Preprints.org

Article

Not peer-reviewed version

An Adaptive Photovoltaic Power Forecasting Approach Integrating Double Q-Learning and Stacking Ensemble

Jiajun Wang , Chen Ye , Yudong Yang * , Yulong Pan , Yabo Sun , Jianbo Jiang

Posted Date: 27 February 2026

doi: 10.20944/preprints202602.1890.v1

Keywords: photovoltaic power forecasting; Double Q-Learning; stacking ensemble; dynamic weighting; XGBoost; adaptive prediction

doi:10.20944/preprints202602.1890.v1

Article

An Adaptive Photovoltaic Power Forecasting Approach Integrating Double Q-Learning and Stacking Ensemble

Jiajun Wang, Chen Ye, Yudong Yang *, Yulong Pan, Yabo Sun and Jianbo Jiang

School of Engineering, Dali University, Dali, China
* Correspondence: 2456165876@qq.com

Abstract

This paper proposes an adaptive photovoltaic (PV) power forecasting approach integrating Double Q-Learning and Stacking ensemble with XGBoost meta-learner to address the poor adaptability of conventional methods in real-time grid-connected PV systems. Unlike single models with limited generalization or fixed-weight ensembles that merely imitate experience superficially, the proposed approach adapts dynamically to time-varying meteorological and operational conditions. It pre-trains three complementary base models, namely RF, SVR and LightGBM, constructs a Stacking framework with XGBoost as the secondary-learner to generate high-precision baseline predictions via out-of-fold validation, and embeds a Double Q-Learning agent to output adaptive weights by capturing meteorological-temporal features and real-time prediction errors. The final prediction is obtained by fusing the Stacking output and Double Q-Learning adjusted base model outputs. Tests on a 50MW PV station dataset show it outperforms four single models and traditional ensembles in MAE, MSE, RMSE, and R^2, enabling reliable, generalized and adaptive real-time predictions.

Keywords: photovoltaic power forecasting; Double Q-Learning; stacking ensemble; dynamic weighting; XGBoost; adaptive prediction

---

I. Introduction

In response to global carbon neutrality goals[1], the large-scale development of renewable energy has become the core of energy structure transformation. Photovoltaic power generation, leveraging advantages such as abundant resources, cleanliness and pollution-free operation, and flexible deployment[2], has emerged as a key power source in power systems. Accurate PV power forecasting is a core technology for ensuring grid safety, enhancing absorption capacity, and optimizing dispatching[3], which is crucial for reducing curtailment rates.

However, PV power output is influenced by meteorological factors such as irradiance and temperature, exhibiting temporal non-stationarity, strong nonlinearity, and random volatility, posing significant challenges to precise forecasting. Existing single models[4-7] each have their strengths but are limited in learning capacity, making it difficult to adapt to complex data characteristics. Their accuracy tends to decline under extreme operating conditions[8]. Although the heterogeneous Stacking ensemble strategy can integrate the advantages of base learners [9], traditional static weight allocation methods [10] fail to adapt to dynamic data changes, resulting in limited generalization ability under extreme conditions.

Reinforcement learning provides a new approach for dynamic weight optimization [11], but traditional Q-learning suffers from Q-value overestimation bias. Double Q-Learning effectively suppresses overestimation and improves optimization stability by separating action selection and value evaluation through two independent Q-tables [12]. To address these issues, this paper proposes an adaptive PV power forecasting method integrating Double Q-Learning and Stacking, which uses heterogeneous models as base learners and dynamically optimizes weights via Double Q-Learning.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 27 February 2026 doi:10.20944/preprints202602.1890.v1

Its innovations are reflected in: 1) For the first time, deeply integrating Double Q-Learning with Stacking to break through the limitations of static weights; 2) The dual-table collaboration suppresses overestimation and improves the accuracy of weight optimization; 3) Targeted design of the state space to achieve real-time matching between weights and data characteristics.

The structure of this paper is organized as follows: Section II details the architecture of the Stacking model and selection of heterogeneous base learners. Section III elaborates on the Double Q-learning mechanism Section IV describes data, experimental settings, and verifies the model's performance via comparative and robustness tests. Section V ummarizes core findings and outlines future research directions.

II. Proposed PV Power Forecasting Model

A. Principle of Stacking Ensemble Learning

Stacking is an ensemble model consisting of two layers of base learners. The first layer is the primary learner, which is responsible for feature mining on raw PV data and generating preliminary prediction results that are strongly correlated with input data , such as meteorological, temporal, and other features. The second layer is the secondary-learner. Its structure is shown in Figure 1, which takes the prediction results of the primary learners as new features to complete the final predictive integration—its core architecture is well-suited for PV forecasting scenarios.

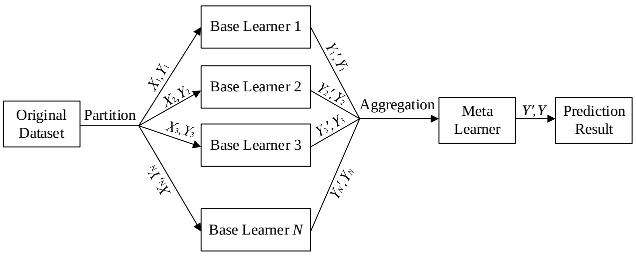


Figure 1. Schematic of Stacking Method.

Stacking excels at integrating the complementary capabilities of different models: it can both capture the periodic patterns of PV power and fit the nonlinear relationships between meteorological factors and power output, making it a high-quality ensemble method for PV power forecasting. During data training, Stacking uses the k-fold cross-validation method to split the dataset. This ensures that the primary learner in each training round is trained on a different subset of samples, and outputs predictions for the subset that did not participate in training. This approach not only avoids data leakage but also improves data utilization efficiency. Its specific workflow is as follows:

*1)*Input the original dataset K={(a_1,b_1),(a_2,b_2),…,(a_n,b_n)}, where ai represents the features of PV data, and bi denotes the corresponding true power value. Define a set of heterogeneous primary learners H_1,H_2,…,H_t and a secondary learner H_0.

*2)*Split the dataset K into k mutually exclusive subsets K_1,K_2,…,K_k. Sequentially select one subset as the validation set, and the remaining k–1 subsets as the training set.

*3)*For each primary learner H_j, train it using the training set in each round, then output predictions for the corresponding validation set. Combine the validation set predictions from all rounds to obtain the preliminary prediction matrix F of the primary learners for the entire dataset K.

*4)*Train the secondary learner H_0 using the preliminary prediction matrix F as new features and the true values b_i as labels.

*5)*During the prediction phase, first use all primary learners to output preliminary predictions for new PV samples, then input these predictions into the trained secondary learner H_0 to get the final power prediction result.

B. Selection Basis of Primary Learners

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 27 February 2026 doi:10.20944/preprints202602.1890.v1

Stacking ensemble learning typically adopts a heterogeneous strategy, which requires base learners to be sufficiently diverse to leverage the complementary advantages of ensemble learning. Given that no universal algorithm fits all supervised learning tasks, Stacking must also balance task requirements and data characteristics to achieve well-rounded performance improvements. For PV power forecasting, this study selects RF and SVR and LightGBM as base learners after comprehensively evaluating algorithm adaptability and data compatibility; their specific performance metrics are presented in Table 1.

Table 1. Primary Learner Performance of Stacking.

Algorithm	Applicable Data	Training Speed	Key Features	Limitations
RF	Large-scale data	Relatively fast	Random sampling ensemble	Weak temporal modeling
SVR	low-to-medium dimensions	Slow	Non-linear fitting via kernel trick	Computationally expensive
LightGBM	Large-scale data	Very fast	Efficient parallel training	Leaf-wise overfitting risk

All three algorithms are theoretically rigorous and perform well in prediction tasks. RF generates multiple independent sample sets via random sampling for separate training and prediction, enabling efficient processing of large-scale nonlinear data with strong generalization ability to fit PV data's complex features. LightGBM is a targeted improvement over traditional gradient boosting algorithms; it discretizes samples using a leaf-wise growth strategy, optimizes feature calculation with histogram algorithms, and enhances memory utilization and operational efficiency to meet the training needs of large-scale PV datasets. SVR, another representative algorithm for regression tasks, excels in capturing intricate nonlinear relationships in time-series PV power data and features robust regularization mechanisms, making it highly adaptable to the dynamic fluctuations of PV output.

The secondary learner takes meta-features generated by the base learners. These features often exhibit complex nonlinear correlations, and since the secondary learner directly determines the final prediction accuracy of the ensemble model, it is more susceptible to overfitting. Thus, the secondary learner algorithm must possess strong nonlinear fitting capability, comprehensive overfitting control mechanisms, and good adaptability to ensemble learning scenarios. Considering the complexity and time cost of the ensemble model training process, XGBoost is selected as the secondary learner to balance fitting accuracy and training efficiency. It can accurately capture the nonlinear correlations among meta-features, suppress overfitting through built-in L1 and L2 regularization, and meet the demands of ensemble training with efficient gradient calculation and parallel processing capabilities. Compared with linear regression, XGBoost better integrates the strengths of the base learners, further improving the prediction accuracy of the Stacking model and maximizing the value of heterogeneous ensemble learning.

 doi:10.20944/preprints202602.1890.v1

III. Weight Optimization Algorithm for Ensemble Models Based on Double Q-Learning

In Stacking ensemble models, the weight allocation of base learners directly affects the ensemble prediction performance. Traditional static weight strategies are difficult to adapt to the temporal non-stationarity of photovoltaic power data and the dynamic characteristics of meteorological operating conditions, which easily leads to insufficient generalization ability of the model in complex scenarios. To address this, this paper adopts an improved Double Q-learning algorithm to optimize the weight allocation strategy of the ensemble model. This algorithm introduces a double Q-table mechanism based on traditional Q-learning, which improves the convergence speed and stability of weight optimization.

A. Objective Function and Constraints

The core optimization goal of this algorithm is to adjust the weight allocation of each base learner under different data distribution scenarios, thereby minimizing the prediction error of the ensemble model. The specific objective function is as follows:

$$E = \min(\mathrm{RMSE}) = \min\left(\sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\hat{P}_i - P_i\right)^2}\right) \tag{1}$$

where $\hat{P}_i$ is the predicted power of the ensemble model, P_i is the actual power value, and N is the number of samples.

At the same time, the weight allocation must satisfy the non-negativity and normalization constraints, the weight of each base learner is not less than 0, and the sum of all weights is 1. The constraints are as follows:

$$\begin{cases} \omega_j \ge 0 \\ \sum_{j=1}^{M} \omega_j = 1 \end{cases} \tag{2}$$

where ω_j is the weight of the j-th base learner, and M is the number of base learners.

B. Double Q-Learning Mechanism

In traditional Q-learning, the update of Q-values often relies on a single Q-table, which is prone to the problem of overestimation bias. Double Q-learning splits the update task by introducing two independent Q-tables ,denoted as Q1 and Q2, and each Q-table uses the Q-value of the other Q-table to update the next state, thereby reducing the risk of overestimation. In Double Q-learning, action selection and Q-value update are performed alternately between Q1 and Q2, and the update formulas for the two Q-tables are as follows:

$$\begin{cases} Q_1(s,a) \leftarrow Q_1(s,a) + \alpha\left[r + \gamma Q_2(s',a^*) - Q_1(s,a)\right] \\ Q_2(s,a) \leftarrow Q_2(s,a) + \alpha\left[r + \gamma Q_1(s',a^*) - Q_2(s,a)\right] \end{cases} \tag{3}$$

where s is the current state, a is the currently selected weight allocation action, α is the learning rate, r is the immediate reward, γ is the discount factor, $s'$is the next state, and a^* is the optimal action selected in the next state.

In this algorithm, the introduction of the double Q-table mechanism effectively reduces the risk of overestimating action values in weight optimization that exists in traditional Q-learning, ensuring the stability of the algorithm.

IV. Experimental Analysis

doi:10.20944/preprints202602.1890.v1

To verify the effectiveness, robustness, and optimization efficiency of the proposed hybrid stacking adaptive PV power forecasting model, comprehensive experiments are designed and conducted. This section details the experimental setup, data preparation, comparative models, and evaluation metrics, followed by in-depth analysis of the experimental results to validate the model's superiority.

A. Data Source and Preprocessing

The experimental data are derived from the open dataset of the Chinese State Grid Renewable Energy Generation Forecasting Competition[13], which contains on-site operational and meteorological data from six wind farms and eight solar stations across China. This dataset was collected over two consecutive years (2019–2020) with a 15-minute sampling interval, ensuring sufficient temporal resolution to capture the dynamic fluctuations of renewable energy output.

