(Received 11 April 2025, Accepted 6 June 2025, Scheduled 2 July 2025)

Intelligent MPPT Framework with Reinforcement Learning and Dynamic Search Region Optimization for Photovoltaic Systems Under Variable Environmental Conditions

Xiaoping Lei*

School of Electronics and Internet of Things, Chongqing Polytechnic University of Electronic Technology Chongqing 401331, China

ABSTRACT: This paper introduces an intelligent Maximum Power Point Tracking (MPPT) framework for photovoltaic systems that achieves significant performance gains through two primary innovations: a dynamic search space optimization that intelligently constrains the search region to approximately 2% of the conventional area and a sophisticated Q-learning algorithm operating within this optimized region. The framework establishes a real-time relationship between environmental conditions and maximum power point parameters for this aggressive search space reduction. For complex partial shading conditions, an adaptive switching mechanism dynamically activates an enhanced meta-heuristic optimization component with improved convergence properties, ensuring appropriate algorithm selection based on detected operating conditions. Experimental results demonstrate that under uniform irradiance, the framework achieves 99.12% tracking efficiency (a 3.34% improvement over P&O). Under rapidly changing conditions, it maintains 97.83% efficiency (compared to P&O's 90.12%), and under partial shading, it achieves 95.89% global MPPT efficiency (versus 76.25% for P&O). The proposed method significantly reduces steady-state oscillations to 0.41% (from 1.87% for P&O) and offers 42.3% faster convergence. While requiring moderately higher computational resources, the approach is implementable on medium-range microcontrollers, balancing performance with practical deployment.

1. INTRODUCTION

With the increasing severity of global energy crisis and environmental pollution issues, renewable energy, especially solar power generation systems, has received widespread attention [1, 2]. However, photovoltaic (PV) systems suffer from low energy conversion efficiency and output power highly dependent on environmental conditions [3]. Maximum Power Point Tracking (MPPT) technology is crucial for maximizing energy output by ensuring PV systems operate at their Maximum Power Point (MPP) under varying conditions [4]. While traditional MPPT algorithms like Perturb and Observe (P&O) and Incremental Conductance (INC) are widely used, they exhibit poor performance under rapidly changing environmental conditions and Partial Shading Conditions (PSC), often leading to power losses, slow tracking, and steady-state oscillations [5, 6]. Recent advancements in artificial intelligence, particularly Reinforcement Learning (RL) combined with search space optimization, show significant promise in improving tracking accuracy, reducing search time, and enhancing system robustness, especially under PSC [7, 8].

The primary limitations of traditional MPPT algorithms stem from their large search space, leading to slow convergence and poor adaptability [9]. While methods like fuzzy logic control (FLC) and artificial neural networks (ANN) have been proposed to narrow the search range [12, 13], they often require costly sensors or complex designs. RL offers an alternative by enabling autonomous learning of optimal strategies [14], dynamically adjusting search based on system states and environmental changes [15]. However, RL-MPPT methods can face challenges like complex training and instability under PSC [16, 17]. This paper proposes an adaptive hybrid MPPT method combining RL with search space optimization. It quickly estimates MPP voltage to establish effective search boundaries and uses an improved Q-learning algorithm for precise MPP location within this reduced space, aiming for more efficient and robust MPPT.

This paper introduces an adaptive hybrid MPPT method that integrates search space optimization with RL techniques. The proposed approach first establishes an empirical relationship between open circuit voltage and MPP voltage to quickly estimate the MPP position without expensive irradiance sensors. Based on this estimation, the system defines optimized search boundaries that significantly reduce the computational burden. Within this constrained region, an improved Q-learning algorithm precisely locates the MPP while dynamically adjusting to environmental changes. For PSC, an enhanced grey wolf optimization (EGWO) component is activated to identify the global MPP, with an adaptive switching mechanism selecting the appropriate algorithm based on real-time conditions.

2. RELATED WORKS

The development of renewable energy, particularly solar, is critical due to global environmental concerns. However, PV

* Corresponding author: Xiaoping Lei (xiaopingleilei2@163.com).

systems' efficiency is hampered by nonlinear output characteristics fluctuating with temperature and irradiance, often preventing operation at the MPP. MPPT algorithms are thus essential. While being diverse, existing MPPT technologies face challenges in accurately tracking the MPP, especially under dynamic environmental conditions.

Traditional methods like P&O, though simple, suffer from oscillations near the MPP [18]. Variable step size P&O [19] and beta MPPT [20] offered improvements but have limitations. Consequently, research has shifted towards optimizing MPPT by reducing the search space using intelligent algorithms like FLC [21] and ANN. However, these often increase complexity or require additional sensors [22, 25]. For instance, Mo et al. [22] used FLC with irradiance measurements, and Harrison et al. [23] combined INC with integral backstepping. While methods based on limited search spaces have been proposed [26, 27], Harrison et al.'s recent work on reducing MPP voltage search space showed improved efficiency over traditional P&O [14].

RL has emerged as a promising alternative, learning optimal control policies without explicit mathematical models [28]. Q-learning based MPPT has demonstrated superior performance over P&O in dynamic conditions [29], and Deep Q-Network (DQN) methods have achieved high tracking efficiency [30]. However, handling PSC with multiple local maxima remains a challenge for these RL methods [31].

Meta-heuristic optimization algorithms have been studied for PSC. Modified Particle Swarm Optimization (PSO) [32] and improved GWO [33] have shown better efficiency under partial shading. However, these methods often exhibit high computational complexity and slow convergence [34].

Hybrid approaches combining different algorithms have gained attention. ANN with PSO [35] and fuzzy logic with INC [36] have shown promise but still struggle with the trade-off between convergence speed and accuracy [37].

The concept of search space reduction as a preprocessing step for MPPT algorithms represents a significant advancement in the field. Basheer et al. [38] pioneered this approach by establishing a framework that effectively narrows the search region based on the relationship between open-circuit voltage and MPP voltage. Their work demonstrated that reducing the search space to approximately 2% of the original area could significantly improve convergence speed while maintaining high tracking accuracy. Building on this concept, Meineri et al. [39] integrated a simplified neural network model to predict the approximate MPP voltage range, further enhancing the efficiency of the search process.

Despite these advancements, existing approaches have not fully explored the integration of RL with search space optimization techniques, particularly for handling complex environmental conditions [40, 41]. Additionally, most current methods lack adaptive mechanisms to intelligently switch between different algorithms based on environmental conditions [42]. This gap presents an opportunity to develop a more comprehensive hybrid MPPT approach that combines the strengths of search space optimization, RL, and meta-heuristic algorithms while addressing their individual limitations.

3. METHODOLOGY

This section details the adaptive hybrid MPPT method based on RL and search space optimization. As depicted in Fig. 1, the proposed MPPT framework integrates four key components: predictive search space optimization, an RL controller, a meta-heuristic global search mechanism, and an adaptive switching strategy. The core innovation of this method lies in combining Rahul and Hariharan's space reduction framework with deep Q-learning and an improved GWO algorithm [43], forming a comprehensive solution that effectively handles MPPT problems under various environmental conditions. Under standard operating conditions, the system first predicts the approximate MPP voltage position using an improved neural network model and dynamically sets search boundaries based on prediction uncertainty, significantly reducing the search space. Then, the RL controller precisely locates the MPP within this reduced search space. When PSC is detected, the system automatically switches to the meta-heuristic global search mechanism to handle multi-peak Power-Voltage (P-V) curves. The entire process is monitored by the adaptive switching strategy module, which intelligently selects the most appropriate MPPT algorithm based on real-time environmental conditions and system state, thereby achieving optimal energy harvesting performance.

The system integrates four key components: predictive search space optimization using deep neural networks with Bayesian confidence intervals, RL controller based on DQN, metaheuristic global search mechanism using modified GWO for PSC, and adaptive switching strategy that selects the optimal algorithm based on environmental conditions. Solid arrows represent data flow, while dashed lines indicate decision paths.

3.1. Predictive Search Space Optimization

The predictive search space optimization component forms the foundation of our hybrid MPPT approach by significantly reducing the search area for the MPP. While Rahul and Hariharan [43] established that limiting the search space to a narrow region around the predicted MPP voltage can improve algorithm convergence, their method relied on simplified voltage relationships that lack robustness under complex environmental conditions. