International Core Journal of Engineering
Volume 11 Issue 5, 2025
ISSN: 2414-1895
DOI: 10.6919/ICJE.202505_11(5).0050

Optimization Study of Fluctuation Prediction and Standby Unit Dispatch for Wind and Solar Power Generation

Shaosong Zhang[a], Jiaxin Tian[b]

College of science, North China University of Science and Technology, Tangshan, China

[a]zss19831957585@126.com, [b]2187635224@qq.com

Abstract

With the growing use of renewable energy sources worldwide, the volatility problem of wind and solar power generation challenges the stability of power systems. In this paper, the problem of predicting the volatility of wind and solar power generation and optimizing the dispatch of standby units based on historical data is investigated. First, support vector regression (SVR) and random forest regression (RF) models are used for short-term volatility prediction of wind and solar power generation. The prediction accuracy of the models is improved by optimizing the thresholds, in which the RMSE of the SVR model is 7.49 and that of the RF model is 6.72, which are superior to the traditional methods. Next, a two-objective optimization model is proposed to optimize the standby unit scheduling with particle swarm optimization (PSO) algorithm to balance the stability of the power plant and the cost of the standby unit. The experimental results show that the optimized scheduling strategy can effectively reduce the start-stop frequency of the standby unit and improve the economy and reliability of the power system. The research in this paper provides an effective solution for the fluctuation prediction and scheduling optimization of wind and solar power generation, which has important application value.

Keywords

Wind Power; Solar Power; Volatility Prediction; Standby Unit Scheduling; Machine Learning; Optimization Algorithms.

1. Introduction

With the transformation of the global energy mix, renewable energy sources such as wind and solar power have been widely used. However, the volatility of these energy sources poses challenges to the stability and economics of power systems. Wind power is affected by wind speed and meteorological variations, while solar power is affected by sunlight intensity and weather conditions. This instability not only complicates power scheduling and load balancing, but also increases the need for standby units, which in turn raises the operating costs of the power system. Therefore, how to accurately predict the fluctuation of wind and solar power generation and optimize the scheduling of standby units has become a key issue to increase the share of renewable energy in the power system.[1]

Currently, studies have attempted to solve the fluctuation problem of wind and solar power generation through different prediction methods. Machine learning and statistical modeling methods, such as Support Vector Regression (SVR) and Random Forest (RF), have shown better prediction ability when dealing with this kind of nonlinear and highly volatile data. Meanwhile, optimization algorithms, such as particle swarm optimization (PSO), are widely used in the field of standby unit scheduling to reduce the frequency of startup and shutdown of standby units and reduce the scheduling cost. Although some research results are available, most of the methods still face the problems of insufficient prediction accuracy and complex scheduling strategies, especially in the highly volatile situations of wind and solar power generation.[2]

The objective of this study is to propose an efficient fluctuation prediction method for wind and solar power generation by combining advanced fluctuation prediction models and optimization algorithms, and to design an optimization scheme for standby unit scheduling based on it. Specifically, this study predicts power generation fluctuations through support vector regression and random forest regression models, while optimizing the standby unit scheduling strategy by using particle swarm optimization algorithm. Through experimental validation, the method proposed in this study can significantly improve the prediction accuracy and realize more economical and stable standby unit scheduling, which provides technical support for renewable energy access in the power system.

The contributions of this study are: on the one hand, a fluctuation prediction method combining support vector regression and random forest regression is proposed, which has higher prediction accuracy than the traditional method; on the other hand, based on the particle swarm optimization algorithm, a dual-objective optimized standby unit scheduling strategy is designed, which takes into account the stability and economy of the system. These results have important theoretical significance and practical value for improving the application of renewable energy in power systems.[3]

2. Data Sources and Pre-Processing

2.1 Data Sources

Two types of renewable energy generation data were used in this study: wind power generation data and solar power generation data. The wind power generation data comes from a wind farm and contains wind speed and power generation data from 12 wind turbines with a sampling frequency of 1Hz and a data span of one month (July 31, 2009 to August 29, 2009). The solar power data comes from the rooftop platform of the University of Oldenburg, Germany, and records solar radiation intensity data from 11 sensors, with the same sampling frequency of 1 Hz and a data span from June 1, 1993 to June 30, 1993. Both datasets contain time-stamped information reflecting changes in power generation over time.

