International Journal of Power and Energy Systems, Vol. 45, No. 10, 2025
An Open Access Paper

EFFICIENCY OPTIMIZATION METHOD FOR PHOTOVOLTAIC POWER GENERATION SYSTEM BASED ON REINFORCEMENT LEARNING AND ADAPTIVE MODEL PREDICTIVE CONTROL

Zhichao Zhang,* Qihang Liu,* Chuyuan Wang,* and Yujie Zong*

Abstract

Photovoltaic power generation is an important component of achieving sustainable development of renewable energy, and improving the efficiency of photovoltaic power generation is crucial. This paper proposes an efficiency optimisation method for photovoltaic power generation systems based on reinforcement learning and adaptive model predictive control (MPC). The method combines reinforcement learning algorithms with MPC to optimise the control parameters through reinforcement learning algorithms, achieving dynamic adaptive control of photovoltaic power generation systems. Firstly, the reinforcement learning algorithms and interactive learning optimal control strategies are adopted in order to increase adaptability and robustness in different environmental conditions. Secondly, the rolling optimisation of predictive control is achieved to increase efficiency and stability in photovoltaic power generation systems. In addition, the adaptive control mechanism dynamically adjusts control parameters by monitoring environmental parameters and system status in real-time, ensuring that the system maintains optimal performance under various operating conditions. Finally, experimental results demonstrate that the proposed optimisation algorithm not only significantly increases accuracy and control efficiency of a system but also significantly boosts stability and reliability in complex environments for greater application potential.

Key Words

Photovoltaic power generation system, efficiency optimisation, enhance learning, adaptive control, model predictive control (MPC)

* College of Engineering and Physical Sciences, University of Birmingham, Edgbaston, United Kingdom B15 2TT; e-mail: zhichao_zhang123@163.com; {784773104, 1705384328}@qq.com; zongyujie590@gmail.com
Corresponding author: Zhichao Zhang

Recommended by Ramesh Devarapalli
(DOI: 10.2316/J.2025.203-0591)

1. Introduction

With advances in technology, photovoltaic power generation systems continue to gain in efficiency and reliability, leading to their deployment across more centralised photovoltaic power stations as well as distributed systems [1]–[3]. Centralised photovoltaic power stations tend to be located in areas where there is plenty of sunshine, generating electricity through large-scale photovoltaic cell arrays and transmitting it to the grid. The distributed photovoltaic power generation systems are more commonly used on the roofs of urban buildings, industrial parks, *etc.* [4], [5]. By generating and using electricity on-site, the energy transmission losses are reduced and energy utilisation efficiency is improved.

Although significant progress has been made in photovoltaic power generation technology, it still faces many challenges in practical applications [6]. Photovoltaic power generation systems' efficiency is significantly influenced by environmental conditions, such as lighting intensity, temperature changes and weather shifts, which can lead to instability in power generation [7]. Secondly, although photovoltaic cells' conversion efficiency has seen significant gains over time, there is still significant room for improvement [8], [9]. Again, existing control strategies and management systems often exhibit certain limitations when dealing with complex and changing environmental conditions [10].

Solving these problems and improve the overall performance of photovoltaic power generation systems, significant development has been made based on artificial intelligence control of photovoltaic power generation systems [11]. However, there are still many shortcomings in existing technologies in practical applications, which limit the overall performance and application effectiveness of photovoltaic power generation systems [12], [13]. Environmental conditions play a substantial role in influencing the energy production efficiency of photovoltaic

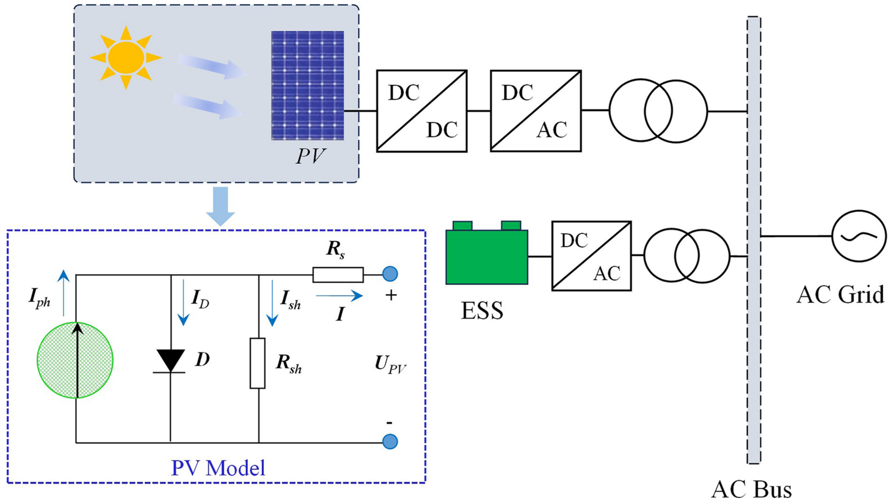

Figure 1. Photovoltaic power generation system model.

power systems, especially changes in light intensity, temperature, and weather [14]. Existing photovoltaic cells exhibit different conversion efficiencies under different lighting conditions, leading to instability in the system's output of electrical energy. Cloudy or windy conditions may adversely impact photovoltaic cells' energy production efficiency and temperature changes may also interfere with their performance, as excessively high or low temperatures can reduce their conversion efficiency [15], [16]. In addition, factors, such as fouling and obstruction of photovoltaic cells can also lead to fluctuations in power generation [17].

