



OPEN

Forecasting Renewable energy and electricity consumption using evolutionary hyperheuristic algorithm

Yang Cao¹, Jun Yu², Rui Zhong³✉ & Masaharu Munetomo³

This research utilizes time series models to forecast electricity generation from renewable energy sources and electricity consumption. The configuration of optimal parameters for these models typically requires optimization algorithms, but conventional algorithms may struggle with fixed search patterns and limited robustness. To address this, we propose an auto-evolution hyper-heuristic algorithm named AE-GAPB. AE-GAPB integrates a genetic algorithm (GA) at the high-level component and employs particle swarm optimization (PSO) and the bat algorithm (BA) at the low-level component. The GA continuously finds the best hyperparameters for PSO and BA based on prediction accuracy, which significantly accelerates the optimization and improves the accuracy. Additionally, the crossover and mutation rates of GA evolve over iteration time and fitness value space, further enhancing its adaptability. We validated AE-GAPB on six time series forecasting models and compared it with five well-known optimization algorithms as well as GAPB without auto-evolution at the high-level component. As a result, AE-GAPB achieved excellent results on renewable energy generation and electricity consumption datasets from the Hokkaido, Kyushu, and Tohoku regions of Japan.

Keywords Electricity Forecasting, Renewable Energy, Electricity Consumption, Auto-Evolution Hyper-heuristics, Time Series Models

As areas such as artificial intelligence, electric vehicles, and the electrification of production and manufacturing continue to grow, the electricity demand is increasing¹. Effective electricity planning is essential for achieving energy conservation and emission reduction. In contrast, poor electricity planning can result in various problems, including stalled renewable energy development, fluctuations in wholesale electricity prices, and worsening grid aging challenges². In regions like Europe and the United States, where wind and solar power generation have surged dramatically, there has even been the phenomenon of negative electricity prices due to oversupply³. Thus, accurate electricity forecasting is essential for the proper planning of electricity.

To achieve accurate electricity forecasting, deep learning techniques and time series models are often employed^{4,5}. However, due to the influence of seasonal factors on electricity data, selecting appropriate features using deep learning approaches can be challenging, while time series models are a more commonly used approach⁶. Common time series forecasting models include Simple Exponential Smoothing (SES)⁷, Double Exponential Smoothing (DES)⁸, the Holt-Winters Model (HWM)⁹, Weighted Moving Average (WMA)¹⁰, Auto Regression (AR)¹¹, and Long Short-Term Memory (LSTM)¹². When utilizing these time series models, the setting of initial parameters typically requires user expertise. Previous studies have employed methods such as Maximum Likelihood Estimates¹³, and meta-heuristic algorithms like Particle Swarm Optimization (PSO)⁶ and Bat Algorithm (BA)¹⁴ to automatically determine optimal initial parameters. However, Traditional approaches often rely on fixed hyperparameters, leading to limited adaptability and suboptimal performance across diverse datasets. Furthermore, these methods typically struggle to balance exploration and exploitation, resulting in susceptibility to local optima and reduced efficiency when handling complex or volatile electricity demand data.

To address this issue, we propose a hyper-heuristic algorithm named AE-GAPB (AutoEvolution-Genetic Algorithm-(PSO, BA)) that can automatically evolve with exploration time and fitness value space. The proposed AE-GAPB generally consists of two main parts: a high-level and a low-level¹⁵. In our research, we employ the genetic algorithm (GA) as the high-level component, and PSO and BA as the low-level components. GA can

¹Graduate School of Information Science and Technology, Hokkaido University, Sapporo, Japan. ²Institute of Science and Technology, Niigata University, Niigata, Japan. ³Information Initiative Center, Hokkaido University, Sapporo, Japan. ✉email: zhongrui@iic.hokudai.ac.jp

continuously adjust the hyperparameters of PSO and BA based on forecast accuracy feedback until the optimal hyperparameter combination is found. Meanwhile, GA, as a high-level component, evolves by exploring both time and fitness value spaces. Specifically, The crossover and mutation rate of GA can be dynamically adjusted in real time according to the number of generation cycles and the quality of the fitness values. This enables AE-GAPB to dynamically evolve hyperparameters in real-time, adapt to varying complexities, and deliver consistently superior results across multiple forecasting scenarios. The unique ability of AE-GAPB to address these challenges makes it the most suitable method for electricity forecasting tasks. We validated the proposed approach using six different time series models: SES, DES, HWM, WMA, AR, and LSTM. Experiments were conducted on datasets of renewable energy generation and electricity demand-side consumption in three regions of Japan: Hokkaido, Kyushu, and Tohoku. These datasets were sourced from the Federation of Electric Power Companies (FEPC) (<https://www.fepc.or.jp/index.html>) of Japan. Our proposed approach outperforms other algorithms on six datasets from three regions in Japan, which demonstrates strong performance and robustness. Industry professionals can leverage the difference between renewable energy generation and power demand to optimize fossil fuel generation, optimize the utilization of renewable energy, minimize fossil fuel usage, and lower carbon emissions. We summarize the related studies on electricity forecasting, along with our proposed AE-GAPB method, in Table 1. A detailed introduction to these studies is provided in Section 2.

The main contribution of this work lies in the development of the AE-GAPB, which integrates a high-level Genetic Algorithm (GA) with low-level heuristic algorithms, namely PSO and BA. This innovative approach addresses key limitations of traditional optimization algorithms, such as fixed search patterns and lack of robustness across diverse datasets.

Specifically, the originality of this study is highlighted in the following aspects:

- Dynamic Parameter Evolution:** AE-GAPB introduces a mechanism where GA dynamically adjusts the hyperparameters of PSO and BA in real-time based on fitness evaluations. This ensures continuous adaptability and enhanced optimization accuracy.
- Adaptive Exploration-Exploitation Balance:** By dynamically modifying crossover and mutation rates based on fitness value feedback and iteration progress, the algorithm effectively balances exploration and exploitation, enabling robust performance across different time series models.
- Multi-Level Optimization Framework:** The hierarchical structure of AE-GAPB, with a high-level GA guiding the optimization of low-level heuristics, represents a novel application of hyper-heuristic principles to time series forecasting.
- Practical Validation:** The proposed algorithm is validated on real-world datasets from the Hokkaido, Kyushu, and Tohoku regions of Japan, covering both renewable energy generation and electricity consumption. This demonstrates the algorithm's capability to handle diverse forecasting tasks. The experimental results confirm that AE-GAPB consistently outperforms state-of-the-art optimization algorithms and standard forecasting methods in terms of prediction accuracy and robustness. By bridging gaps in traditional methods and addressing the challenges of forecasting in fluctuating energy systems, this work significantly contributes to the field of electricity demand and renewable energy forecasting.

The rest of this paper is structured as follows: Section 2 presents an overview of related work on electricity consumption prediction and hyper-heuristic algorithms. Section 3 provides a detailed explanation of the proposed hyper-heuristic algorithm. Section 4 covers the experimental setup and outcomes. Section 5 offers an analysis of the experimental findings. Lastly, Section 6 concludes the study and discusses potential future research directions.

Study	Optimization Method	Time Series Models Used	Key Contributions	Limitations
Thejus et al. (2021) ¹⁶	RNN	LSTM	Stacked LSTM showed best performance for power consumption and solar energy prediction; RNN excelled in wind energy forecasting.	Limited scalability for large datasets; model performance depends on hyperparameter tuning.
Elsaraiti et al. (2021) ¹⁷	ACF, PCF	ARIMA	Achieved a MAPE of 4.332% using ACF and PACF for parameter tuning, demonstrating high precision.	ARIMA struggles with highly non-linear data; requires manual analysis for parameter setting.
Deng et al. (2021) ⁶	PSO	SES	PSO-optimized SES demonstrated higher accuracy and speed compared to other methods.	May not generalize well to datasets with seasonality shifts or complex patterns.
Qureshi et al. (2024) ¹⁸	Deep Learning	LSTM	Achieved 95% accuracy in forecasting electricity usage trends in hospitals.	Focused on a specific domain, limiting generalizability to other sectors or applications.
Lambert et al. (2023) ¹⁹	Statistical + Adaptive Methods	Generalized Additive Models (GAM)	Adaptive method balances individual variability and national demand; incorporates transfer learning for scalability.	Computational challenges remain for real-time large-scale implementation.
Bhutta et al. (2024) ²⁰	Hybrid Machine Learning	Hybrid Convolutional-LSTM Net	HCLN significantly improved solar power predictions, enabling real-time optimization of energy systems.	High computational cost; may require specialized hardware for deployment.
Mauricio et al. (2023) ²¹	CSA, GA	Holt-Winters	Both CSA and GA yield similar MSE for HW models, optimizing coefficients for additive and multiplicative smoothing.	No significant difference observed in optimization speed or accuracy between CSA and GA.
Present Work	AE-GAPB	SES, DES, HWM, WMA, AR, LSTM	Proposed dynamic hyper-heuristic framework integrating GA, PSO, and BA, achieving robust performance across datasets.	Potentially higher computational complexity.

Table 1. Comparison of the Present Work with Related Studies.

Related works

Literature review of electricity forecasting

There has been extensive prior research on power forecasting. Thejus et al.¹⁶ focused on forecasting power consumption and generation using deep learning in Python. Models like LSTM and Recurrent Neural Network (RNN) were tested, with the stacked LSTM showed the best performance in predicting power consumption and solar energy, while the RNN excelled in forecasting wind energy. This research utilized datasets from multiple sources, including hourly power consumption and renewable energy generation data. Elsaraiti et al.¹⁷ employed the autoregressive integrated moving average (ARIMA) model to forecast future electricity consumption. By analyzing data with ACF and PACF plots, the model's parameters were set. The results demonstrated the model's high precision, achieving a Mean Absolute Percentage Error (MAPE) of 4.332%, proving it to be a strong competitor to current forecasting methods. Deng et al.⁶ applied seasonal exponential smoothing (SES) models, optimized through the particle swarm optimization (PSO), to forecast electricity consumption. Testing with real data showed that the PSO-based SES model is more accurate and faster than other methods, making it a strong choice for forecasting electricity consumption. Qureshi et al.¹⁸ focused on optimizing models and forecasting electricity use with an LSTM-based approach. Tested on real data from a hospital, their methods accurately predicted electricity trends, achieving a 95% accuracy rate. Lambert et al.¹⁹ concentrated on predicting day-ahead electricity loads for more than a thousand substations across France, addressing the variability of individual consumption while maintaining the stability of national demand. Their approach employs an adaptive method that integrates generalized additive models with state-space representations. To address the computational challenges of scaling this approach, a simplified variant is developed that reduces parameter estimation through expert aggregation and transfer learning. This method effectively balances accuracy and efficiency, while maintaining model interpretability for practical applications. Bhutta et al.²⁰ employed hybrid machine learning models, specifically the Hybrid Convolutional-LSTM Net (HCLN), to forecast the fluctuating output of renewable energy sources such as solar and wind farms. Smart grids equipped with AI systems can optimize energy production and distribution in real-time, enhancing efficiency, reducing emissions, and improving energy security. Their research shows that the HCLN model significantly increases the accuracy of solar power generation predictions, thereby boosting overall system efficiency. And Mauricio et al.²¹ applied the Holt-Winters (HW) forecasting technique to predict transformer overload, comparing the Cuckoo Search Algorithm (CSA) and Genetic Algorithm (GA) for optimizing mean squared error (MSE) and smoothing coefficients. Results showed no significant difference in optimization speed or forecast accuracy between CSA and GA, with both algorithms yielding similar MSE values for the HW additive and multiplicative models.

Based on the previous research, electricity forecasting is primarily approached in two ways: through deep learning and time series models. Deep learning methods face challenges in selecting feature values due to the influence of weather conditions and seasonal factors, both on the generation and consumption sides. As a result, time series models are often preferred for electricity forecasting. To enhance the accuracy of these models, optimization algorithms like PSO and BA are employed to determine the initial parameters. However, traditional optimization algorithms often suffer from fixed hyperparameters and limited robustness across different datasets. To address these issues, a hyper-heuristic algorithm that incorporates multiple operators presents a promising solution and offers improved robustness.