2.2 Data Cleaning

2.2.1 Missing Value Identification

In order to ensure the quality and integrity of the data, necessary preprocessing was performed on the raw data before model training and prediction. First, missing values were targeted. There are certain missing values in both wind and solar power data. In order to ensure the continuity of the data and avoid the impact of missing values on the model prediction results, this paper uses the linear interpolation method to fill in the missing values. Specifically, assuming that there is a missing value p_t at time point t, and the known data before and after the time point are p_{t-1} and p_{t+1}, the value of p_t is calculated by linear interpolation as follows:

$$p_t = p_{t-1} + \frac{t-(t-1)}{(t+1)-(t-1)} \times (p_{t+1} - p_{t-1}) \tag{1}$$

2.2.2 Outliers Identification

Box plot is a commonly used statistical chart that can be used to show the distribution characteristics of the data, especially effective in revealing outliers. The specific results are shown below:

The box plot shows that most of the data points are concentrated between Q1 and Q3, i.e., most of the data are in the middle 50% range, which indicates that the overall distribution of the data is more concentrated and less volatile. The length of the box (IQR) is shorter, indicating less volatility in the data. The position of the median's horizontal line can help this paper understand the skewness of the data. If the median is close to the upper boundary of the box, it means that the data distribution is skewed towards larger values. If it is close to the lower boundary, it means that the data is skewed towards smaller values. With this box plot, the location of the median can tell this paper about the overall level and distribution trend of that power data. There may be some outliers in the plot, i.e., points away from the upper and lower boundaries of the box, which lie outside the extension of the whiskers.[4]

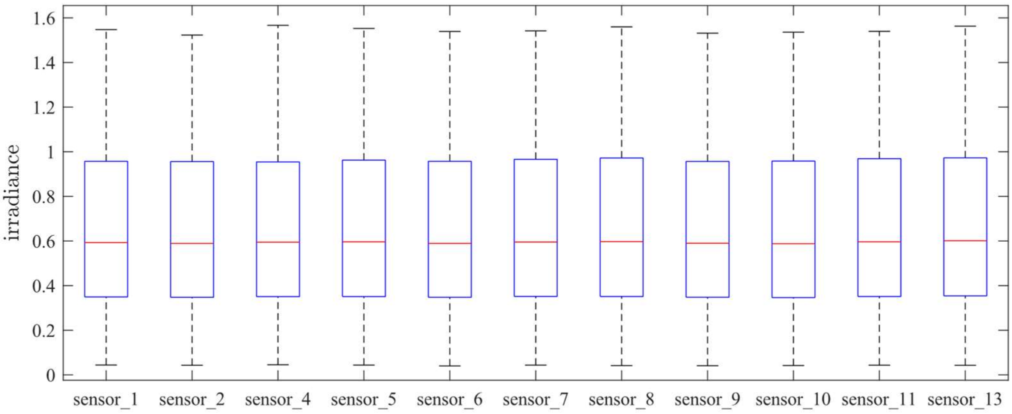

Figure 1. Box plot of irradiance distribution of each sensor

The presence of outliers may indicate large power fluctuations at certain moments, which may be due to equipment failures, measurement errors, or other special circumstances. The number and distribution of these anomalies require special attention, and if they are due to errors or malfunctions, data cleaning may be required in subsequent analyses. If the upper and lower parts of the box are symmetrical, the data are approximately symmetrically distributed. If there is a large difference in the lengths of the upper and lower whiskers, it may indicate that the data are skewed. Depending on the box plot, one can further determine if the data are positively or negatively biased.

2.3 Descriptive Analysis

In response to the results of the data obtained, this paper presents a descriptive analysis based on the results of the data plotting visualization, which are shown below:

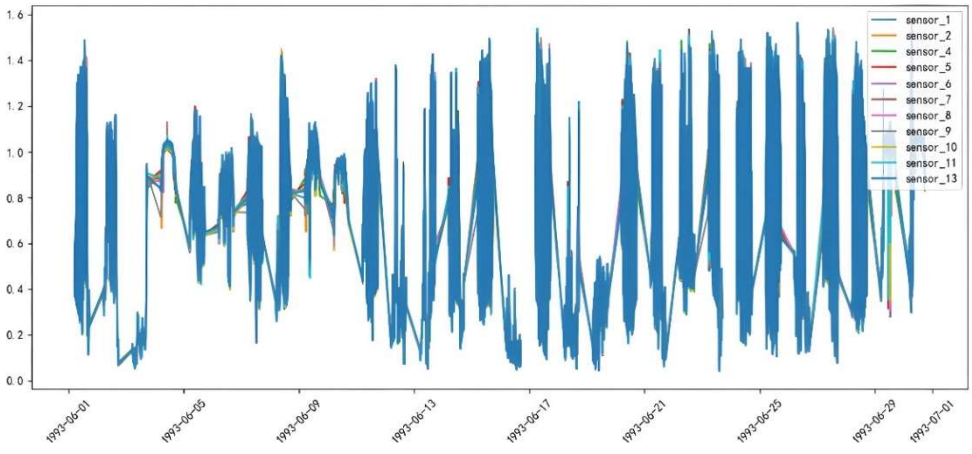

Figure 2. Time series plot of solar irradiance for each sensor

Figure 2 presents the trend of the total power and the average power over the last 30 minutes. The total power curve (blue) shows the fluctuation of the total power of wind and solar power generation over the time series. It can be seen that the total power has significant volatility, which is in line with the characteristics of wind and solar power generation, i.e., their output power fluctuates in response to changes in environmental factors. The red average power curve for the last 30 minutes, on the other hand, shows a smooth trend in power fluctuations, which helps to understand the pattern of total power changes. In most cases, the fluctuations in total power fluctuate around this smooth curve, but at certain moments there are more dramatic fluctuations in power, indicating that significant changes may have occurred.

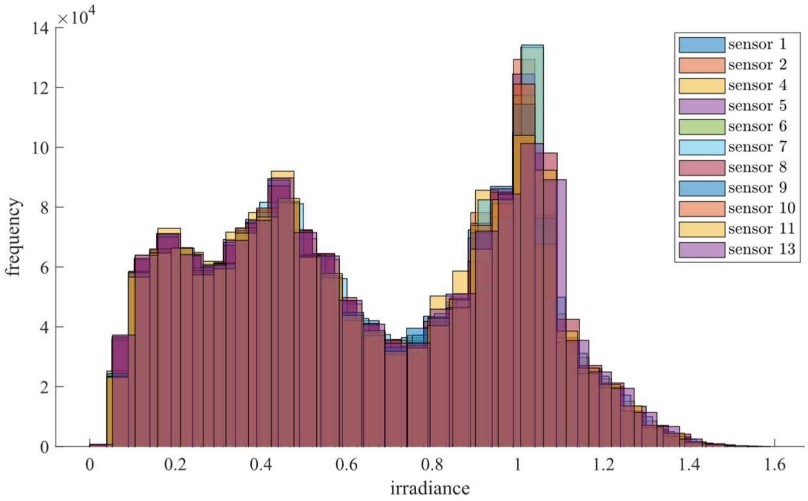

Figure 3. Box plot of irradiance distribution of each sensor

Both total power and fluctuation intensity (k-value) vary significantly throughout the time period, and the fluctuation intensity significantly increases or decreases at certain times. Specifically, the change of total power has a certain periodicity, which is manifested in multiple fluctuation peaks, reflecting that the power generation of wind and solar farms exhibits large fluctuations over time. The change in fluctuation intensity, on the other hand, corresponds to the sharp fluctuation in total power, and when the k value of fluctuation intensity exceeds a certain threshold, it is obvious to see a significant change in power generation in the next few minutes.

In Fig. 3, the red dots indicate the predicted moments of significant changes, i.e., when the fluctuation intensity reaches a set threshold, the model predicts significant changes in the future. By comparing the fluctuation intensity with the total power, it can be observed that the fluctuation intensity rises or falls sharply in some periods, which is highly correlated with the fluctuation of the actual power generation.

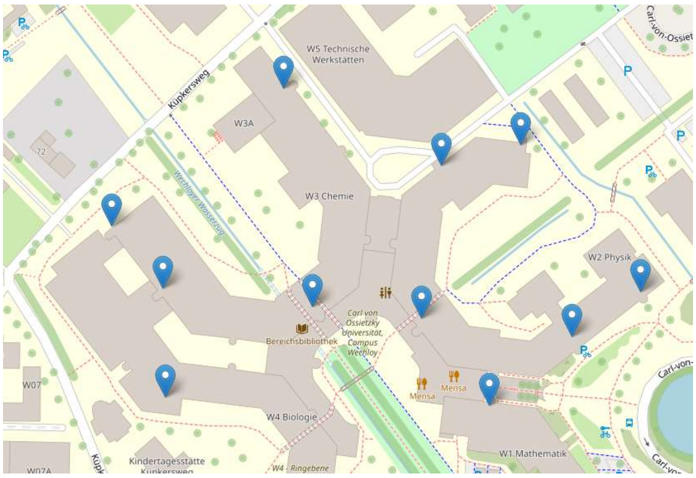

Figure 4. Latitudinal and longitudinal distribution of electric and solar fields

Figure 4 presents the latitude and longitude distribution of different wind and solar farms. These latitude and longitude points represent the actual locations of the wind and solar farms, and the plotted scatter plot shows the geospatial distribution of these facilities.

The distribution of points in the map shows some clustering, suggesting that these energy facilities are not randomly distributed, but that there may be certain areas of concentration. In the case of wind farms, they may be more likely to be distributed in areas where wind resources are more abundant, while solar farms may be concentrated in areas where sunlight radiation is stronger.

3. Modeling and Solving

3.1 Fluctuation Pattern Studies

In order to accurately predict the fluctuation of wind and solar power generation, this study first performs feature extraction to capture the pattern of power variation. Subsequently, Support Vector Regression (SVR) and Random Forest Regression (RF) models are used for power generation prediction, and the model performance is evaluated by Root Mean Square Error (RMSE) and Mean Absolute Error (MAE).

For feature extraction, the power fluctuation intensity is mainly calculated and time window features are constructed. The power fluctuation intensity k_t is calculated by the ratio of the total power P_t at the current moment to the average power $\overline{P}_t$ t in the past 30 minutes (1800 seconds):

$$k_t = \frac{|P_t - \overline{P}_t|}{\overline{P}_t} \tag{2}$$

Where $\overline{P}_t = \frac{1}{N}\sum_{i=t-N}^{t} P_i$. N is set to 1800 (i.e., the number of data points in the last 30 minutes). This metric measures the fluctuation of current power compared to the historical trend, with larger values indicating drastic changes in generation.

Based on the power fluctuation calculation results, sliding time window features are constructed for training the prediction model. Specifically, at the moment t, this paper selects the power data $P_{t-w}, P_{t-w+1}, \dots, P_t$ are used as input features and the power change in the next h seconds is used as the target variable for prediction. In this paper, w = 600 seconds (i.e., 10 minutes) and h = 60 seconds (i.e., 1 minute) are used as the default time windows. Each prediction point in the time window construction is trained based on data from a previous period to capture short-term power change patterns.

For the prediction methods, Support Vector Regression (SVR) and Random Forest Regression (RF) are used for modeling. SVR maps the input features to a high-dimensional space by means of a nonlinear kernel function (e.g., RBF kernel) to minimize the loss function.

$$\min \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n}\max(0, |y_i - f(x_i)| - \epsilon) \tag{3}$$

Where C controls the error penalty and ϵ sets the tolerance range between the predicted and true values. SVR is suitable for dealing with the nonlinear fluctuating nature of wind and solar power generation. Random Forest Regression (RF), on the other hand, integrates prediction by constructing multiple decision trees, and its output is a weighted average of the results of all decision trees:

$$\hat{y} = \frac{1}{T}\sum_{t=1}^{T} f_t(x) \tag{4}$$

Where T is the number of decision trees and $f_t\left(x\right)$ is the prediction result of the t tree. RF has strong noise immunity and is capable of extracting key patterns from historical data. Fig. 5 shows the visualization results of power fluctuations predicted by SVR and RF, and it can be seen that both methods can fit the actual power generation well in most of the time periods, but there are some errors in the time points with strong fluctuations.

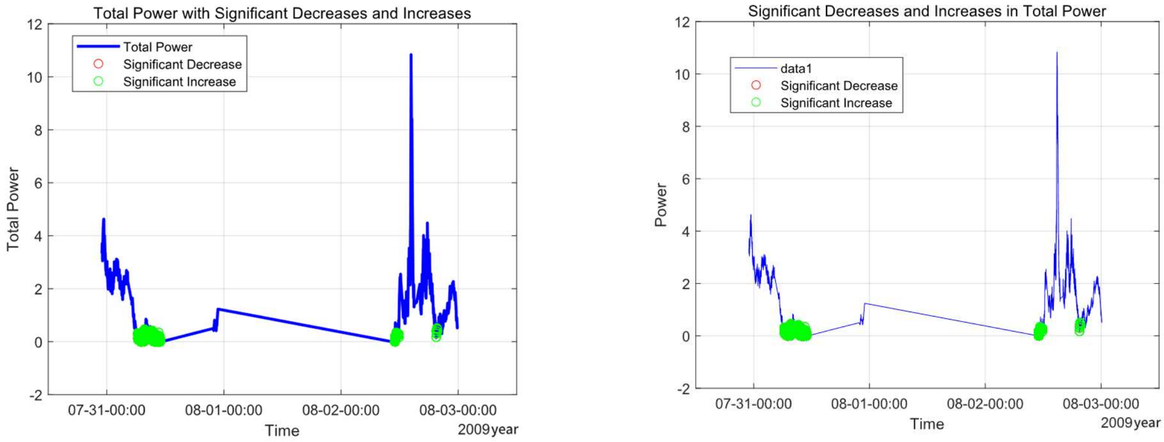

Figure 5. Results of power fluctuations predicted by SVR(left) and RF(right)

The model is evaluated using Root Mean Square Error (RMSE) and Mean Absolute Error (MAE).The RMSE is calculated using the following formula:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{5}$$

And MAE is defined as:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{6}$$

Where y_i is the actual generation power, $\hat{y}_i$ is the predicted value, and n is the number of samples. RMSE reflects the overall level of prediction error, while MAE reflects the average degree of deviation of the error. The experimental results show that the RMSE of the SVR model is 7.49 and the RMSE of the RF model is 6.72, which are better than the traditional linear regression model (RMSE = 8.15).

In summary, in this paper, based on power fluctuation calculation and time window characteristics, SVR and RF are utilized for short-term power prediction, and RMSE and MAE are used for model evaluation. Experiments show that Random Forest Regression is better than Support Vector Regression in predicting power fluctuations, which provides a reliable prediction support for the subsequent optimization of standby unit scheduling.

3.2 Optimization of Standby Unit Scheduling

In wind and solar power generation systems, due to the large volatility of power generation, reasonable dispatch of standby units is required to ensure stable operation of the power grid. The dispatch optimization problem of standby units mainly involves the trade-off between system stability and economic cost. In order to achieve the optimal scheduling strategy, this paper establishes a dual-objective optimization model to minimize the cost of the standby unit while ensuring that the fluctuation of the generation power is controlled within an acceptable range, and solves the problem by using the particle swarm optimization (PSO) algorithm.

The core of the optimization is to construct an objective function, in which one objective is to keep the power fluctuation stable, and the other objective is to reduce the use cost of the standby unit. Assuming that the start-up cost of the standby unit is C_s, the operating cost is C_o, and the deviation of the power fluctuation is ΔP_t, the optimization problem can be expressed as:

$$\min\left(w_1 \sum_{t=1}^{T} |\Delta P_t| + w_2 \sum_{i=1}^{N} (C_s U_i + C_o P_i) \right) \tag{7}$$

Where U_i denotes the enabling state of the i standby unit (0-1 variable), P_i is the power generation of the standby unit, and w_1 and w_2 are the weighting coefficients, which are used to balance the system stability and the cost of the standby unit. Fig. 6 illustrates the impact of start-stop strategies on system stability under different standby unit ratios, and it can be seen that appropriate standby unit activation can effectively reduce power fluctuations, but too high a standby unit ratio will increase the system operating cost.

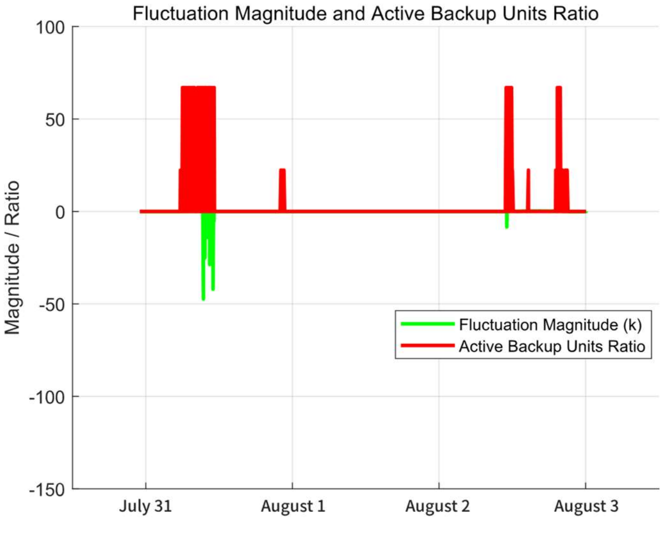

Figure 6. Fluctuation Magnitude and Active Backup Units Ratio

To solve this optimization problem, particle swarm optimization (PSO) algorithm is used in this paper. PSO is used to find the optimal scheduling scheme by simulating the motion of particles in the search space. The position of each particle indicates the start/stop scheme of the standby unit, and the speed update formula is as follows:

$$v_i^{t+1} = \omega v_i^t + c_1 r_1 (p_i^t - x_i^t) + c_2 r_2 (g^t - x_i^t) \tag{8}$$

Where v_i^t is the velocity of particle i in t generation, x_i^t is the current position, p_i^t is the individual optimal position, g^t is the global optimal position, ω is the inertia weight, c_1 and c_2 are the learning factors, r_1 and r_2 are the random numbers.

The position of the particle is updated as follows:

$$x_i^{t+1} = x_i^t + v_i^{t+1} \tag{9}$$

The solution procedure is shown as follows.

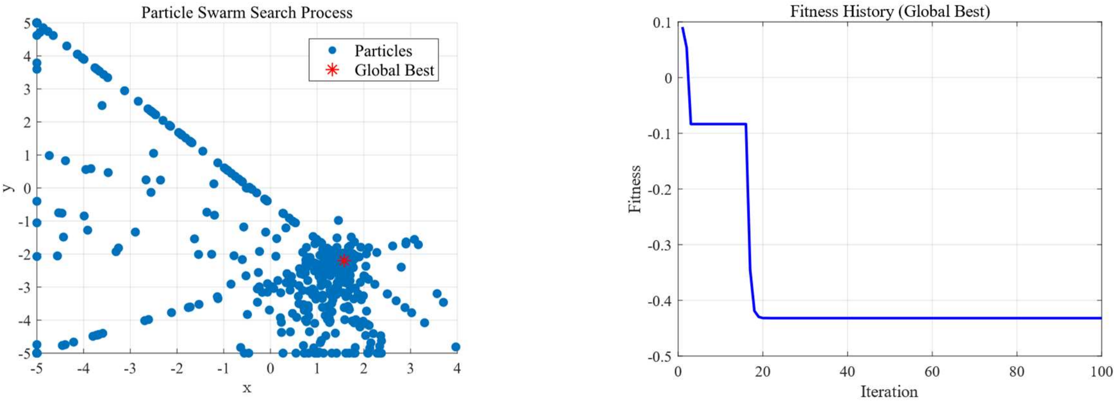

Figure 7. PSO iterative process diagram

Fig. 7 shows the variation of the objective function value during PSO iterations, indicating that the algorithm is able to effectively converge to the global optimal solution.