The data preprocessing process strictly follows engineering specifications: outliers caused by sensor failures and equipment maintenance are removed using the 3σ principle; missing data accounting for less than 0.5% is supplemented by linear interpolation; all features are mapped to the [0, 1] interval through min-max normalization to eliminate the influence of dimensional differences. The dataset is randomly divided into training and test sets at a ratio of 7:3.

B. Comparison Models and Results Analysis

To fully demonstrate the superiority of the adaptive Stacking model, 5 typical models are selected as baselines, covering single base learners, traditional Stacking variants, to verify the effectiveness of heterogeneous ensemble and Double Q-Learning based dynamic weight optimization.

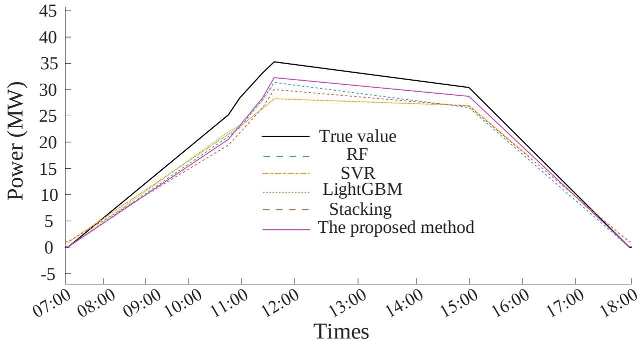
Table 2 presents the comprehensive prediction performance of all models on the full test set. It can be clearly observed that:

Table 2. Overall Prediction Performance Comparison of Different Models.

Model	MAE	MSE	RMSE	R^2
RF	1.5779	13.3648	3.6558	0.9288
SVR	1.9272	20.4287	4.5198	0.8912
LightGBM	1.5941	12.8810	3.5890	0.9314
Stacking	2.2478	14.3770	3.7917	0.9234
The proposed method	1.5281	12.5000	3.5355	0.9334

Photovoltaic power generation is significantly affected by weather conditions. To verify the superiority of the proposed model in short-term prediction under different weather conditions, the prediction results of two typical days—sunny day and rainy day—are selected for comparative analysis, The prediction results are shown in Figure 2.

doi:10.20944/preprints202602.1890.v1



a.Sunny day

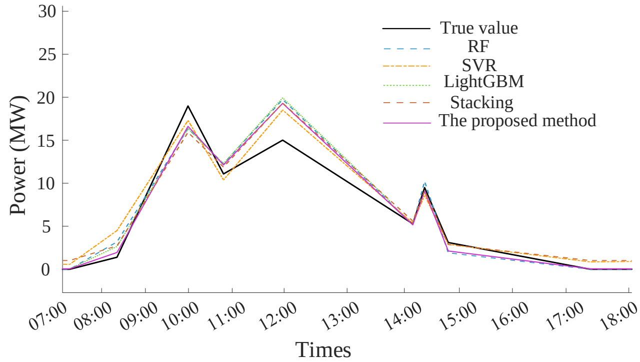


Figure 2. Comparison of Prediction Results of Five Models Under Different Weather ConditionsIdentify the Headings.

As shown in Table 2 and Figure 2, compared with the single SVR model, the RF model reduces the RMSE by 16.27% and increases R^2 by 4.17%, indicating that the ensemble strategy can effectively improve the model's fitting ability for photovoltaic data. However, when combined with the LightGBM model, the traditional static Stacking model increases the RMSE by 5.61% and decreases R^2 by 0.86% compared with the single LightGBM model, suggesting that a simple static ensemble strategy cannot effectively integrate the advantages of base learners and even causes performance degradation. The proposed method achieves the lowest MAE, MSE, and RMSE as well as the highest R^2 among all models; compared with the traditional Stacking model, it reduces the RMSE by 6.76% and improves R^2 by 1.08%, and compared with the best-performing single model LightGBM, it reduces the RMSE by 1.38% and improves R^2 by 0.21%, demonstrating that the dynamic weight optimization strategy proposed in this paper can effectively break through the performance bottleneck of traditional ensemble methods and significantly improve the accuracy of photovoltaic power prediction.

V. Conclusions and Prospects

This study focuses on the key challenges in short-term photovoltaic power forecasting, such as data non-stationarity, limitations of static weight strategies, and overestimation bias in traditional reinforcement learning. A hybrid Stacking model optimized by Double Q-learning is proposed for dynamic weight allocation, combining heterogeneous ensemble learning and reinforcement learning while alleviating overestimation. Experiments show that the model outperforms single models on RMSE, MSE, MAE, and R^2. Future work will extend the model to medium- and long-term forecasting, expand feature dimensions, improve interpretability and efficiency, integrate deep learning and multi-agent reinforcement learning, and combine it with edge computing and smart grid systems for practical engineering applications.

Funding: This work was supported in part by Yunnan Fundamental Research Projects under Grant 202401AT070078, in part by Yunnan College Students' Innovation Training Projects under Grant S202510679050 and in part by the Undergraduate Research Project of Dali University under Grant KYSX2025252.

References

1. International Energy Agency (IEA), "Net zero roadmap: A global pathway to keep the 1.5 °C goal in reach – 2023 update," IEA Report, Revised ed., Nov. 2024, [Online]. Available: https://www.iea.org.
2. H. Chen, W. Wu, C. Li, G. Lu, D. Ye, C. Ma, L. Ren, and G. Li, "Ecological and environmental effects of global photovoltaic power plants: A meta-analysis," Journal of Environmental Management, vol. 373, 2025, art. 123785, doi: 10.1016/j.jenvman.2024.123785.
3. P. Di Leo, A. Ciocia, G. Malgaroli, and F. Spertino, "Advancements and challenges in photovoltaic power forecasting: A comprehensive review," Energies, vol. 18, no. 8, p. 2108, 2025, doi: 10.3390/en18082108.
4. N. Kumari, K. Namrata, M. Kumar and R. P. Gupta, "Random Forest Algorithm for Solar Forecasting in Jamshedpur – India," 2022 4th International Conference on Energy, Power and Environment (ICEPE), Shillong, India, 2022, pp. 1-6, doi: 10.1109/ICEPE55035.2022.9797975.
5. Z. Ullah, S. M. Shahid, J. Park, B. Min, and S. Kwon, "LSTM-based online learning for real-time solar power forecasting," IEEE Access, vol. 13, pp. 160456–160465, 2025, doi: 10.1109/ACCESS.2025.3606910.
6. E. Safishahi, A. A. Safavi and P. Khalili, "Real-Time PV Power Forecasting Using LSTM-Attention, Transformer & XGBoost Hybrid Model," 2025 11th International Conference on Control, Instrumentation and Automation (ICCIA), Tehran, Iran, Islamic Republic of, 2025, pp. 1-4, doi: 10.1109/ICCIA69223.2025.11286049.
7. J. Wang, J. Jiang and W. Zhang, "Power Load Forecasting Based on GA-Optimized CNN-LSTM Model," 2025 IEEE 7th International Conference on Civil Aviation Safety and Information Technology (ICCASIT), Yinchuan, China, 2025, pp. 1555-1559, doi: 10.1109/ICCASIT66611. 2025.11348648.
8. P. Di Leo, A. Ciocia, G. Malgaroli, and F. Spertino, "Advancements and challenges in photovoltaic power forecasting: A comprehensive review," Energies, vol. 18, no. 8, p. 2108, 2025, doi: 10.3390/en18082108.
9. F. Schwenker, "Ensemble Methods: Foundations and Algorithms [Book Review]," in IEEE Computational Intelligence Magazine, vol. 8, no. 1, pp. 77-79, Feb. 2013, doi: 10.1109/MCI.2012.2228600.
10. J. Chen, E. Dou, and X. Li, "Dynamic weight allocation and extension optimization for intelligent evaluation," Procedia Computer Science, vol. 266, pp. 652–659, 2025, doi: 10.1016/j.procs.2025.08.082.
11. R. S. Sutton and A. G. Barto, Reinforcement Learning: An Introduction, Cambridge, MA, USA: MIT Press, 1998.
12. H. Mohammadi Rouzbahani, H. Karimipour and L. Lei, "Optimizing Resource Swap Functionality in IoE-Based Grids Using Approximate Reasoning Reward-Based Adjustable Deep Double Q-Learning," in IEEE Transactions on Consumer Electronics, vol. 69, no. 3, pp. 522-532, Aug. 2023, doi: 10.1109/TCE.2023.3279138.
13. Y. Chen and J. Xu, "Solar and wind power data from the Chinese State Grid Renewable Energy Generation Forecasting Competition," Scientific Data, vol. 9, p. 577, 2022, doi: 10.1038/s41597-022-01696-6.