Literature review of hyper-heuristics

Hyper-heuristics are an advanced strategy in heuristic search, developed to streamline and improve the process of selecting, combining, generating, and fine-tuning various basic heuristics²². Unlike conventional approaches that depend on a single heuristic, hyper-heuristics exploit the combined strength of multiple heuristic search operators to generate superior solutions²³. Research has demonstrated that this combined method surpasses the performance of individual heuristics, providing more robust and efficient problem-solving capabilities. Fig.1 illustrates a typical structure of a hyper-heuristic algorithm, highlighting the framework that supports this powerful technique.

Li et al.²⁴ employed neural networks to improve heuristic selection in selection hyper-heuristics, which serve as adaptable solvers for a wide array of optimization challenges. Unlike traditional algorithms, these hyper-heuristics manage a collection of heuristics to solve various problems. They implemented two neural network models, LSTM and TCN, to predict heuristic sequences, training the models using data from 20 existing hyper-heuristics. Evaluated on the 1-dimensional Bin Packing problem, the results indicated that neural networks successfully replicated and unified heuristic selection strategies across different hyper-heuristics. Zhong et al.²⁵ presented an evolutionary status-guided hyper-heuristic algorithm (ES-HHA) for continuous optimization. This algorithm features a low-level component containing operator pools for exploitation and exploration, and a high-level component that employs a probabilistic selection function driven by fitness distance correlation (FDC) and population diversity (PD). These components work in tandem to balance exploration and exploitation while dynamically sequencing heuristics. Extensive experiments across several benchmark tasks, and comparisons with 14 optimization methods, showed that ES-HHA performs competitively, with the probabilistic selection function effectively guiding the optimization process. Vela et al.²⁶ introduced the Squared Hyper-Heuristic (SHH), a new hyper-heuristic model that adds an additional layer of generalization by using hyper-heuristics (HHs) instead of low-level heuristics (LLHs). SHH was validated through a four-stage testing process across various scenarios, demonstrating that, under the right conditions, it outperforms traditional HHs. Its flexibility allows for additional layers, enhancing complexity and customization for diverse training instances, opening the door to a transfer learning approach in hyper-heuristics. Singh et al.²⁷ examined ant-based hyper-heuristics, focusing on how different parameters affect generality, optimality, and computational efficiency. Their study also looked at the potential for unnecessary complexity, or “bloat,” in generated heuristics under different conditions. The research tackled three key problems: the traveling salesman problem, movie scene scheduling, and the

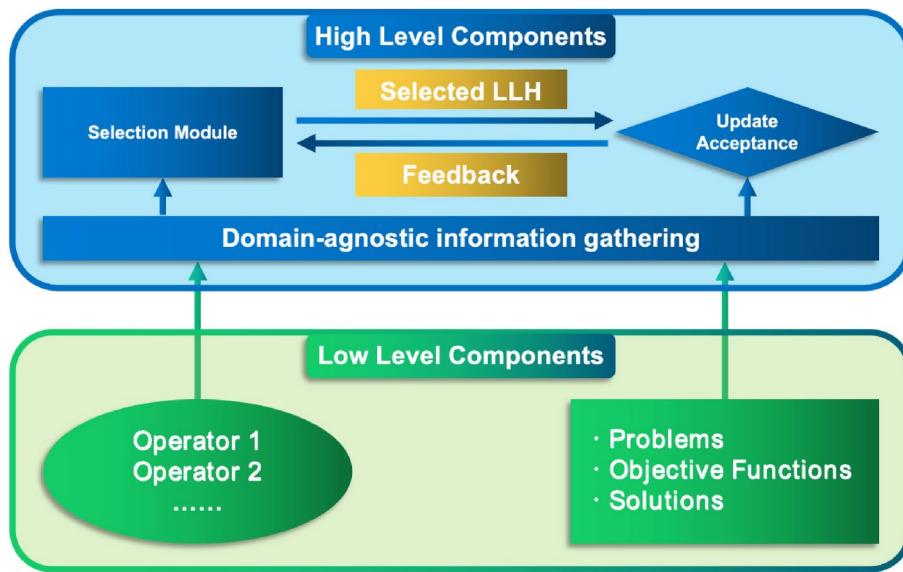


Fig. 1. Composition of the hyper-heuristic algorithm.

capacitated vehicle routing problem. The findings showed that selecting appropriate pheromone maps and operator constraints is crucial for broad applicability, while the number of ants and iterations plays a significant role in efficiently achieving optimal solutions. Their work also revealed that ant-based hyper-heuristics produce compact and streamlined solutions, even with larger operator limits, and are resilient to unnecessary complexity. Arpacı et al.²⁸ introduced sequence-based hyper-heuristics using a hidden Markov model (HMM) alongside various acceptance strategies. Their method was tested on real-world telecommunications network instances, with results indicating that HMM, particularly when using an extended learning period and a threshold acceptance strategy, generates high-quality solutions for large-scale network design problems. Anwar et al.²⁹ explored the many-objective Pickup and Delivery Problem (MaOPDP), employing 15 low-level heuristics (LLHs) during the perturbation and local search phases, optimized using a q-learning-based Hyper-heuristic (HH). They proposed a high-level selection mechanism balancing exploration and exploitation for LLH selection. Their method, validated on benchmark datasets of varying sizes, was compared to state-of-the-art hyper-heuristics and meta-heuristics. Their approach outperformed in most cases, surpassing the leading HH, HH-ILS, by 646.7% in HV and 100% in the Additive Epsilon Indicator (AEI). Yin et al.³⁰ introduced a hyper-heuristic evolutionary algorithm based on proximal policy optimization, termed HHEA-PPO, designed for solving multi-objective truss optimization problems. The algorithm features a two-layer structure: a high-level strategy utilizing proximal policy optimization and low-level heuristics comprising ten predefined heuristic operators. During iterations, the high-level strategy dynamically selects the most suitable low-level heuristic based on the state of individuals and the population. To enhance convergence and maintain the distribution of the external Pareto archive, a dynamic crowding distance mechanism is employed. HHEA-PPO is evaluated on eight multi-objective truss optimization problems and benchmarked against thirteen state-of-the-art optimization algorithms, considering success rate, average computation time, and fitness evaluations. The results demonstrate that HHEA-PPO achieves higher search efficiency and greater stability, showcasing its effectiveness in addressing large-scale engineering design challenges.

The above studies highlight the latest advancements in hyper-heuristic algorithms from various perspectives. Dynamic adaptability and diversity management are exemplified by algorithms such as ES-HHA and ant-based hyper-heuristics, which enhance flexibility and robustness by dynamically adjusting heuristic selection and parameters. The integration of deep learning and reinforcement learning, as demonstrated by the works of Li et al. and Anwar et al., introduces forward-looking tools for heuristic selection, though their computational cost remains relatively high. Additionally, the versatility of hyper-heuristic algorithms in addressing complex problems is emphasized in the studies by Vela et al. and Singh et al., particularly their remarkable performance in large-scale optimization tasks. Yin et al. further contribute to this domain by proposing the HHEA-PPO algorithm, which effectively combines proximal policy optimization with hyper-heuristics to solve multi-objective truss optimization problems. Its dynamic two-layer structure, incorporating a high-level PPO strategy and ten low-level heuristic operators, exemplifies an innovative approach to balancing convergence and solution diversity through dynamic crowding distance mechanisms. The algorithm's demonstrated efficiency and stability in large-scale engineering design challenges highlight its potential as a powerful tool for multi-objective optimization. However, future research should focus on reducing computational complexity, improving the efficiency of hyper-heuristic algorithms in high-dimensional and dynamic environments, and validating their long-term stability in practical applications.

Time series models

This section provides a detailed explanation of the six different time series models used in our experiments.

Simple exponential smoothing (SES)

Simple Exponential Smoothing (SES)³¹ is a fundamental forecasting technique that estimates future values using historical data. It operates by assigning exponentially decreasing weights to prior observations, giving more significance to recent data points compared to older ones. SES is particularly well-suited for forecasting time series data that lacks trends or seasonal patterns. The formula for Simple Exponential Smoothing is as follows:

$$S_t = \alpha \cdot X_t + (1 - \alpha) \cdot S_{t-1} \quad (1)$$

Where:

- S_t is the smoothed value (or forecast) at time t .
- X_t is the actual observed value at time t .
- α represents the smoothing parameter that controls the rate at which the impact of past data decreases over time.
- S_{t-1} represents the smoothed value (or forecast) at time $t - 1$.

Double exponential smoothing (DES)

Double Exponential Smoothing (DES)³² builds upon Simple Exponential Smoothing by adapting it to handle data with trends. Unlike Simple Exponential Smoothing, which only smooths the level of the series, DES addresses trends by applying two separate smoothing components: one for the level and another for the trend. This two-fold approach allows the model to adjust forecasts by accounting for both the current level and the trend, making it more effective for data that consistently follows an upward or downward trajectory over time. The formulas for Double Exponential Smoothing are given as follows:

- **Level Equation:**

$$L_t = \alpha \cdot X_t + (1 - \alpha) \cdot (L_{t-1} + T_{t-1}) \quad (2)$$

- **Trend Equation:**

$$T_t = \beta \cdot (L_t - L_{t-1}) + (1 - \beta) \cdot T_{t-1} \quad (3)$$

- **Forecast Equation:**

$$F_{t+m} = L_t + m \cdot T_t \quad (4)$$

Where:

- L_t represents the level at t .
- T_t represents the trend at t .
- F_{t+m} denotes the forecast for m periods into the future.
- α is the smoothing factor applied to the level.
- β is the smoothing factor applied to the trend.
- X_t represents the actual observed value at t .

Holt-Winters Model (HWM)

The Holt-Winters Model (HWM), also known as Triple Exponential

Smoothing, is a method that builds on the Double Exponential Smoothing model by adding a third component to account for seasonality. This makes it especially effective for data exhibiting both trends and seasonal variations. The model can handle both additive and multiplicative seasonality. The formulas for the Holt-Winters Model are outlined as follows:

- **Equation for the level component:**

$$L_t = \alpha \cdot \frac{X_t}{S_{t-p}} + (1 - \alpha) \cdot (L_{t-1} + T_{t-1}) \quad (5)$$

- **Equation for the trend component:**

$$T_t = \beta \cdot (L_t - L_{t-1}) + (1 - \beta) \cdot T_{t-1} \quad (6)$$

- **Equation for the seasonal component:**

$$S_t = \gamma \cdot \frac{X_t}{L_t} + (1 - \gamma) \cdot S_{t-p} \quad (7)$$

- **Equation for the forecast:**

$$F_{t+m} = (L_t + m \cdot T_t) \cdot S_{t-p+m} \quad (8)$$

Where:

- L_t represents the level component at t .
- T_t denotes the trend component at t .
- S_t represents the seasonal component at t .
- F_{t+m} denotes the prediction for m periods into the future.
- α, β , and γ are the smoothing factors for the level, trend, and seasonal components, respectively.
- p represents the duration of the seasonal cycle.
- X_t represents the actual value observed at t . For the **additive** model, the seasonal component is added rather than multiplied. The corresponding equations would be adjusted as follows:
- **Equation for the seasonal component (Additive):**

$$S_t = \gamma \cdot (X_t - L_t) + (1 - \gamma) \cdot S_{t-p} \quad (9)$$

- **Equation for the forecast component (Additive):**

$$F_{t+m} = L_t + m \cdot T_t + S_{t-p+m} \quad (10)$$

Weight moving averaging (WMA)

The Weighted Moving Average (WMA) is a forecasting technique that applies varying weights to past data points, assigning more significance to certain observations based on their position in the time series. Unlike the Simple Moving Average (SMA), which treats all data points equally, WMA offers greater flexibility by allowing recent observations to have a stronger influence on the forecast, depending on the assigned weights. This makes WMA particularly useful for identifying trends or patterns where recent data is more predictive of future outcomes. The formula for the Weighted Moving Average is given as follows:

$$WMA_t = \frac{\sum_{i=0}^{n-1} w_i \cdot X_{t-i}}{\sum_{i=0}^{n-1} w_i} \quad (11)$$

Where:

- WMA_t is the weighted moving average at t .
- X_{t-i} represents the actual observed values at $t - i$.
- w_i is the weight assigned to the i -th observation.
- n represents the number of observations included in the moving average.

Auto Regression (AR)

The Auto Regression (AR) model³³ is a forecasting technique that uses prior values of a variable to predict future values. In the AR model, the current value is expressed as a linear combination of its previous values, known as lags, along with a random error term. This approach is especially effective when a strong correlation exists between past and present values, making it highly useful in fields such as financial and economic analysis. The general form of the Auto Regression model of order p (AR p) is:

$$X_t = c + \sum_{i=1}^p \phi_i X_{t-i} + \epsilon_t \quad (12)$$

Where:

- X_t represents the value of the time series at t .
- c is an optional constant term.
- ϕ_i represents the coefficients of the model.
- X_{t-i} denotes the previous values (lags) of the time series.
- p is the model's order, indicating the number of lags incorporated.
- ϵ_t represents the error term at t , which is assumed to follow a white noise process.

Long Short-Term Memory (LSTM)

Long Short-Term Memory (LSTM)³⁴ is a distinct architecture within recurrent neural networks (RNNs) designed to process sequential data while overcoming limitations found in traditional RNNs, such as the vanishing gradient problem. LSTMs excel at tasks requiring the capture of long-term dependencies, such as time series forecasting, language modeling, and sequence prediction. Through a system of gates, LSTMs regulate the flow of information, allowing selective retention or removal of past data, which enables learning of complex temporal relationships. The key gates in an LSTM cell are:

- **Forget Gate Component:**

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (13)$$

- **Input Gate Component:**

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (14)$$

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \quad (15)$$

- **Output Gate Component:**

$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (16)$ The cell state is updated in the following manner:

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t \quad (17)$$

The hidden state is updated in the following way:

$$h_t = o_t \cdot \tanh(C_t) \quad (18)$$

- **Weight Update:**

$$W = W - \eta \cdot \nabla L \quad (19)$$

Where:

- x_t represents the input at t .
- h_t denotes the hidden state at t (which is also the output of the LSTM cell).
- C_t is the cell state at t .
- f_t , i_t , and o_t correspond to the forget, input, and output gates, respectively.
- \tilde{C}_t is the candidate cell state.
- W_f , W_i , W_C , W_o are the weight matrices for each gate.
- b_f , b_i , b_C , b_o are the bias terms for the respective gates.
- σ denotes the sigmoid activation function.
- \tanh represents the hyperbolic tangent activation function.
- W is the weight parameter, ∇L is the gradient of the loss function L with respect to that parameter, and η is the learning rate. LSTMs are extensively utilized in deep learning due to their capability to learn from long data sequences and effectively capture temporal dependencies.

Mean Absolute Percentage Error (MAPE)

Mean Absolute Percentage Error (MAPE)³⁵ is a common metric used to assess the accuracy of forecasting models. It measures the average absolute percentage difference between the predicted and actual values, offering a straightforward way to evaluate the model's performance. MAPE is widely favored because it presents the error as a percentage, making it easy to interpret across various datasets³⁶. The formula for MAPE is:

$$\text{MAPE} = \frac{1}{n} \sum_{t=1}^n \left| \frac{A_t - F_t}{A_t} \right| \times 100\% \quad (20)$$

Where:

- n represents the number of observations.
- A_t denotes the actual value at t .
- F_t refers to the forecasted value at t . MAPE offers a simple method to compare the accuracy of various forecasting models, where lower values signify higher accuracy.

Our proposal: AE-GAPB

The primary structure of the proposed AE-GAPB is illustrated in Fig. 2. In AE-GAPB, a Genetic Algorithm (GA) is utilized as the high-level strategy, while Particle Swarm Optimization (PSO) and Bat Algorithm (BA) function as the low-level methods. Our method employs the Mean Absolute Percentage Error of Prediction (MAPE) as the fitness measure. During optimization, the GA continuously selects hyperparameter combinations for PSO and BA based on the MAPE values returned, until the optimal combination is identified. Meanwhile, the GA's crossover and mutation rates are dynamically adjusted according to the exploration progress and the fitness values of individuals in the solution space. This approach increases the adaptability of both the low-level and high-level heuristics within the hyper-heuristic algorithm, thereby enhancing overall accuracy and robustness. The following sections provide a detailed explanation of the high-level component and low-level heuristics (LLHs).

High-level heuristics of AE-GAPB

Our proposed AE-GAPB utilizes GA as the high-level component. GA, a meta-heuristic algorithm inspired by natural evolutionary processes, continuously searches for the optimal hyperparameter combinations for PSO (inertia weights w , learning factors c_1 and c_2) and BA (loudness A and pulse emission rate r) within the LLHs. These hyperparameter combinations are represented as encoded individuals within the GA. Initially, a set of hyperparameter individuals is randomly generated and evaluated based on the MAPE values returned by the time series prediction model. The GA then performs crossover and mutation operations on these individuals, selecting and refining them to identify effective hyperparameter combinations for the LLHs.

At the same time, the GA, as the high-level component, adaptively adjusts its crossover and mutation rates based on the exploration time (iterations) and prediction accuracy (measured by MAPE) using Eq. (21). Specifically, as the number of iterations T increases, the crossover rate P_c and mutation rate P_m decrease, leading the population to gradually converge towards the optimal solution. Conversely, when an individual's fitness value is low, the crossover and mutation rates increase, promoting greater diversity and allowing the algorithm to explore a wider range of potential solutions. As fitness values improve, the rates are reduced, enabling the population to concentrate more on refining the optimal solution. Traditional methods rely on static hyperparameters, which can limit their adaptability to diverse datasets. AE-GAPB leverages a high-level GA to dynamically adjust the hyperparameters of low-level heuristics (PSO and BA) in real-time, ensuring improved adaptability and search performance. Effective forecasting requires a careful balance between exploring the solution space and refining

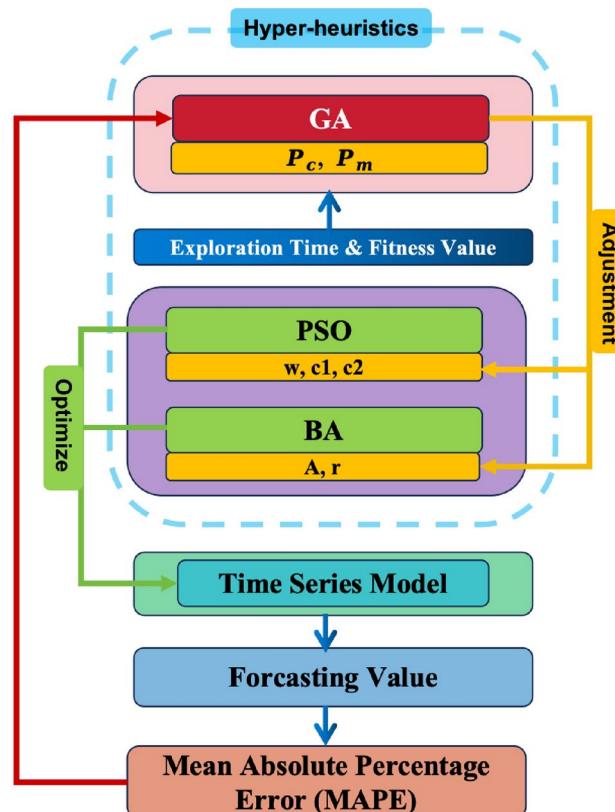


Fig. 2. The structure of the proposed AE-GAPB hyper-heuristic algorithm.

promising areas. AE-GAPB achieves this through an adaptive adjustment mechanism for crossover and mutation rates in GA, guided by fitness feedback, significantly reducing the likelihood of falling into local optima.

$$P = P_{min} + 2 * \left(1 - \frac{T}{T_{max}}\right) * \frac{f_i}{f_{total}} * P_{max} \quad (21)$$

Where:

- P is the current crossover (or mutation) rate.
- P_{min} represents the minimum crossover (or mutation) rate.
- T denotes the current number of iterations in the algorithm.
- T_{max} is the total number of iterations allowed in the algorithm.
- f_i indicates the fitness value for individual x_i .
- f_{total} refers to the sum of all individual fitness values.
- P_{max} is the maximum crossover (or mutation) rate.

Low-level heuristics of AE-GAPB

In AE-GAPB, search operators from PSO and BA are utilized as LLHs.

LLHs from PSO

PSO utilizes particle velocity v_i and position x_i , combined with empirically-driven updating mechanisms, to explore the parameters of the globally optimal time-series model. The fundamental search operators for velocity and position updates are detailed in Eqs. (22) and (23).

$$\begin{aligned} v_i(t+1) = & w \cdot v_i(t) + c_1 \cdot r_1 \cdot (p_{i,best} - x_i(t)) \\ & + c_2 \cdot r_2 \cdot (g_{best} - x_i(t)) \end{aligned} \quad (22)$$

In this context, $v_i(t)$ denotes the velocity of the i -th particle at iteration t ; w represents the inertia weight; c_1 and c_2 are the cognitive and social learning factors, respectively; r_1 and r_2 are random values within the interval $[0, 1]$; $p_{i,best}$ signifies the best position achieved by the i -th particle so far; g_{best} refers to the best overall position found by the swarm; and $x_i(t)$ is the position of the i -th particle at iteration t .

$$x_i(t+1) = x_i(t) + v_i(t+1) \quad (23)$$

Here, $x_i(t+1)$ represents the position of the i -th particle at iteration $t+1$, $x_i(t)$ is its position at iteration t , and $v_i(t+1)$ denotes its velocity at iteration $t+1$.

LLHs from BA

The key search operators from BA are outlined as follows.

$$\text{Frequency update: } f_i = f_{min} + (f_{max} - f_{min})\beta \quad (24)$$

In this case, β is a random value within the interval $[0, 1]$, and f_{min} and f_{max} correspond to the minimum and maximum frequency values, respectively.

$$\text{Velocity update: } v_i^{(t+1)} = v_i^{(t)} + (x_i^{(t)} - x_*)f_i \quad (25)$$

Here, $v_i^{(t)}$ denotes the velocity of bat i at time t , $x_i^{(t)}$ represents its position at time t , and x_* indicates the current global best solution.

$$\text{Position update: } x_i^{(t+1)} = x_i^{(t)} + v_i^{(t+1)} \quad (26)$$

$$\text{Local search: } x_{\text{new}} = x_{\text{best}} + \epsilon A^{(t)} \quad (27)$$

The bat randomly generates a number, and if this number surpasses its pulse rate r , it initiates a random flight. In this case, x_{new} represents the newly generated position, x_{best} denotes the current best position, ϵ is a random value within the range $[-1, 1]$, and A indicates the average loudness of all bats in the current generation.

The hierarchical framework of AE-GAPB allows it to perform effectively on different forecasting models (e.g., SES, DES, HWM, WMA, AR, LSTM) and datasets from regions such as Hokkaido, Kyushu, and Tohoku. Its robustness and scalability make it highly suitable for electricity forecasting applications. In summary, the Algorithm of AE-GAPB is presented in Algorithm 1.