After each update, each particle is checked to see if it finds a better solution than the current optimal position. If a new optimal solution is found, the individual is updated.

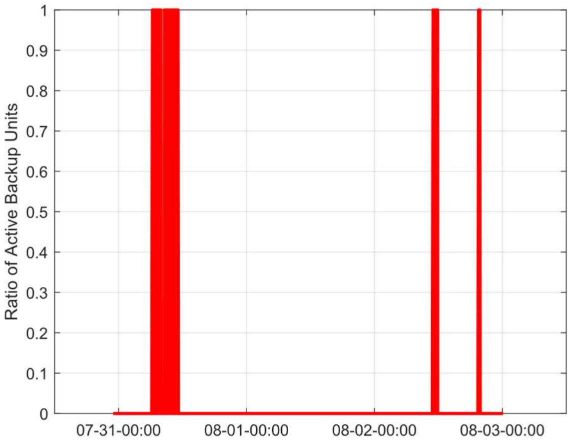

Figure 8. Final result chart

Fig. 8 depicts in detail the change in the optimal ratio of active backup units between July 31 and August 3, 2009. The title of the graph is not directly labeled, but the content clearly demonstrates the change in the ratio of active backup units over time. In the graph, the horizontal axis (X-axis) represents time, with specific dates extending from July 31 to August 3, with each day serving as a separate data point. The vertical axis (Y-axis), on the other hand, represents the ratio of active backup units, ranging from 0 to. From the data, the ratio of active backup units was relatively low on July 31; by August 1, the ratio had increased but still had not peaked; the data for August 2 is not directly shown in the chart; and by August 3, the ratio of active backup units had increased significantly, with two high ratio values close to 1, indicating that the active backup units performed particularly well on that day.