At present, most photovoltaic power generation systems adopt traditional PID control or simple model predictive control (MPC). These control strategies, although capable of regulating the system to a certain extent, exhibit significant limitations when facing complex and changing environmental conditions [18]. Traditional control methods are often based on fixed control parameters, which are difficult to adjust in real-time to cope with dynamic changes in the environment. In addition, these methods often rely on preset models, which cannot effectively address various nonlinear and uncertain problems that arise in practical applications, resulting in slow system response speed and inability to quickly adapt to environmental changes [19], [20].

In summary, existing photovoltaic power generation technologies still have many shortcomings in terms of energy production efficiency, control strategies, adaptive capabilities, multi-objective optimisation, and grid connected operation. This article proposes an efficiency optimisation method for photovoltaic power generation systems based on reinforcement learning and adaptive MPC, based on an in-depth analysis of the shortcomings of existing technologies. The main contribution of the proposed method can be summarised as follows.

1) The method proposed in this article achieves dynamic adaptive control of photovoltaic power generation systems by introducing reinforcement learning and adaptive MPC, significantly improving the system's adaptability and robustness under different environmental conditions.
2) The photovoltaic power generation optimisation system proposed in this article achieves efficient and stable operation of the system through the combination of multiple advanced control algorithms.
3) The photovoltaic power generation optimisation system proposed in this article significantly improves the overall performance of the system through advanced control algorithms and optimisation techniques. The system performs excellently in multiple aspects, such as power generation efficiency, stability, response speed, and equipment lifespan.

2. Modelling of Photovoltaic Power Generation System

As illustrated in Fig. 1, the photovoltaic power generation system investigated in this article comprises various components including its model and photovoltaic cells models, inverter model, and energy storage model. The detailed descriptions of each component are as follows:

1) Photovoltaic Cell Model

Photovoltaic cells form the cornerstone of photovoltaic power generation systems and their mathematical models can be described by following formula:

$$I = I_{\text{ph}} - I_0 \left(\exp\left(\frac{V + IR_s}{nV_t} \right) - 1 \right), \tag{1}$$

where I is the output current of the cell; I_{ph} is the photovoltaic current, which is usually proportional to the solar irradiance; I_0 is the reverse saturation current; The output voltage of V photovoltaic cells; R_s is a series resistor; n is the ideal factor of the diode; V_t is the thermal voltage, defined as $V_t = kT/q$, where k is the Boltzmann constant, T is the absolute temperature, and q is the charge.

$\mathrm{P_{pv}(t)}$ of the system can be represented using mathematical formula:

$$P_{\text{pv}}(t) = \eta_{\text{pv}} \cdot A \cdot G(t) \cdot \eta_{\text{inv}}, \tag{2}$$

where η_{pv} is the efficiency of the photovoltaic cell, A is the area of the photovoltaic cell, $G(t)$ is the solar irradiance at time t, and η_{inv} is the efficiency of the inverter.

2) Inverter Model

An inverter's primary function is to convert direct current (DC) generated from photovoltaic cells into alternating current (AC), for use either by power grid operators or end users. The efficiency of the inverter η_{inv} depends on its design and operating conditions. Pinv can be expressed in terms of output power for inverter applications:

$$P_{\text{inv}} = \eta_{\text{inv}} \cdot P_{\text{dc}} \tag{3}$$

where P_{dc} is the dc input power.

The input–output relationship of inverters usually adopts the following simplified model:

$$V_{\text{ac}} = k_{\text{inv}} \cdot V_{\text{dc}} \tag{4}$$

$$I_{\text{ac}} = \frac{I_{\text{dc}}}{k_{\text{inv}}}, \tag{5}$$

where V_{ac} and I_{ac} are ac voltage and current, V_{dc} and I_{dc} are dc voltage and current, respectively, and k_{inv} is the voltage gain of the inverter.

3) Energy Storage Device Model

Energy storage devices are used to store excess electricity from photovoltaic power generation systems during peak periods and provide power support when there is insufficient sunlight [18]. The mathematical model of an energy storage device can be expressed as:

$$\begin{aligned} E_{\text{stored}}(t) = {} & E_{\text{stored}}(t-1) + \eta_{\text{charge}} \cdot P_{\text{charge}}(t) \\ & - \frac{P_{\text{discharge}}(t)}{\eta_{\text{discharge}}}, \end{aligned} \tag{6}$$

where $E_{\text{stored}}(t)$ stands for the total stored energy at time t; η_{charge} and $\eta_{\text{discharge}}$ are charging efficiency and discharging efficiency metrics respectively for energy storage device; P_{charge} (t) denotes charging power at that momentous moment and $P_{\text{discharge}}$ (t) is discharging power at that same point in time.

The state change equation of the energy storage device can be expressed as:

$$\begin{aligned} \text{SOC}(t) = {} & \text{SOC}(t-1) \\ & + \frac{\eta_{\text{charge}} \cdot P_{\text{charge}}(t) - P_{\text{discharge}}(t)/\eta_{\text{discharge}}}{E_{\max}}, \end{aligned} \tag{7}$$

where $\text{SOC}(t)$ is the state of charge, and E_{max} is the maximum capacity of the energy storage device.

The photovoltaic power generation system based on reinforcement learning and adaptive MPC proposed in this paper achieves dynamic adaptive control of the photovoltaic power generation system by introducing advanced control algorithms and optimisation techniques [16]. The system architecture includes the following key components:

1) Enhanced Learning Module

This module continuously learns and optimises the control strategy of photovoltaic power generation systems using reinforcement learning algorithms such as Q-learning or deep Q-networks, with the objective being to maximize long-term cumulative returns using (8) as its criteria for optimisation,

$$\pi^{=} \arg\max_{\pi} \mathbb{E}\left[\sum_{t=0}^{\infty} \gamma^t r_t \mid \pi\right], \tag{8}$$

where π is the strategy, r_t is the reward for the time step, $\gamma \in [0, 1]$ is the discount factor.

2) MPC Module

The MPC module is based on an enhanced learning optimised control strategy to perform rolling optimisation of control inputs for a period of time in the future [19]. MPC ensures optimal performance of the system during each control cycle by adjusting the control input in real-time, as shown in (9),

$$\min_{u(t),\ldots,u(t+N-1)} \sum_{k=t}^{t+N-1} \left(\| y(k) - y_{\text{ref}}(k) \|_Q^2 + \| u(k) \|_R^2 \right). \tag{9}$$

3) Adaptive Control Module

The adaptive control module dynamically adjusts the control parameters of MPC by monitoring environmental parameters and system status in real-time to cope with dynamic changes in environmental conditions. The adaptive control mechanism ensures that the system maintains optimal performance under various operating conditions, and its control process can be expressed as:

$$\theta(t) = \theta(t-1) + \alpha(y(t) - y_{\text{ref}}(t)), \tag{10}$$

where $\theta(t)$ is the control parameter, α is the adjustment coefficient, $y(t)$ is the current system output, and $y_{\text{ref}}(t)$ is the current reference output.

4) Multi-objective Optimisation Module

This module utilises a multi-objective optimisation algorithm which takes into account multiple performance indicators like energy production efficiency, system stability, response speed, and equipment lifespan. Multi objective optimisation problems are usually expressed as Pareto optimal solutions, which means that there are no other solutions that can improve all objectives without sacrificing one objective, as shown in (11),

$$\min_{x \in \mathcal{X}} F(x) = [f_1(x), f_2(x), \ldots, f_m(x)]^T. \tag{11}$$

The architecture design of the entire system is shown in Fig. 2. Through the above system architecture design, dynamic adaptive control of the photovoltaic power generation system has been achieved, significantly improving the system's energy production efficiency, stability, and overall performance.

3. Efficiency Optimisation Algorithm Based on Reinforcement Learning and Adaptive MPC

3.1 Enhancement Learning Algorithm

Reinforcement learning is a machine learning method based on experimentation and feedback, which learns

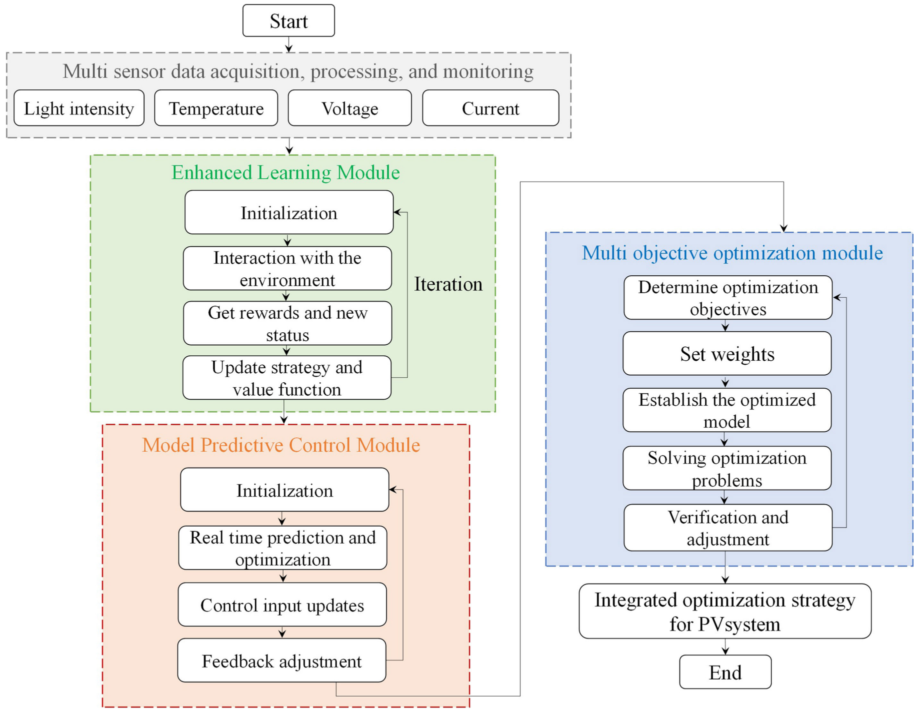

Figure 2. Iteration steps with reinforcement learning and adaptive model predictive control algorithm.

optimal control strategies through interaction with the environment. The basic principle is to optimise the control strategy by maximising cumulative rewards, enabling the system to adaptively adjust under different environmental conditions and achieve efficient and stable operation. The core of reinforcement learning algorithms includes elements, such as states, actions, rewards, and strategies. Specifically, the reinforcement learning algorithm can be described as follows.

1) State Space:

The state space represents the set of states of a system at a certain moment, denoted as S. In photovoltaic power generation systems, the state space can include environmental and system parameters, such as light intensity, temperature, voltage, and current.