Require: Population size: N , Dimension: D , Max. iteration: T

Ensure: Optimal hyper-parameters: $x_{best}^{(t)}$

- 1: $t = 0$
- 2: $X^{(t)} \leftarrow \text{initial}(N, D)$
- 3: $x_{best}^{(t)} \leftarrow \text{best}(X^{(t)})$
- 4: **while** $t < T$ **do**
- 5: Applying LLHs to optimize the model.
- 6: Calculating the fitness value f_{MAPE}^t .
- 7: Updating the crossover rate p_c and variance rate p_r .
- 8: Applying crossover and mutation operators.
- 9: Generating a new offspring population $X^{(t+1)}$.
- 10: Surviving elite individuals.
- 11: $t = t + 1$.
- 12: **end while**
- 13: **return** Optimal Value

Algorithm 1. AE-GAPB**Numerical experiments**

This section presents a summary of the numerical experiments, detailing the experimental setup and outcomes, to enable a fair evaluation and comparison of the performance of the proposed AE-GAPB.

Experimental settings

We conducted experiments on datasets of renewable energy generation and electricity consumption in three regions of Japan: Hokkaido, Kyushu, and Tohoku. The training set spans from 2000 to 2011(75% of the total data set), while the test set covers the period from 2012 to 2015(25% of the total data set). These datasets were sourced from the Federation of Electric Power Companies of Japan (FEPC).

To assess the performance of our proposed AE-GAPB, we utilize six distinct time series models: SES, DES, HWM, WMA, AR, and LSTM. We also compare our approach with PSO, BA, Differential Evolution (DE)³⁷, Grey Wolf Optimizer (GWO)³⁸, and Whale Optimization Algorithm (WOA)³⁹. To evaluate the impact of the auto-evolutionary ability of the high-level components in the proposed method, we also compared GAPB without the auto-evolutionary in our experiments. The parameter settings, including hyperparameters, are summarized in Table 2. And the hyperparameters α of SES, α, β of DES, α, β, γ of HWM, ω of WMA, ϕ of AR, η of LSTM which are automatically searched by optimization algorithm. Additionally, each algorithm underwent 30 trial runs to minimize the impact of randomness.

Experimental results

The performance of the model is evaluated using the MAPE between the predicted and actual electricity consumption values. The experimental results are shown in Tables 3 through 8. A lower MAPE value signifies greater forecasting accuracy (Tables 4,5,6,7,8). Furthermore, comparison curves are displayed in Figs. 3, 4, 5, 6,

Parameters	Values
Population size (N)	30
Number of iterations (T)	500
Maximum crossover rate (P_c^{max})	0.99
Minimum crossover rate (P_c^{min})	0.4
Maximum variation rate (P_r^{max})	0.01
Minimum variation rate (P_r^{min})	0.005
Inertia weight (w)	1
Cognitive coefficient (c_1)	1.5
Social coefficient (c_2)	1.5
Loudness (A)	0.5
Pulse emission rate (r)	0.5
Attraction factor (C)	1.5
Learning Factors (l_1, l_2)	0.5

Table 2. Parameters selection.

Models		AE-GAPB	GAPB	PSO	BA	DE	GWO	WOA
SES	α	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	MAPE	30.89%	30.89%	30.89%	30.89%	30.89%	30.89%	30.89%
DES	α, β	0.91, 1.26	0.85, 1.49	0.61, 2.68	0.52, 2.70	0.53, 2.89	0.51, 2.72	0.52, 2.91
	MAPE	35.95%	41.56%	72.98%	74.79%	53.03%	78.64%	52.26%
HWM	α, β, γ	0.47, 0.26, 0.64	1.01, 0.18, 0.92	0.08, 0.50, 0.34	0.18, 0.36, 0.92	0.75, 0.13, 0.93	0.35, 0.51, 0.49	0.97, 0.18, 0.02
	MAPE	44.01%	48.16%	53.45%	49.10%	51.01%	66.63%	49.68%
WMA	ω	1.24	1.24	1.24	1.24	1.24	1.24	1.24
	MAPE	75.21%	75.21%	75.21%	75.21%	75.21%	75.21%	75.21%
AR	ϕ	0.82	0.82	0.82	0.82	0.82	0.82	0.82
	MAPE	82.45%	82.45%	82.45%	82.45%	82.45%	82.45%	82.45%
LSTM	η	0.22	0.27	0.25	0.12	0.16	0.15	0.10
	MAPE	36.03%	41.17%	53.44%	42.06%	53.74%	44.86%	48.83%

Table 3. Results of Hokkaido Renewable Energy Generation Dataset.

Models		AE-GAPB	GAPB	PSO	BA	DE	GWO	WOA
SES	α	0.67	0.67	0.67	0.67	0.67	0.67	0.67
	MAPE	7.35%	7.35%	7.35%	7.35%	7.35%	7.35%	7.35%
DES	α, β	1.20, 0.35	1.17, 0.13	1.17, 0.13	1.17, 0.13	1.18, 0.13	1.18, 0.13	1.18, 0.13
	MAPE	6.17%	12.51%	12.54%	12.56%	12.53%	12.53%	12.62%
HWM	α, β, γ	0.07, 0.10, 0.05	0.03, 0.25, 0.44	0.30, 0.01, 0.71	0.02, 0.83, 0.62	1.00, 0.27, 0.35	0.15, 0.57, 0.17	0.01, 0.86, 0.16
	MAPE	5.47%	6.41%	8.90%	11.07%	9.23%	9.39%	14.77%
WMA	ω	1.09	1.09	1.09	1.09	1.09	1.09	1.09
	MAPE	6.19%	6.19%	6.19%	6.19%	6.19%	6.19%	6.19%
AR	ϕ	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	MAPE	11.11%	11.11%	11.11%	11.11%	11.11%	11.11%	11.11%
LSTM	η	0.26	0.38	0.33	0.35	0.56	0.87	0.43
	MAPE	10.01%	15.18%	20.62%	30.77%	47.77%	23.48%	25.69%

Table 4. Results of Hokkaido Electricity Consumption Dataset.

Models		AE-GAPB	GAPB	PSO	BA	DE	GWO	WOA
SES	α	0.76	0.76	0.76	0.76	0.76	0.76	0.76
	MAPE	9.08%	9.08%	9.08%	9.08%	9.08%	9.08%	9.08%
DES	α, β	0.39, 0.06	0.38, 0.06	0.38, 0.06	0.38, 0.06	0.38, 0.06	0.38, 0.06	0.38, 0.06
	MAPE	9.76%	9.76%	9.76%	9.76%	9.76%	9.76%	9.76%
HWM	α, β, γ	0.01, 0.36, 0.27	0.43, 0.14, 0.48	0.46, 0.10, 0.62	0.96, 0.03, 0.76	1.00, 0.03, 0.02	0.45, 0.12, 0.59	0.44, 0.14, 0.59
	MAPE	7.32%	9.44%	9.53%	9.77%	9.98%	9.51%	9.51%
WMA	ω	0.96	0.96	0.96	0.96	0.96	0.96	0.96
	MAPE	8.82%	8.82%	8.82%	8.82%	8.82%	8.82%	8.82%
AR	ϕ	0.99	0.99	0.99	0.99	0.99	0.99	0.99
	MAPE	11.17%	11.17%	11.17%	11.17%	11.17%	11.17%	11.17%
LSTM	η	0.88	0.44	0.46	0.48	0.75	0.56	0.37
	MAPE	9.42%	11.64%	34.46%	14.66%	49.55%	41.57%	12.55%

Table 5. Results of Kyushu Renewable Energy Generation Dataset.

7, 8. Additionally, we generated a graph showing the difference between electricity consumption and renewable energy generation in Figs. 9, 10, 11, which can assist in scheduling fossil energy generation while optimizing the use of renewable energy.

Discussion

On the Hokkaido Renewable Energy Generation Dataset, the proposed method consistently demonstrates superior performance across the DES, HWM, WMA, and LSTM models compared to the algorithm GAPB without the auto-evolutionary and other algorithms. It also achieves comparable prediction accuracy on the

Models		AE-GAPB	GAPB	PSO	BA	DE	GWO	WOA
SES	α	0.84	0.84	0.84	0.84	0.84	0.84	0.84
	MAPE	5.32%	5.32%	5.32%	5.32%	5.32%	5.32%	5.32%
DES	α, β	1.11, 0.99	1.18, 0.10	1.18, 0.10	1.20, 0.10	1.18, 0.10	1.18, 0.10	1.32, 0.08
	MAPE	5.20%	5.92%	5.94%	6.00%	5.94%	5.94%	6.35%
HWM	α, β, γ	0.07, 1.18, 0.14	0.23, 0.04, 0.28	0.26, 0.04, 0.30	0.47, 1.29, 0.23	0.25, 0.51, 0.27	0.23, 0.04, 0.28	0.16, 0.14, 0.06
	MAPE	2.98%	6.86%	6.89%	7.20%	7.06%	6.86%	8.08%
WMA	ω	1.10	1.10	1.10	1.10	1.10	1.10	1.10
	MAPE	5.22%	5.22%	5.22%	5.22%	5.22%	5.22%	5.22%
AR	ϕ	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	MAPE	8.20%	8.20%	8.20%	8.20%	8.20%	8.20%	8.20%
LSTM	η	0.39	0.40	0.56	0.55	0.40	0.71	0.39
	MAPE	9.20%	31.89%	48.15%	32.88%	32.94%	32.52%	36.65%

Table 6. Results of Kyushu Electricity Consumption Dataset.

Models		AE-GAPB	GAPB	PSO	BA	DE	GWO	WOA
SES	α	0.67	0.67	0.67	0.67	0.67	0.67	0.67
	MAPE	12.39%	12.39%	12.39%	12.39%	12.39%	12.39%	12.39%
DES	α, β	0.44, 0.04	0.61, 0.02	0.61, 0.02	0.61, 0.02	0.61, 0.02	0.61, 0.02	0.61, 0.02
	MAPE	14.06%	14.21%	14.21%	14.21%	14.21%	14.21%	14.21%
HWM	α, β, γ	0.03, 0.19, 0.12	0.51, 0.12, 0.45	0.08, 1.02, 0.13	0.81, 0.28, 0.19	0.40, 0.28, 0.45	0.01, 0.94, 0.11	0.78, 0.30, 0.02
	MAPE	12.56%	15.37%	27.18%	17.61%	16.43%	22.09%	17.74%
WMA	ω	0.75	0.75	0.75	0.75	0.75	0.75	0.75
	MAPE	11.01%	11.01%	11.01%	11.01%	11.01%	11.01%	11.01%
AR	ϕ	0.99	0.99	0.99	0.99	0.99	0.99	0.99
	MAPE	11.62%	11.62%	11.62%	11.62%	11.62%	11.62%	11.62%
LSTM	η	0.28	0.77	0.41	0.36	0.41	0.13	0.50
	MAPE	15.02%	26.69%	38.53%	41.43%	72.71%	78.83%	53.15%

Table 7. Results of Tohoku Renewable Energy Generation Dataset.

Models		AE-GAPB	GAPB	PSO	BA	DE	GWO	WOA
SES	α	0.40	0.40	0.40	0.40	0.40	0.40	0.40
	MAPE	5.75%	5.75%	5.75%	5.75%	5.75%	5.75%	5.75%
DES	α, β	1.18, 0.11	1.20, 0.11	1.20, 0.11	0.95, 0.71	1.23, 0.10	1.20, 0.12	0.95, 0.82
	MAPE	2.01%	2.06%	2.09%	2.54%	2.20%	2.07%	2.55%
HWM	α, β, γ	0.11, 0.12, 0.21	0.11, 0.13, 0.21	0.14, 0.71, 0.15	0.05, 0.23, 0.14	0.01, 0.63, 0.37	0.19, 0.11, 0.25	0.11, 0.20, 0.21
	MAPE	5.11%	5.14%	8.67%	6.33%	18.55%	5.63%	6.93%
WMA	ω	1.12	1.12	1.12	1.12	1.12	1.12	1.12
	MAPE	2.67%	2.67%	2.67%	2.67%	2.67%	2.67%	2.67%
AR	ϕ	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	MAPE	8.55%	8.55%	8.55%	8.55%	8.55%	8.55%	8.55%
LSTM	η	0.90	0.42	0.54	0.37	0.44	0.23	0.63
	MAPE	9.19%	9.28%	16.77%	19.74%	54.82%	21.34%	55.83%

Table 8. Results of Tohoku Electricity Consumption Dataset.

