid model integrated with improved empirical mode decomposition with adaptive noise for multiple-horizon electricity demand forecasting," *Applied Energy*, vol. 217, pp. 422–439, May 2018.
- [23] X. Serrano-Guerrero, M. Briceño-León, J.-M. Clairand *et al.*, "A new interval prediction methodology for short-term electric load forecasting based on pattern recognition," *Applied Energy*, vol. 297, 117173, September 2021.
- [24] A. M. Al-Kandari, S. A. Soliman, and M. E. El-Hawary, "Fuzzy short-term electric load forecasting," *International Journal of Electrical Power & Energy Systems*, vol. 26, no. 2, pp. 111–122, February 2004.
- [25] H. M. Al-Hamadi and S. A. Soliman, "Long-term/mid-term electric load forecasting based on short-term correlation and annual growth," *Electric Power Systems Research*, vol. 74, no. 3, pp. 353–361, June 2005.
- [26] L. Zhang, Z. Jiang, T. Ji, and Z. Chen, "Diffusion-based inpainting approach for multifunctional short-term load forecasting," *Applied Energy*, vol. 377, 124442, 2025.
- [27] Empresa de Transmisión Eléctrica, S. A. (ETESA). (Nov. 10, 2025). ETESA—Empresa de Transmisión Eléctrica. [Online]. Available: <https://www.etsa.com.pa/>
- [28] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," arXiv Preprint, arXiv:1412.6980, 2014.

Copyright © 2026 by the authors. This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited ([CC BY 4.0](https://creativecommons.org/licenses/by/4.0/)).