2) Action Space:

The action space represents the set of actions that a system can take in a certain state, denoted as A. In photovoltaic power generation systems, actions can include adjusting the output voltage of the inverter, changing control parameters, and other operations.

3) Reward Function:

Reward functions $R(s,a)$ represent the immediate reward obtained by taking action a in state s. When designed correctly, this type of function should maximise energy production efficiency while simultaneously decreasing fluctuations of system operation. The reward function can be defined as:

$$R(s,a) = \alpha \cdot \eta - \beta \cdot \sigma, \tag{12}$$

where η is the energy production efficiency, σ is the system output, and α and β are weight parameters.

4) Strategy:

Reinforcement learning aims at finding an optimal strategy π^* which maximises cumulative rewards by looking for probability distributions of selecting action a in state s. The strategy $\pi(a|s)$ serves as the probability distribution for selecting action a.

5) Value Function:

The value function $V^\pi(s)$ represents cumulative reward expectations starting from state s under strategy π, starting in state s itself. It can also be represented mathematically as

$$V^\pi(s) = \mathbb{E}^\pi\left[\sum_{t=0}^{\infty}\gamma^t R(s_t,a_t) \mid s_0 = s\right], \tag{13}$$

where γ is the discount factor, representing the discount rate of future rewards.

6) Q-Value Function:

The Q-value function $Q^\pi(s,a)$ represents the expected cumulative reward after taking action a in state s under strategy π. The Q-value function is defined as:

$$Q^\pi(s,a) = \mathbb{E}^\pi\left[\sum_{t=0}^{\infty}\gamma^t R(s_t,a_t) \mid s_0 = s, a_0 = a\right]. \tag{14}$$

The optimal strategy π^* satisfies:

$$\pi^*(s) = \arg\max_a Q^\pi(s,a). \tag{15}$$

3.2 MPC

MPC is an advanced control strategy which optimises control inputs over time to achieve optimal system performance in each control cycle. Here, the MPC module achieves efficient and stable operation of photovoltaic power generation systems by real-time optimising of control strategies using MPC as its foundational control strategy.

1) Establish the System Model:

Firstly, assuming that the state vector of the photovoltaic power generation system is $\mathbf{x}(t)$, the control input vector is $\mathbf{u}(t)$, and the output vector is $\mathbf{y}(t)$. The state space model can be represented as:

$$x(t+1) = f(x(t), u(t)) + w(t) \tag{16}$$

$$y(t) = g(x(t), u(t)) + v(t), \tag{17}$$

where f and g represent state transition function and output function respectively and $\mathbf{w}(t)$ and $\mathbf{v}(t)$ represent process noise and measurement noise, respectively

2) Predicting Output:

In each control cycle, based on the current state $\mathbf{x}(t)$, the system model is used to predict the future N-step system output $y(t{+}k)$, where k=1,2,...,N is the prediction step size.

3) Define Optimisation Objectives:

Goal of optimisation: to minimise deviation between predicted output and reference trajectory, while minimising the change in control input. The optimisation objective can be expressed as the following cost function:

$$J = \sum_{k=0}^{N-1} \left(\| y(t+k|t) - y_{\mathrm{ref}}(t+k) \|_Q^2 + \| u(t+k) \|_R^2 \right), \tag{18}$$

where $y(t{+}k|t)$ is the output of the future k time predicted at time t, $y_{\mathrm{ref}}(t{+}k)$ is the reference trajectory, and are the weighted quadratic norms, respectively.

4) Solving Optimisation Problems:

Utilising quadratic programming methods to solve optimisation problems effectively, obtain the optimal control input sequence $u^*(t), u^*(t+1), \ldots, u^*(t+N-1)$ for the next N steps

5) Apply Optimal Control Input:

In practical applications, only the first optimal control input $\mathbf{u}^*(t)$ is executed, and then the next control cycle is entered, repeating the above steps.

3.3 Adaptive Control Mechanism

The adaptive control dynamically adjusts control parameters by monitoring environmental parameters and system status in real-time to ensure optimal performance of the system under various operating conditions. This mechanism combines reinforcement learning and MPC to form a closed-loop adaptive control system. Below are the steps involved with implementation:

1) Real-time Monitoring of System Status and Environmental Parameters

Real time collection of status parameters (such as voltage, current, power, *etc.*) and environmental parameters (such as light intensity, temperature, *etc.*) of photovoltaic power generation systems through data acquisition and monitoring modules. These data are used to evaluate the current operating status of the system and environmental changes.

2) Adaptive Adjustment of MPC Parameters

Based on real-time monitoring data, the adaptive control mechanism dynamically adjusts the parameters of MPC to cope with changes in environmental conditions. The core of adaptive control lies in parameter adjustment strategies, usually using gradient descent or other optimisation methods to update control parameters.

3) Parameter Adjustment Strategy

Assuming the control parameter vector is $\theta(t)$, the control parameters are adjusted by the error between the real-time monitoring data $\boldsymbol{y}(t)$ and the reference output $\mathbf{y}_{\mathrm{ref}}(t)$. The adjustment strategy can be expressed as:

$$\theta(t+1) = \theta(t) + \alpha \cdot \nabla_\theta J(\theta, y(t), y_{\mathrm{ref}}(t)), \tag{14}$$

where α represents the learning rate, and $\nabla_\theta J$ denotes the gradient of the loss function J concerning the parameter θ. The loss function J is typically characterised as the mean squared error between the predicted output and the target output:

$$J(\theta, y(t), y_{\mathrm{ref}}(t)) = \frac{1}{2}\sum_{i=1}^{n}(y_i(t) - y_{\mathrm{ref},i}(t))^2, \tag{15}$$

where n is the dimension of the output vector, $y_i(t)$ and $y_{\mathrm{ref},i}(t)$ are the ith output and the reference output, respectively.

4) Real-time Feedback Control

The adaptive control mechanism utilises updated control parameters to adjust the system's control inputs in real-time, making the system output as close as possible to the reference output. The modification procedure of the control input can be articulated as:

$$u(t) = f(x(t), \theta(t)), \tag{16}$$

where $\mathbf{u}(t)$ is the control input, $\mathbf{x}(t)$ is the system state, and f is the control strategy function.

5) Closed-Loop Control System

The combination of adaptive control mechanism and MPC forms a closed-loop control system. Real time monitoring data is used to update control parameters, optimise MPC control strategies, enable the system to quickly respond to environmental changes, and maintain efficient and stable operation.

3.4 Multi-objective Optimisation Algorithm

In photovoltaic power generation systems, multi-objective optimisation algorithms are used to comprehensively consider multiple performance indicators, such as energy production efficiency, system stability, response speed, and equipment life, in order to achieve overall performance optimisation of the system. Multi objective optimisation problems typically involve multiple conflicting objectives and require finding a balance point between different objectives. To achieve this goal, an adaptive multi-objective optimisation algorithm utilising Pareto optimal solutions was devised.

1) Problem Description

By allocating varying weights to each objective function and merging all these functions into a unified comprehensive objective function, a multi-objective optimisation problem can be articulated as follows.

$$\min_{x\in\mathcal{X}} J(x) = w_1 f_1(x) + w_2 f_2(x) + w_3 f_3(x) + w_4 f_4(x) \tag{17}$$

$$J(x) = \sum_{i=1}^{m} w_i f_i(x). \tag{18}$$

Among them, $\mathbf{x}$ is the decision variable, $\mathcal{X}$is the feasible solution space, $\mathbf{F}(\mathbf{x})$ is the objective function vector, $f_i(\mathbf{x})$ is the ith objective function, m is the number of objective functions, w_i is the weight of the ith objective function, and $\sum_{i=1}^{m} w_i = 1$.

2) Pareto Optimal Solution

The Pareto optimal solution is a solution that cannot further improve a certain objective without compromising other objectives. Mathematically, the solution $\mathbf{x}^*$ is a Pareto optimal solution, If and only if there is no other solution $x \in \mathcal{X}$, such that:

$$\begin{cases} f_i(x) \leq f_i(x^*), \forall i = 1, 2, \ldots, m \\ f_j(x) < f_j(x^*), \text{at least } j \in \{1, 2, \ldots, m\} \end{cases}. \tag{19}$$

Based on the above analysis, Fig. 2 shows the processing and optimisation iteration steps of the reinforcement learning and adaptive MPC algorithm. Through these steps, the proposed algorithm can achieve overall performance optimisation of photovoltaic power generation systems by finding an ideal equilibrium point between various optimisation objectives, ultimately optimising overall power production efficiency.

4. Results and Discussion

The testing was conducted on MATLAB/Simulink platform and actual experimental platform. The experimental model includes main components, such as photovoltaic cells, inverters, energy storage devices, and control systems, which can truly reflect the operating characteristics and control behaviour of solar energy generation systems.

The experimental platform includes high-efficiency monocrystalline silicon photovoltaic modules, high-efficiency inverters, lithium battery energy storage systems, and environmental sensors (used to monitor light intensity, temperature, voltage, current, power, etc.). Real time collection of environmental parameters and system status data for data processing and monitoring. The data collection frequency is set to once every 0.1 s to ensure real-time and accurate data.

4.1 Parameter Settings

The parameter settings cover both experimental model parameters and algorithm parameters, ensuring the rationality and reliability of the test results.

The experimental model parameter settings are shown in Table 1.

The algorithm parameter settings are shown in Table 2.

Table 1
Experimental Model Parameter Settings

Module	Parameter Name	Parameter Values
PV	Types of photovoltaic cells	Single-crystal silicon
	Rated power	250 W
	Efficiency	20%
	Total area of array	20 m^2
Inverter	Rated power	5 kW
	Efficiency	95%
	Input voltage range	300–600 V
	AC output voltage	230 V
ESS	Type of energy storage	Lithium battery
	Total capacity	10 kWh
	Efficiency	90%
	Range of power	0–5 kW

The hardware module parameter settings for experimental parameters refer to Table 1, and other parameter settings are shown in Table 3.