SES and AR models, underscoring its versatility. The method delivers strong prediction results on the DES, HWM, and LSTM models for the Hokkaido Electricity Consumption Dataset, outperforming other algorithms. Additionally, it matches the prediction accuracy of different algorithms on the SES, WMA, and LSTM models, showcasing its reliability.

On the Kyushu renewable energy generation dataset, the proposed method consistently outperforms GAPB and other algorithms on HWM and LSTM models. It also achieves comparable prediction accuracy on SES, DES, WMA, and AR models, demonstrating its versatility. Similarly, for the Kyushu electricity consumption dataset, the method delivers strong prediction results on DES, HWM, and LSTM models, surpassing GAPB

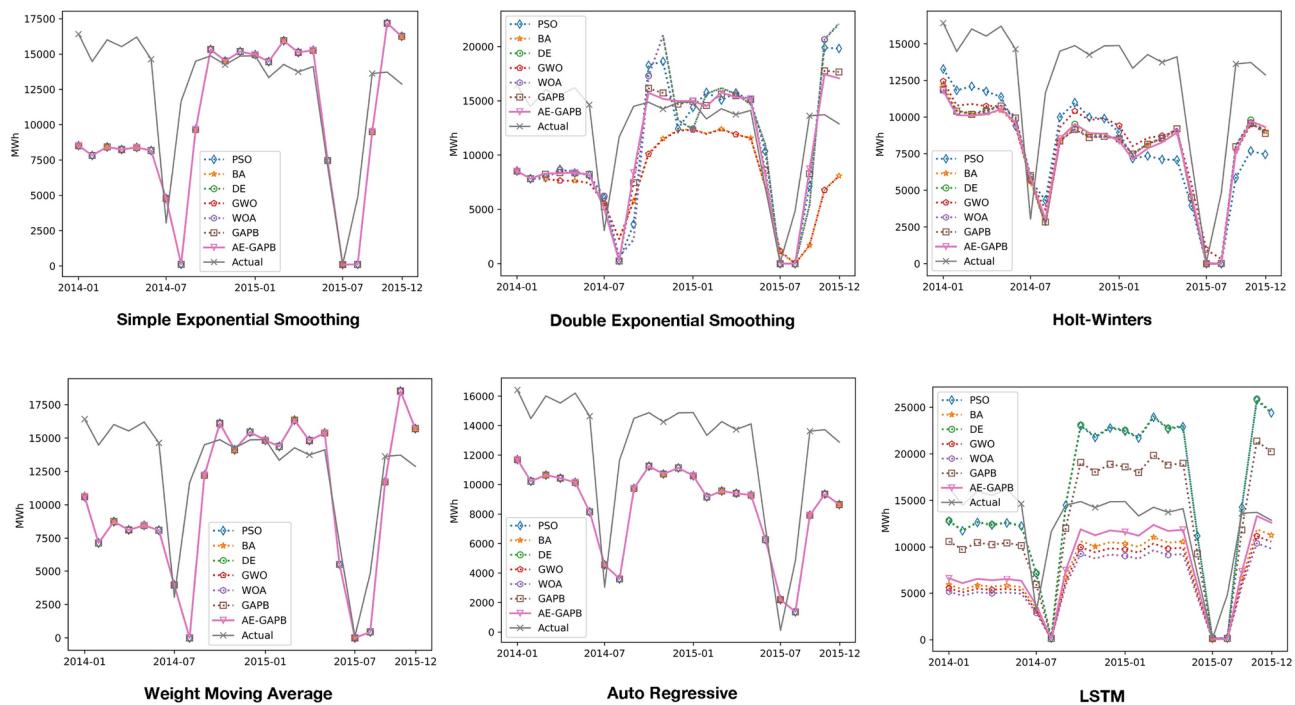


Fig. 3. Compare with actual data in Hokkaido Renewable Energy Generation Dataset.

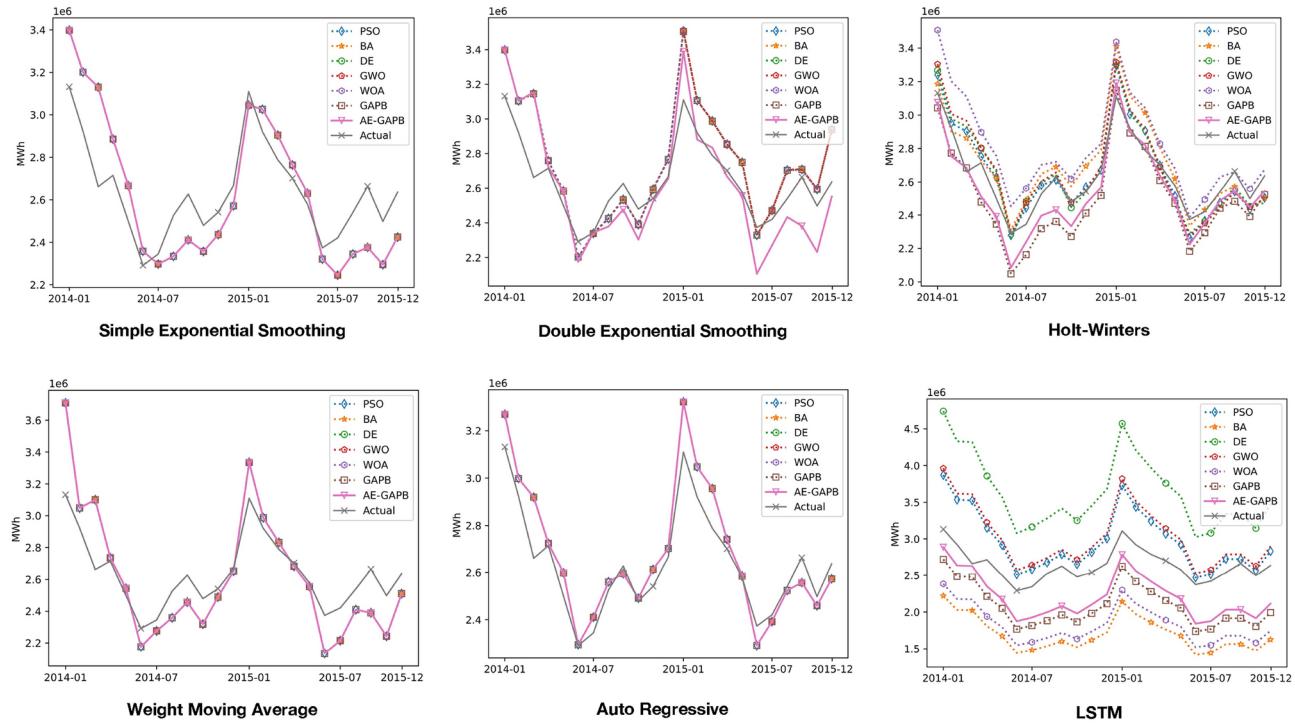


Fig. 4. Compare with actual data in Hokkaido Electricity Consumption Dataset.

and other methods. Moreover, its performance on SES, WMA, and AR models is on par with GAPB and other algorithms, further showcasing its reliability. In the Tohoku region, the proposed approach surpasses GAPB and other algorithms on the DES, HWM, and LSTM models in both renewable energy generation and electricity consumption datasets, while demonstrating comparable effectiveness on the SES, WMA, and AR models.

The experimental results clearly show that GAPB without auto-evolution outperforms other optimization algorithms across multiple time series forecasting models on various datasets. This indicates that GAPB not only

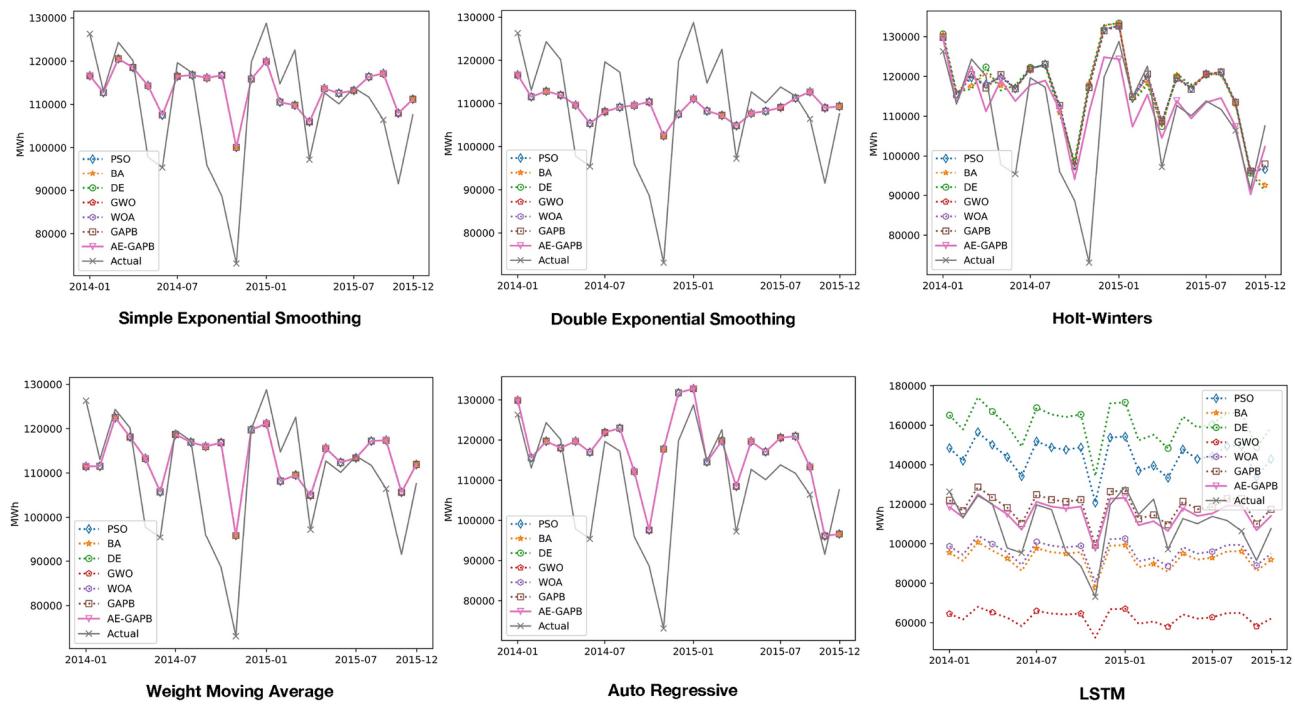


Fig. 5. Compare with actual data in Kyushu Renewable Energy Generation Dataset.

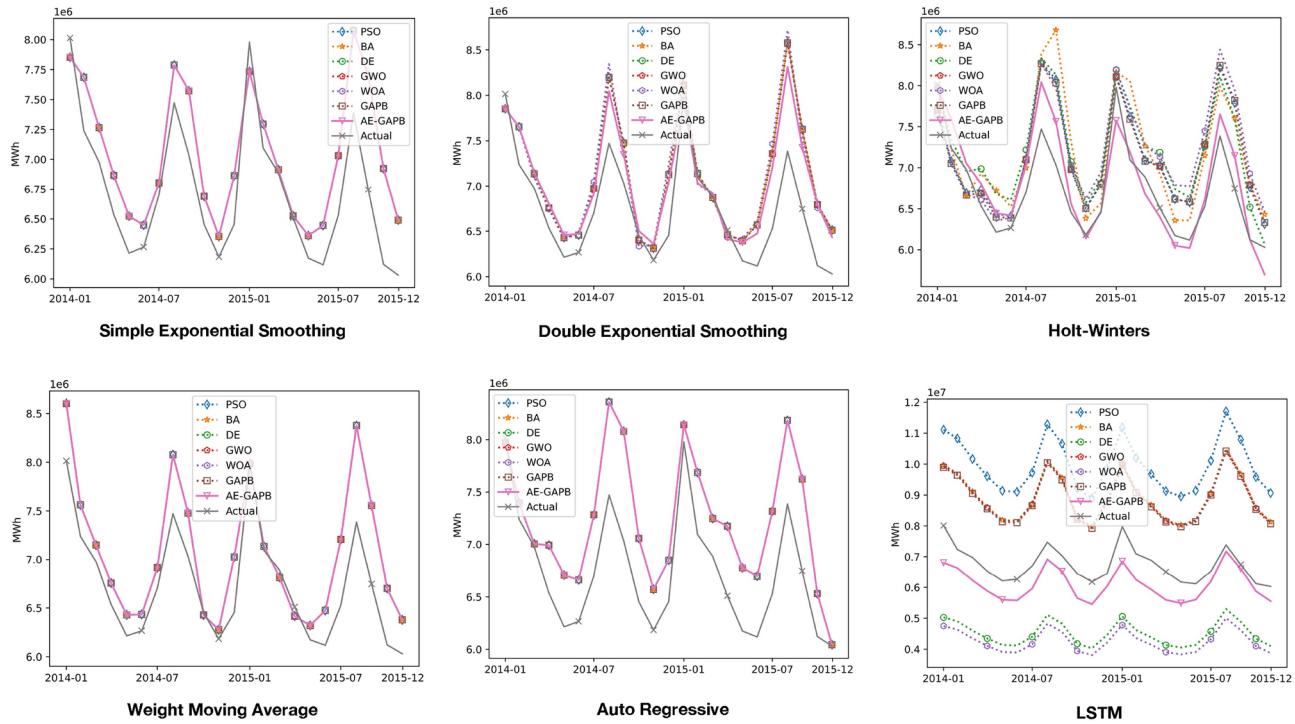


Fig. 6. Compare with actual data in Kyushu Electricity Consumption Dataset.

achieves superior optimization accuracy but also exhibits strong robustness across diverse real-world datasets. Additionally, AE-GAPB delivers even better results, suggesting that the auto-evolution mechanism in the high-level GA component effectively enhances the overall performance of the algorithm. AE-GAPB dynamically adjusts the crossover and mutation rates based on the algorithm's iterations and the returned MAPE. As iterations increase, the crossover and mutation rates gradually decrease, reducing the frequency of these operations and allowing the solution to converge toward the optimal result. At the same time, when the returned MAPE is

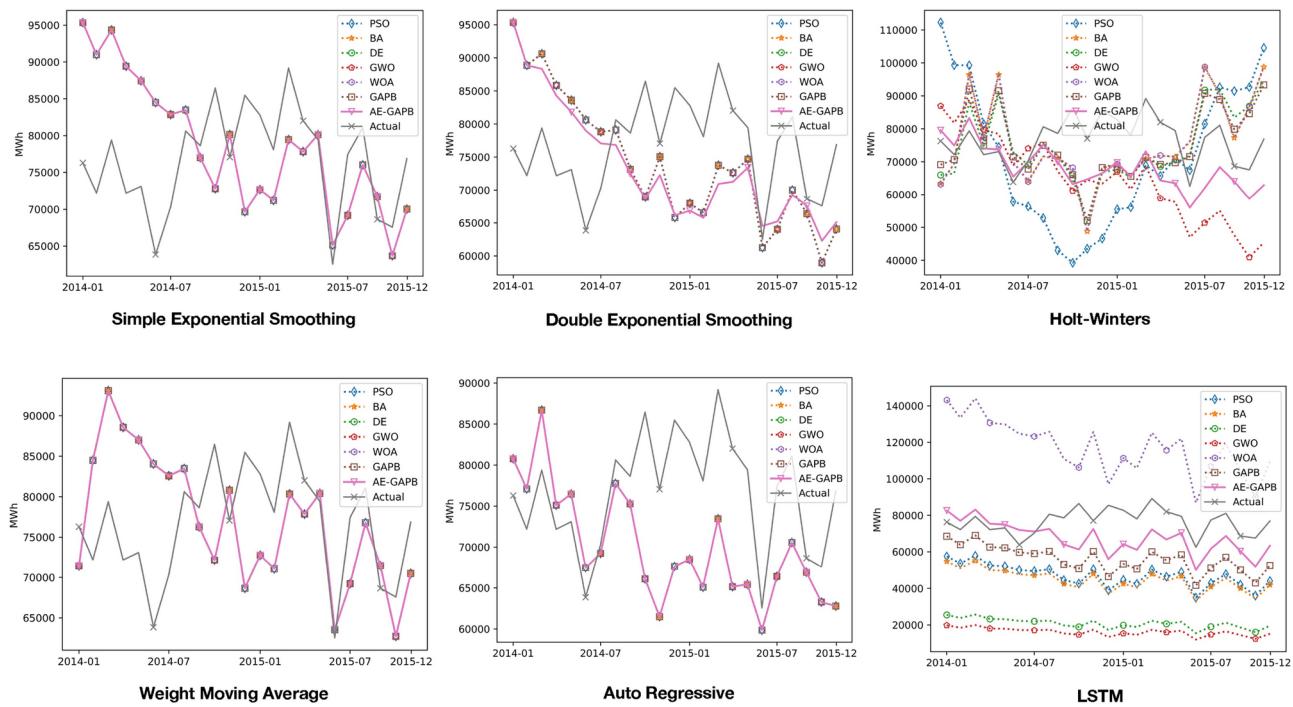


Fig. 7. Compare with actual data in Tohoku Renewable Energy Generation Dataset.

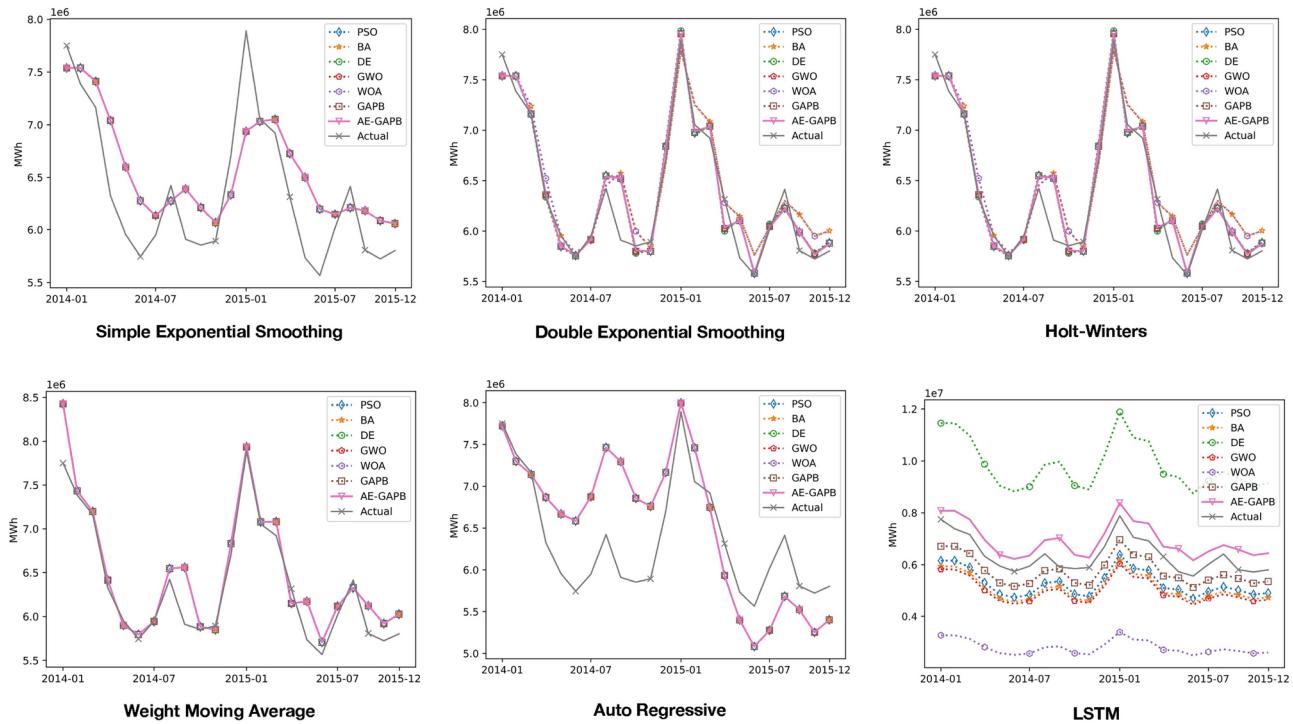


Fig. 8. Compare with actual data in Tohoku Electricity Consumption Dataset.

poor, the crossover and mutation rates are large, and the chance of crossover and mutation operations occurring increases, thus allowing the solution set to improve. On the contrary, when the returned MAPE is better, it means that the solution is in a better condition at this time, and the crossover rate and mutation rate decrease, thus enabling the solution to concentrate on the better solution.

Additionally, as depicted in Fig. 9, 10, 11, the difference between the forecasted electricity consumption and renewable energy generation using the proposed approach closely aligns with the actual difference. This

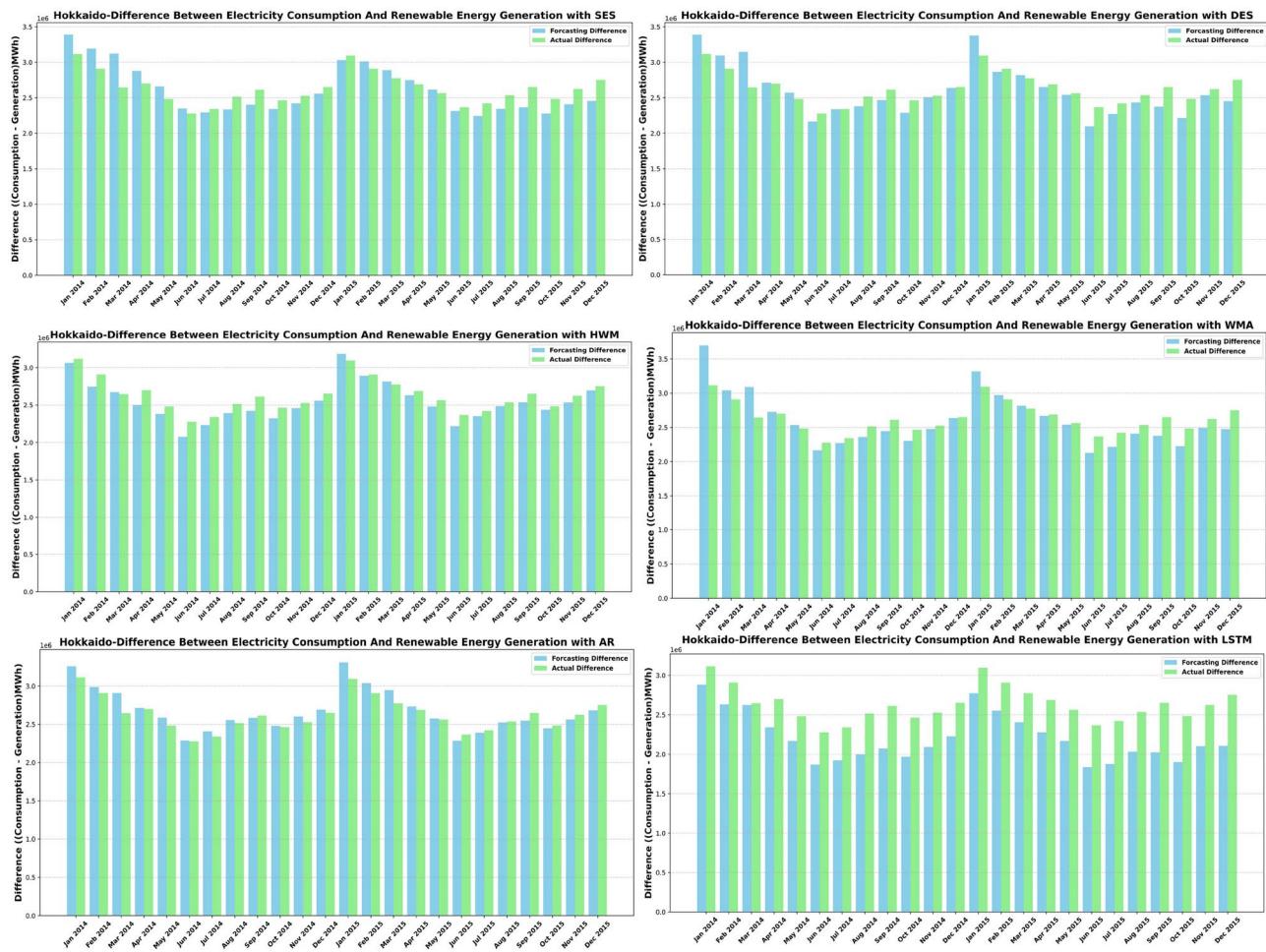


Fig. 9. Hokkaido-Difference Between Electricity Consumption And Renewable Energy Generation.

alignment facilitates efficient planning of thermal power generation capacity, maximizing the use of renewable energy and significantly reducing greenhouse gas emissions.