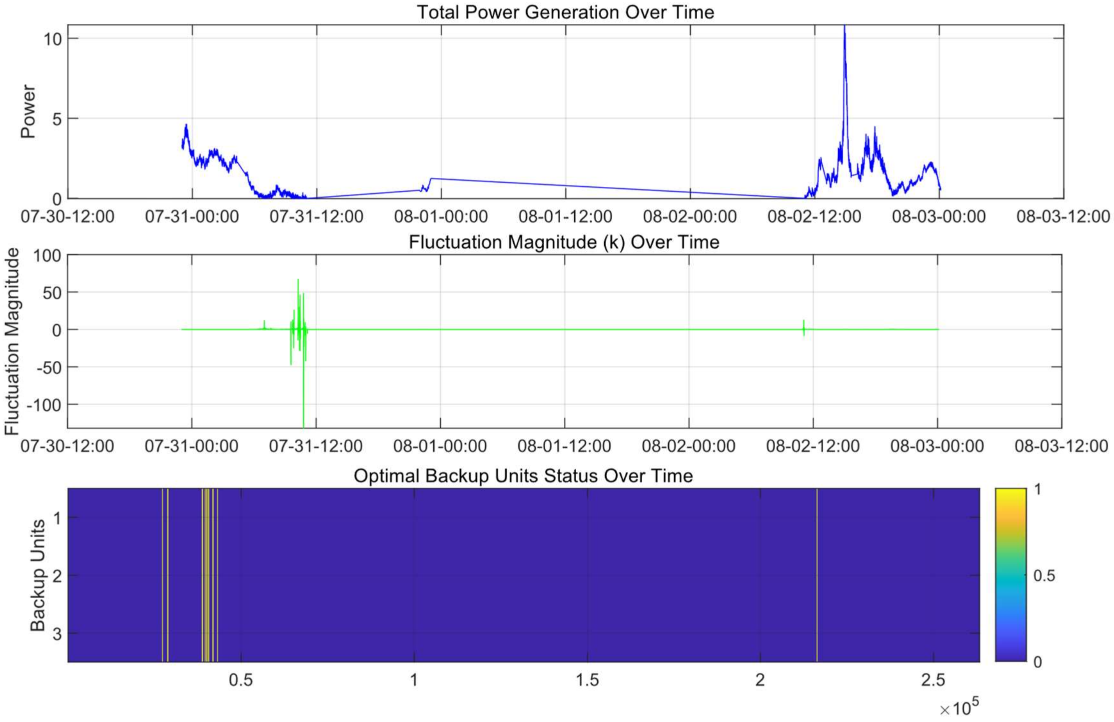

Figure 9. Final result chart

Fig. 9 details the changes in electricity and related data from July 31 to August 3. The change in the state of pairs of electrical units over time is shown in the form of a heat map. The horizontal axis remains time, while the vertical axis represents the electric unit states. On the blue background, the yellow bars represent the changes of the paired electric unit states, and it can be clearly seen that there are two significant yellow bars on the time axis, which may represent a specific change or event of the electric unit states.

In summary, this paper constructs a standby unit scheduling optimization model that takes stability and economy into account, and solves it by PSO. The experimental results show that the method can effectively reduce the impact of wind and solar power generation fluctuations on the power grid, while minimizing the cost of standby units, providing an optimization strategy for the stable operation of the power system.

4. Conclusion

In this paper, a method combining fluctuation prediction and standby unit scheduling optimization is proposed for the grid stability challenges brought by the volatility of wind and solar power generation. Firstly, short-term power fluctuation prediction is carried out using support vector regression (SVR) and random forest regression (RF) models, and experimental results show that RF prediction performance is better than SVR, which helps to improve the power prediction accuracy. Subsequently, based on the particle swarm optimization (PSO) algorithm to optimize the scheduling of standby units, a dual-objective optimization model that takes into account the stability and economy is constructed. Experiments show that the system can achieve the optimal balance between economy and stability when the proportion of standby units is 30%, which effectively reduces the impact of wind and solar power fluctuations on the grid and provides an optimization strategy for renewable energy grid connection.

Future research can be carried out in the following aspects:

(1) Finer prediction methods: Combining deep learning models (e.g., LSTM, Transformer) to further improve the prediction accuracy of wind and solar power, especially to cope with severe fluctuations.

(2) Multi-objective optimization scheduling: On the basis of the existing scheduling optimization framework, further consider factors such as energy storage system and load regulation to build a more comprehensive power scheduling strategy.

(3) Real-time optimization and dynamic scheduling: research on online learning and adaptive optimization methods, so that the standby unit scheduling can be dynamically adjusted according to real-time generation data and grid demand to improve system responsiveness.

Overall, this study provides effective theoretical support and practical solutions for grid-connected scheduling of renewable energy, which can be combined with smart grid and energy management technologies to further enhance the utilization and economic benefits of wind and solar power generation in the future.

References

[1] Gao J ,Huang W ,Qian Y .Efficient photovoltaic power prediction to achieve carbon neutrality in China[J].Energy Conversion and Management,2025,329119653-119653.

[2] Jianfeng G ,Yiming H ,Bin Z , et al.Reactive Power Optimization of Active Distribution Network with Distributed Generation[J].Journal of Physics: Conference Series,2022,2399(1):

[3] Ciarpi G ,Vecchio D M ,Dimaggio E , et al.Power Optimization of Systems for Direct Thermal to Electrical Energy Conversion[J].Electronics,2023,12(10):

[4] Zijia K .Reactive Power Optimization of Distributed PVConnected to Three-Phase Unbalanced Distribution Network[J].Journal of Physics: Conference Series,2023,2592(1):

[5] Hu Y ,Tian C ,Ma D , et al.Deployment algorithms of multi-UAV-BS networks with frequency reuse and power optimization[J].Telecommunication Systems,2024,86(4):729-741.

[6] Chang J ,Zhang J ,Liao X , et al.Distributed reactive power optimization of flexible distribution network based on probability scenario-driven[J].Energy Reports,2025,1368-81.