4.2 Experimental Result Analysis

Figure 3 depicts the variation of energy production efficiency over time of photovoltaic power generation systems with various control strategies applied. The trend of energy production efficiency over time clearly reflects the variation of light intensity within 24 h of a day. The energy production efficiency under the four control strategies showed similar trends, increasing with increasing light intensity and decreasing with decreasing light intensity. At night (when the light intensity was zero), the energy production efficiency was also zero. The energy production efficiency under PID control strategy is relatively low, with an average energy production efficiency of about 0.29 and a peak energy production efficiency of about 0.79. Although PID control is simple, its adaptability to environmental changes is poor, which has resulted in considerable fluctuations in energy production efficiency; the implementation of MPC has significantly enhanced this efficiency, with an average energy production efficiency of about 0.31 and a peak energy production efficiency of about 0.84. MPC control has improved the stability and energy production efficiency of the system to a certain extent. The energy production efficiency is further improved under the adaptive MPC strategy, with an average energy production efficiency of about 0.33 and a peak energy production efficiency of about 0.89. Adaptive MPC can dynamically adjust control parameters according to environmental changes, considerably enhancing the efficiency and stability of the power generation system. The method described in this article corresponds to the highest energy production

Table 2
Algorithm Parameter Settings

Module	Parameter Name	Parameter Values
RL	Algorithm type	Deep Q Network
	Reward function weight	1,0.01
	Learning rate	0.001
	Discount factor	0.9
	Experience replay buffer size	10000
	Target network update frequency	1000
MPC	Predicting time domain	10 step
	Control time domain	5 step
Adaptive control	Initial control parameters	[1,0.5]
	Learning rate	0.01
	Loss function	mean squared error

Table 3
Experimental Parameter Settings

Module	Parameter Name	Parameter Values
Sensors	Light intensity sensor	0–1200 W/m^2
	Temperature sensor	−20 to 50°C
	Voltage sensor	0–600 V
	Current sensor	0–50 A

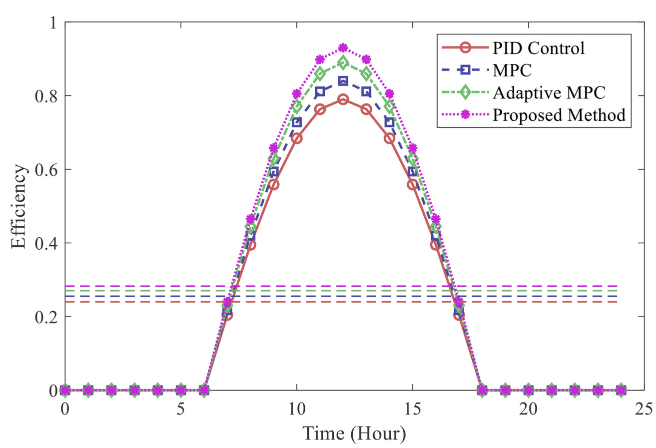

Figure 3. Photovoltaic energy production efficiency under different strategies.

efficiency, with an average energy production efficiency of about 0.35 and a peak energy production efficiency of about 0.93. This method combines reinforcement learning and adaptive MPC to effectively learn and optimise control strategies, achieving dynamic adaptive control of photovoltaic power generation systems and significantly improving the overall performance.

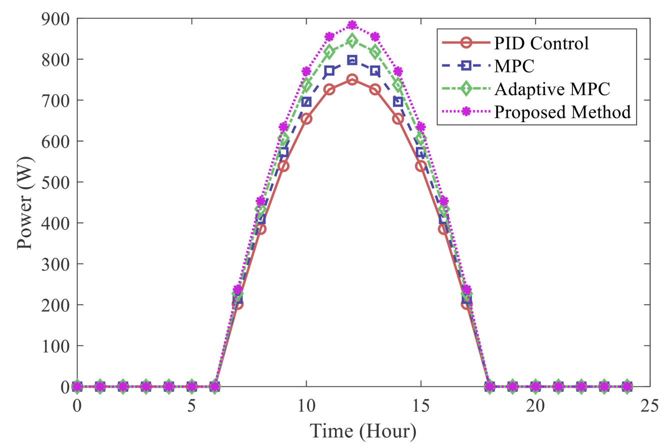

Figure 4. Photovoltaic output power under different strategies.

Figure 4 shows the variation of output power of photovoltaic power generation system under different control strategies under the conditions of light intensity and temperature changes. As light intensity increases, so too does output power; On the contrary, as the intensity of light decreases, the output power gradually decreases. In addition, temperature changes also have an impact on output power. The temperature gradually increases from morning to noon and gradually decreases in the evening. Due to the decrease in energy production efficiency caused by temperature rise, the output power slightly decreases at the highest temperature. Although PID control is simple, its adaptability to environmental changes is poor, resulting in significant fluctuations in output power. The output power under MPC strategy has been improved with less fluctuation. Average output power levels remain relatively high, with a peak output power of about 840 W. MPC control has enhanced the system's stability and output power to some degree. By combining reinforcement

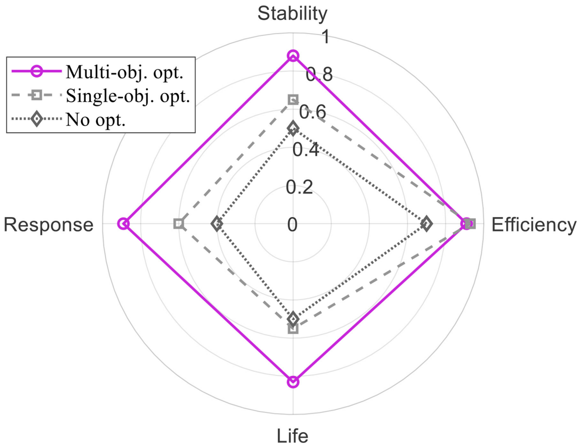

Figure 5. Analysis of multi-objective and single objective optimisation effects.

learning and adaptive MPC, dynamic adaptive control of photovoltaic power generation systems has been achieved, significantly improving the output power and stability of the system.

Figure 5 illustrates the comparison of different performance indicators for the method proposed in this paper across three scenarios: without optimisation, with single-objective optimisation, and with multi-objective optimisation, reflecting the impact of different optimisation strategies on four key performance indicators: energy production efficiency, system stability, response speed, and equipment life. In terms of energy production efficiency, the single objective optimisation strategy has the highest energy production efficiency, reaching 0.93. This is because the single objective optimisation strategy only optimises energy production efficiency and ignores the optimisation of other performance indicators. However, the drawbacks of this single optimisation method are also very obvious, that is, the system stability, response speed, and equipment lifespan have not been effectively improved. The multi-objective optimisation strategy performs well in balancing various performance indicators. While the energy production efficiency following multi-objective optimisation is marginally lower than that of single-objective optimisation, the system stability, response speed, and equipment lifespan are significantly improved compared to non-optimisation and single objective optimisation. This indicates that multi-objective optimisation strategies can significantly improve overall performance without significantly reducing energy production efficiency. In addition, all performance indicators of the non-optimised strategy are at a relatively low level. This indicates that without optimisation measures, the system performance is poor and difficult to meet the needs of practical applications.

Figure 6 presents a comparison and analysis of the average energy production efficiency and average output power across various weather conditions. The results indicate that, regardless of the weather, the average energy production efficiency and average output power of the proposed method are significantly greater than those of the other three control strategies. This demonstrates the effectiveness and advantages of the proposed method in optimising photovoltaic power generation systems, confirming its adaptability and robustness in dynamic environments, but also provides its broad prospects in practical applications. This method combines reinforcement learning and adaptive MPC to effectively optimise the control strategy of photovoltaic power generation systems, achieving efficient, stable, and reliable power generation performance. It has important application value and promotion potential.

5. Conclusion

This article details a system designed to optimise photovoltaic power generation using reinforcement learning and adaptive MPC techniques for dynamic adaptive management of photovoltaic production systems. To improve adaptability and robustness under various environmental conditions, the system employs reinforcement learning

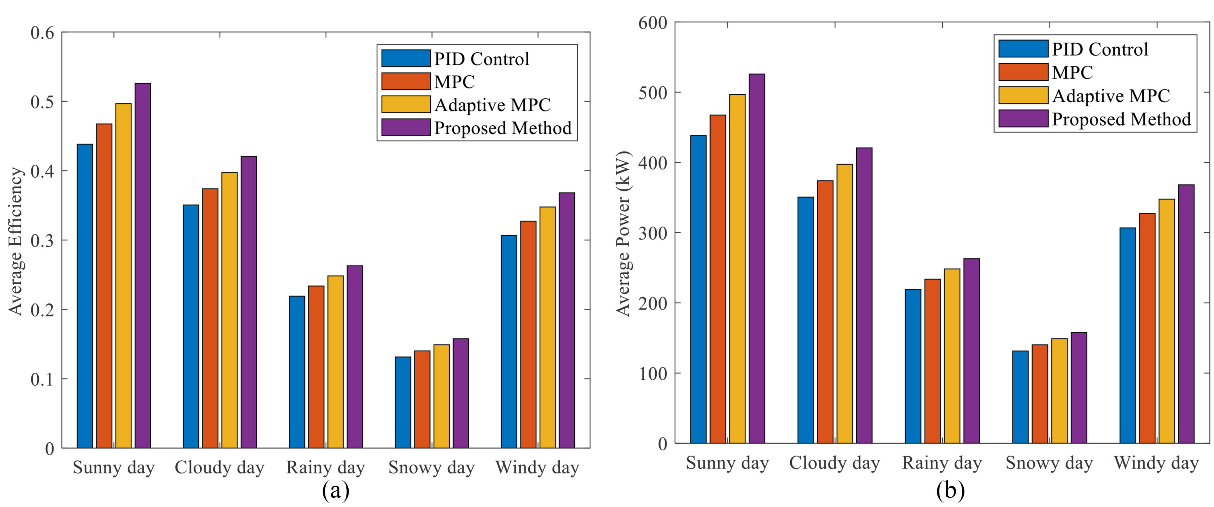

Figure 6. Comparison of average efficiency and output power under different weather conditions from 6 to 18 h: (a) Comparison of average efficiency and (b) Comparison of average output power.

algorithms to discover an optimal control strategy through interaction with its environment. The MPC module then utilises this optimised strategy to forecast future control inputs, ensuring efficient and stable operation of the photovoltaic power generation systems.

Additionally, adaptive control mechanisms dynamically adjust control parameters in real-time by monitoring environmental variables and system status, maintaining peak performance across all operating conditions. This system employs multi-objective optimisation algorithms to enhance energy production efficiency while also considering system stability, response speed, and lifespan requirements, ultimately improving the overall performance of the photovoltaic power generation system.

Finally, the experimental results demonstrate that the proposed optimisation system significantly enhances accuracy, control efficiency, and stability/reliability in complex environments, showcasing its broad application potential.

Future work could aim to reduce the computational complexity of the multi-objective optimisation algorithms. This reduction would lead to faster decision-making processes, making the system more responsive to rapid environmental changes and thus improving its practical application value.

References

[1] D. Zhao, X. Yang, and H. Wang, Recent advances in perovskite solar cells with high efficiency and stability, *IEEE Journal of Photovoltaics*, *14*(1), 2024. 309–326.

[2] Y. Li, X. Zhang, and L. Wang, Interface engineering in perovskite solar cells for improved performance, *IEEE Transactions on Electron Devices*, *70*(10), 2023, 4451–4460.

[3] Q. Wang, X. Liu, and X. Li, Understanding the degradation mechanisms of perovskite solar cells, *IEEE Journal of Photovoltaics*, *13*(5), 2023, 1245–1254.

[4] S .H. Abdallah and S. Tounsi, Optimal design and control of ac-dc-dc wind energy system for electric vehicle batteries recharging stations, *International Journal of Power and Energy Systems*, *44*(396), 2024, 1–9.

[5] X. Yang, W. Zhang, and W. Zou, Design of power tracking model for photovoltaic power generation based on improved quantum swarm algorithm and conductance incremental approach, *International Journal of Power and Energy Systems*, *44*(532), 2024, 1–10.

[6] Y. Zhao, X Sun, and X. Li, Optimization of the rear surface field in silicon solar cells, *IEEE Transactions on Electron Devices*, *70*(8), 2023, 3345–3352.

[7] X. Huang, Y. Wang, and J. Zhang, A novel maximum power point tracking algorithm for photovoltaic systems, *IEEE Transactions on Power Electronics*, *39*(2), 2024, 1567–1576.

[8] H. Li, C. Wang, and Y. Liu, Performance analysis and optimization of photovoltaic arrays under non-uniform irradiation, *IEEE Journal of Photovoltaics*, *13*(6), 2023, 1512–1520.

[9] L. Zhang, Q. Zhao, and H. Yang, Hybrid photovoltaic/thermal systems for enhanced energy conversion efficiency, *IEEE Transactions on Energy Conversion*, *38*(3), 2023, 1205–1213.

[10] X. Wang, Y. Zhang, and M. Li, Reliability analysis of photovoltaic modules under harsh environmental conditions, *IEEE Transactions on Device and Materials Reliability*, *24*(1), 2024, 34–43.

[11] W. Liu, Z. Chen, and J. Wang, Degradation mechanisms and lifetime prediction of photovoltaic systems, *IEEE Journal of Photovoltaics*, *13*(3), 2023, 750–759.

[12] J. Zhao, W. Sun, and Q. Li, Fault diagnosis and tolerant control for photovoltaic systems, *IEEE Transactions on Power Electronics*, *38*(11), 2023, 11057–11066.

[13] X. Li, Y. Wang, and Q. Zhang, Optimal design and control of photovoltaic-energy storage systems for grid integration, *IEEE Transactions on Smart Grid*, *15*(1), 2024, 456–465.

[14] L. Yang, X. Liu, and H. Wang, A novel energy management strategy for photovoltaic-battery systems, *IEEE Journal of Photovoltaics*, *13*(2), 2023, 502–510.

[15] Z. Wang, X. Zhang, and J. Li, Modeling and experimental of photovoltaic-energy storage hybrid systems, *IEEE Transactions on Power Systems*, *38*(4), 2023, 3015–3024.

[16] M. Chen, X. Li, and X. Wang, Novel materials for high-efficiency photovoltaic devices, *IEEE Journal of Photovoltaics*, *14*(2), 2024, 501–512.

[17] Y. Zhang, Y. Liu, and Y. Wang, Characterization and optimization of photovoltaic materials, *IEEE Transactions on Electron Devices*, *70*(9), 2023, 3785–3794.

[18] J. Wang, X. Zhao, and X. Li, Research progress in organic photovoltaic materials and devices, *IEEE Journal of Photovoltaics*, *13*(1), 2023, 201–210.

[19] Y. Huang, X. Wang, and L. Zhang, Intelligent monitoring system for photovoltaic power plants based on big data and machine learning, *IEEE Transactions on Industry Applications*, *50*(1), 2024, 321–330.

[20] M. Li, C. Wang, and W. Liu, Model predictive control of photovoltaic systems for maximum power point tracking, *IEEE Transactions on Power Electronics*, *38*(9), 2023, 9567–9576.

Biographies

Zhichao Zhang was born in 2000 in Taixing, Jiangsu, China. He received the master's degree in electrical power systems from the University of Birmingham in December 2023. He is currently pursuing the Doctoral degree with the School of Electrical Engineering, China University of Mining and Technology. His research interests include new energy power generation technology and power electronics technology.

Qihang Liu was born in 1999 in Lianyungang, Jiangsu, China. He received the masters degree in electrical power systems from the University of Birmingham in December 2023. Since July 2024, he has been conducting integrated technology research with NARI Group Corporation. His research interests include demand-side management, microgrids, and virtual power plants.

Chuyuan Wang was born in 1999 in Hefei, Anhui, China. She received the master's degree in electrical power systems from the University of Birmingham in December 2023. Since July 2024, she has been a Technical Engineer with the Dispatching Control Center, Feixi Power Supply Company of State Grid. Her research interests include design and management methods for photovoltaic DC series-connected grid.

Yujie Zong was born in 2000 in Taixing, Jiangsu, China. He received the masters degree in civil engineering from the University of Birmingham in December 2023. Since July 2024, he has been a Technical Engineer with the Eighth Engineering Bureau of China Construction. His research interests include the field of new intelligent construction based on deep learning and industrial intelligence.