Given the significant influence of various environmental factors on renewable energy generation, the data in the Renewable Energy Generation Dataset tends to be less regular compared to that in the Electricity Consumption Dataset. As a result, the prediction model performs better on the Electricity Consumption Dataset than on the Renewable Energy Generation Dataset, highlighting the challenges of forecasting in more variable conditions.

Conclusion

Traditional electricity forecasting models are limited by fixed hyperparameters and lack robustness. In this paper, we propose a novel hyper-heuristic algorithm for optimizing electricity forecasting models, named AE-GAPB. The algorithm utilizes GA as the high-level component while incorporating search operators from PSO and BA as LLHs. GA continuously adapts to the exploration dynamics, achieving simultaneous adaptability in both the high-level and LLHs components. To evaluate the performance of the proposed method, we conducted numerical experiments on real datasets of renewable power generation and power consumption from the Hokkaido, Kyushu, and Tohoku regions of Japan, using six different time series forecasting models.

The results demonstrate that the GAPB without auto-evolution offers superior accuracy and robustness compared to PSO, BA, DE, GWO, and WOA. At the same time, The experimental results demonstrate that incorporating the auto-evolution mechanism into the high-level GA component significantly enhances the performance of AE-GAPB. This clearly shows that the proposed method effectively optimizes the crossover and mutation operations of high-level GA components, thereby improving the algorithm's exploration accuracy. By dynamically adapting parameters, balancing exploration and exploitation, and delivering robust performance across diverse datasets, AE-GAPB emerges as a highly effective solution for electricity forecasting. Additionally, the difference between the forecasted power consumption and renewable energy generation by the proposed method closely aligns with the actual difference, making it applicable to real-world scenarios. This enables rational planning of thermal power generation and maximization of renewable energy utilization, thereby reducing greenhouse gas emissions.

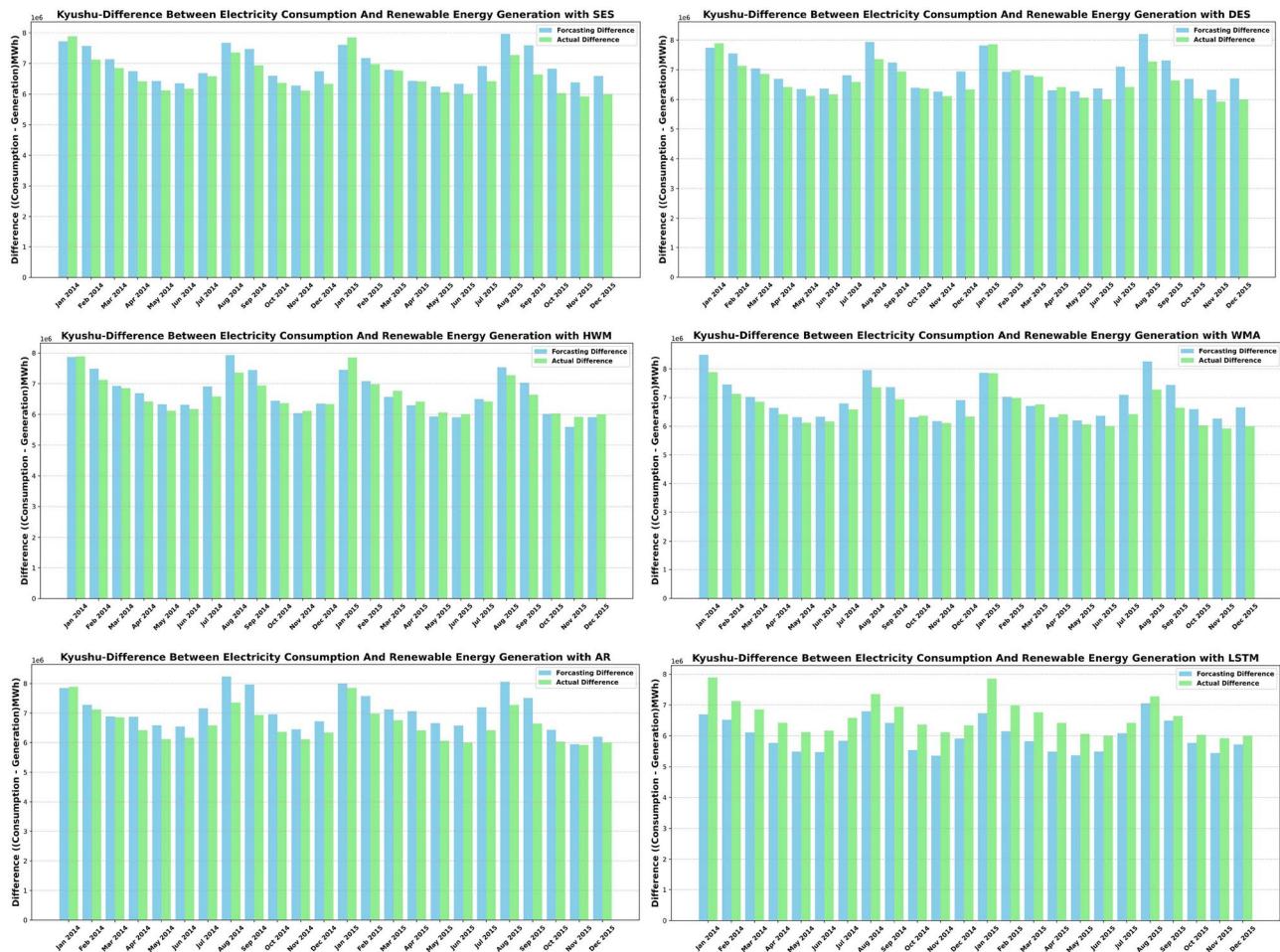


Fig. 10. Kyushu-Difference Between Electricity Consumption And Renewable Energy Generation.

Despite the improvements in accuracy and robustness, the complexity of the algorithm has increased, leading to higher computational costs. In future research, we aim to simplify and enhance the efficiency of the algorithm, reducing its computational burden. Additionally, we plan to integrate other effective low-level heuristics to further improve its performance across diverse domains. Exploring hybrid frameworks that combine AE-GAPB with advanced deep learning models, such as Long Short-Term Memory (LSTM) networks or Transformer architectures, will be considered to capture complex temporal patterns in data more effectively. Extending the algorithm to real-time optimization in dynamic environments, such as smart grids and IoT-enabled systems, is another avenue for enhancement. Furthermore, addressing multi-objective optimization problems will allow the algorithm to manage trade-offs between conflicting objectives, such as minimizing costs while maximizing renewable energy usage. Finally, we aim to validate the algorithm on broader datasets from various regions and apply it to real-world optimization problems in fields like healthcare, transportation, and finance, further demonstrating its versatility and scalability.

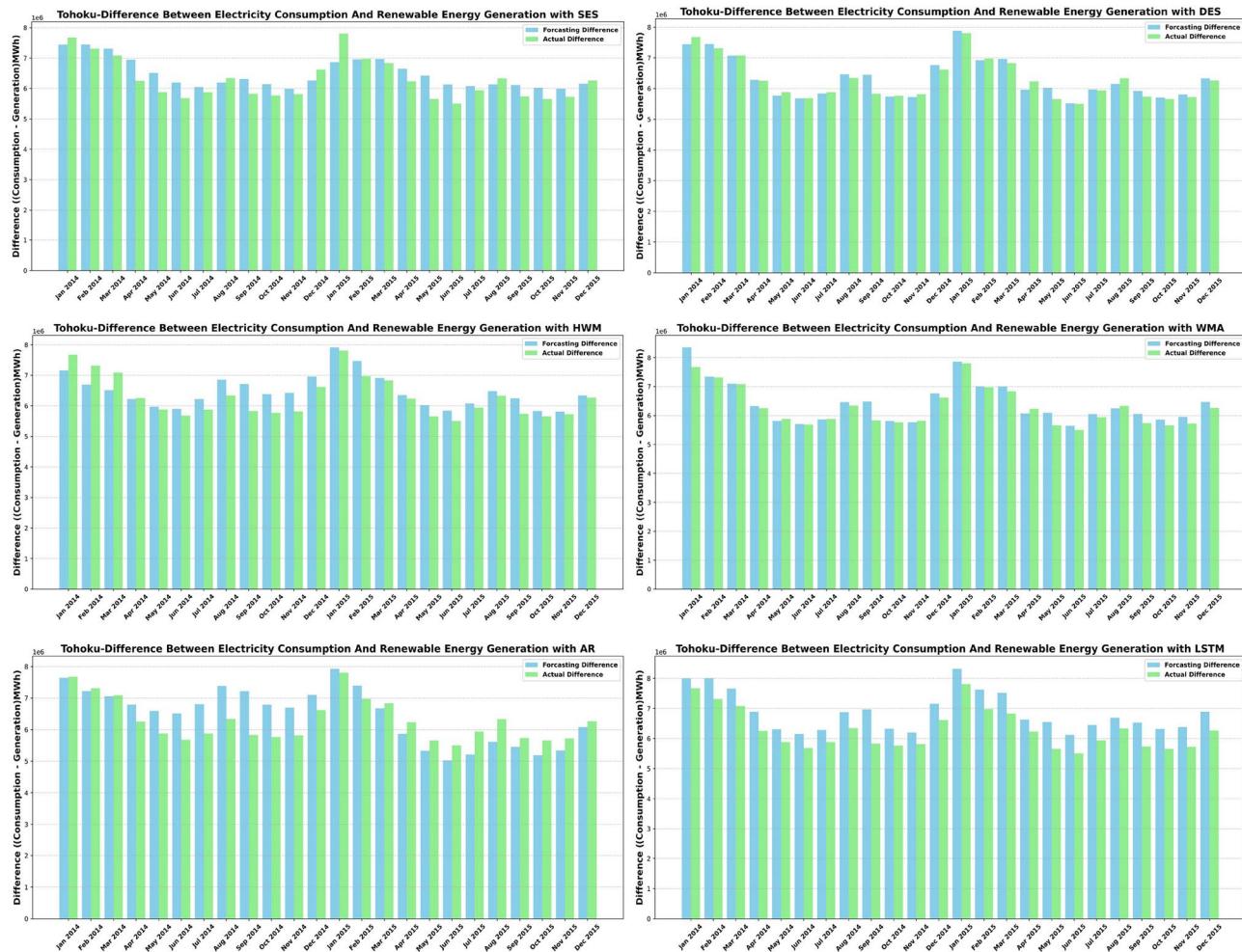


Fig. 11. Tohoku-Difference Between Electricity Consumption And Renewable Energy Generation.

Data availability

The source code of this research can be downloaded from <https://github.com/soyo123/Auto-Evolution-Hyper-Huristic-Algorithm>.

Received: 20 October 2024; Accepted: 15 January 2025

Published online: 20 January 2025

References

1. Wang, R., Ma, J., Sheng, H., Zavala, V. & Jin, S. Exploiting different electricity markets via highly rate-mismatched modular electrochemical synthesis. *Nature Energy*, 1–10 (2024)
2. László, S., Girona, M., Jäger-Waldau, A., Kougias, I., Fahl, F. & Szabó, S. Impacts of large-scale deployment of vertical bifacial photovoltaics on european electricity market dynamics. *Nature Communications*, 15 (2024)
3. Seel, J., Millstein, D., Mills, A., Bolinger, M. & Wiser, R. Plentiful electricity turns wholesale prices negative. *Advances in Applied Energy* 4, 100073 (2021).
4. Eren, Y. İbrahim Küçükdemir: A comprehensive review on deep learning approaches for short-term load forecasting. *Renewable and Sustainable Energy Reviews* 189, 114031 (2024).
5. Bozlak, Çağatay Berke & Yaşar, C. F. An optimized deep learning approach for forecasting day-ahead electricity prices. *Electric Power Systems Research* 229, 110129 (2024).
6. Deng, C., Zhang, X., Huang, Y. & Bao, Y. Equipping seasonal exponential smoothing models with particle swarm optimization algorithm for electricity consumption forecasting. *Energies* 14, 4036 (2021).
7. Mbuli, N. & Pretorius, J.-H.C. Simple exponential smoothing for forecasting the numbers of pole-mounted transformer failures. In: 2023 IEEE AFRICON, pp. 1–5 (2023)
8. Ke, G. et al. Epidemiological analysis of hemorrhagic fever with renal syndrome in china with the seasonal-trend decomposition method and the exponential smoothing model. *Scientific Reports* 6, 39350 (2016).
9. Chicas, S., Valladarez-Cob, J., Omine, K., Sivsankar, V. & Kim, S. Spatiotemporal distribution, trend, forecast, and influencing factors of transboundary and local air pollutants in nagasaki prefecture, japan. *Scientific Reports* 13, 851 (2023).
10. Ali Raza, M., Amin, A., Aslam, M., Nawaz, T., Irfan, M. & Tariq, F. Nonparametric mixed exponentially weighted moving average-moving average control chart. *Scientific Reports*, 14 (2024)
11. Yin, H., Chen, J. & Jiang, Z. Study on spatial auto-regression within soil physical-chemical indicators in typical karst demonstration zone. In: 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS, pp. 6528–6531 (2021)

12. Guo, Q., He, Z. & Wang, Z. Monthly climate prediction using deep convolutional neural network and long short-term memory. *Scientific Reports*, **14** (2024)
13. Cheang, W.-K. & Reinsel, G. Bias reduction of autoregressive estimates in time series regression model through restricted maximum likelihood. *Journal of The American Statistical Association - J AMER STATIST ASSN* **95**, 1173–1184. <https://doi.org/10.1080/01621459.2000.10474318> (2000).
14. Fitria, V. & Widayanti, L. Enhancing accuracy in stock price prediction: The power of optimization algorithms. *MATRIK: Jurnal Manajemen, Teknik Informatika dan Rekayasa Komputer* **23**, 405–418. <https://doi.org/10.30812/matrik.v23i2.3785> (2024).
15. Zhong, R. & Yu, J. $\text{Dea}^2\text{ h}^2$: differential evolution architecture based adaptive hyper-heuristic algorithm for continuous optimization. *Cluster Computing*, 1–28 (2024)
16. Thejus, S. & P, S. Deep learning-based power consumption and generation forecasting for demand side management. In: 2021 Second International Conference on Electronics and Sustainable Communication Systems (ICESC), pp. 1350–1357 (2021)
17. Elsaraiti, M., Ali, G., Musbah, H., Merabet, A. & Little, T. Time series analysis of electricity consumption forecasting using arima model. In: 2021 IEEE Green Technologies Conference (GreenTech), pp. 259–262 (2021)
18. Qureshi, M., Ahmad, A. & Rehman, S.U. Deep learning-based forecasting of electricity consumption. *Scientific Reports*, **14** (2024)
19. Lambert, G., Hamrouche, B. & Vilmarest, J. Frugal day-ahead forecasting of multiple local electricity loads by aggregating adaptive models. *Scientific Reports* **13** (2023)
20. Bhutta, M., Li, Y., Abubakar, M., Almasoudi, F., Alatawi, K., Altimania, M. & Al-Barashi, M. Optimizing Solar Power Efficiency in Smart Grids Using Hybrid Machine Learning Models for Accurate Energy Generation Prediction
21. Mauricio, C.C. & Ostia, C.F. Cuckoo search algorithm optimization of holt-winter method for distribution transformer load forecasting. In: 2023 9th International Conference on Control, Automation and Robotics (ICCAR), pp. 36–42 (2023)
22. Zhong, R., Yu, J., Zhang, C. & Munetomo, M.: Surrogate ensemble-assisted hyper-heuristic algorithm for expensive optimization problems. *International Journal of Computational Intelligence Systems*, **16** (2023)
23. Zhong, R., Zhang, E. & Munetomo, M. Evolutionary multi-mode slime mold optimization: a hyper-heuristic algorithm inspired by slime mold foraging behaviors. *The Journal of Supercomputing* **80**, 1–32 (2024).
24. Li, J., Lin, Y. & Misir, M.: Neural network based heuristic selection for selection hyper-heuristics. In: 2023 IEEE Congress on Evolutionary Computation (CEC), pp. 1–7 (2023)
25. Zhong, R. & Yu, J. A novel evolutionary status guided hyper-heuristic algorithm for continuous optimization. *Cluster Computing* (2024)
26. Vela, A., Cruz-Duarte, J. M., Ortiz-Bayliss, J. C. & Amaya, I. Beyond hyper-heuristics: A squared hyper-heuristic model for solving job shop scheduling problems. *IEEE Access* **10**, 43981–44007 (2022).
27. Singh, E. & Pillay, N. A parameter-based analysis of ant-based generation hyper-heuristics. In: 2022 IEEE Symposium Series on Computational Intelligence (SSCI), pp. 812–819 (2022)
28. Arpacı, A., Chen, J., Drake, J.H. & Glover, T.: Sequence-based selection hyper-heuristics for real-world fibre network design optimisation. In: 2022 IEEE Congress on Evolutionary Computation (CEC), pp. 1–8 (2022)
29. Anwar, A.A. & Zhang, X.: Reinforcement learning based hyper-heuristics for many-objective pickup and delivery problem. In: 2023 IEEE International Conference on Data Mining (ICDM), pp. 924–929 (2023)
30. Yin, S. & Xiang, Z. A hyper-heuristic algorithm via proximal policy optimization for multi-objective truss problems. *Expert Systems with Applications* **256**, 124929. <https://doi.org/10.1016/j.eswa.2024.124929> (2024).
31. Gardner, E. S. Exponential smoothing: The state of the art-part ii. *International Journal of Forecasting* **22**(4), 637–666 (2006).
32. LaViola, J.J. Double exponential smoothing: an alternative to kalman filter-based predictive tracking. In: Proceedings of the Workshop on Virtual Environments 2003. EGVE '03, pp. 199–206. Association for Computing Machinery, New York, NY, USA (2003)
33. Hannan, E. J. & Quinn, B. G. The determination of the order of an autoregression. *Journal of the Royal Statistical Society: Series B (Methodological)* **41**(2), 190–195 (1979).
34. Graves, A. Long Short-Term Memory, pp. 37–45. Springer, Berlin, Heidelberg (2012)
35. de Myttenaere, A., Golden, B., Le Grand, B. & Rossi, F. Mean absolute percentage error for regression models. *Neurocomputing*, **192**, 38–48 (2016). Advances in artificial neural networks, machine learning and computational intelligence
36. Bian, J., Hou, T., Ren, D., Lin, C., Qiao, X., Ma, X., Ma, J., Wang, Y., Wang, J. & Liang, X. Predicting mine water inflow volumes using a decomposition-optimization algorithm-machine learning approach. *Scientific Reports* **14** (2024)
37. Yue, L., Hu, P. & Zhu, J. Advanced differential evolution for gender-aware english speech emotion recognition. *Scientific Reports* **14** (2024)
38. Zhang, H., Yang, T.-T. & Wang, W.-T.: A novel hybrid model for species distribution prediction using neural networks and grey wolf optimizer algorithm. *Scientific Reports*, **14** (2024)
39. Xu, H., Liu, W.-d., Li, L., Yao, D.-j. & Ma, L. Fsrw: fuzzy logic-based whale optimization algorithm for trust-aware routing in iot-based healthcare. *Scientific Reports*, **14** (2024)

Acknowledgements

This work was supported by JSPS KAKENHI Grant Numbers 24K15098 and 21A402.

Author contributions

Yang Cao: Conceptualization, Methodology, Investigation, Writing – original draft, Writing – review & editing. **Jun Yu:** Methodology, Formal Analysis, and Writing – review & editing. **Rui Zhong:** Investigation, Formal Analysis, and Writing – review & editing. **Masaharu Munetomo:** Writing – review & editing, Project administration, and Funding acquisition.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to R.Z.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025

Terms and Conditions

Springer Nature journal content, brought to you courtesy of Springer Nature Customer Service Center GmbH (“Springer Nature”).

Springer Nature supports a reasonable amount of sharing of research papers by authors, subscribers and authorised users (“Users”), for small-scale personal, non-commercial use provided that all copyright, trade and service marks and other proprietary notices are maintained. By accessing, sharing, receiving or otherwise using the Springer Nature journal content you agree to these terms of use (“Terms”). For these purposes, Springer Nature considers academic use (by researchers and students) to be non-commercial.

These Terms are supplementary and will apply in addition to any applicable website terms and conditions, a relevant site licence or a personal subscription. These Terms will prevail over any conflict or ambiguity with regards to the relevant terms, a site licence or a personal subscription (to the extent of the conflict or ambiguity only). For Creative Commons-licensed articles, the terms of the Creative Commons license used will apply.

We collect and use personal data to provide access to the Springer Nature journal content. We may also use these personal data internally within ResearchGate and Springer Nature and as agreed share it, in an anonymised way, for purposes of tracking, analysis and reporting. We will not otherwise disclose your personal data outside the ResearchGate or the Springer Nature group of companies unless we have your permission as detailed in the Privacy Policy.

While Users may use the Springer Nature journal content for small scale, personal non-commercial use, it is important to note that Users may not:

1. use such content for the purpose of providing other users with access on a regular or large scale basis or as a means to circumvent access control;
2. use such content where to do so would be considered a criminal or statutory offence in any jurisdiction, or gives rise to civil liability, or is otherwise unlawful;
3. falsely or misleadingly imply or suggest endorsement, approval , sponsorship, or association unless explicitly agreed to by Springer Nature in writing;
4. use bots or other automated methods to access the content or redirect messages
5. override any security feature or exclusionary protocol; or
6. share the content in order to create substitute for Springer Nature products or services or a systematic database of Springer Nature journal content.

In line with the restriction against commercial use, Springer Nature does not permit the creation of a product or service that creates revenue, royalties, rent or income from our content or its inclusion as part of a paid for service or for other commercial gain. Springer Nature journal content cannot be used for inter-library loans and librarians may not upload Springer Nature journal content on a large scale into their, or any other, institutional repository.

These terms of use are reviewed regularly and may be amended at any time. Springer Nature is not obligated to publish any information or content on this website and may remove it or features or functionality at our sole discretion, at any time with or without notice. Springer Nature may revoke this licence to you at any time and remove access to any copies of the Springer Nature journal content which have been saved.

To the fullest extent permitted by law, Springer Nature makes no warranties, representations or guarantees to Users, either express or implied with respect to the Springer nature journal content and all parties disclaim and waive any implied warranties or warranties imposed by law, including merchantability or fitness for any particular purpose.

Please note that these rights do not automatically extend to content, data or other material published by Springer Nature that may be licensed from third parties.

If you would like to use or distribute our Springer Nature journal content to a wider audience or on a regular basis or in any other manner not expressly permitted by these Terms, please contact Springer Nature at

onlineservice@springernature.com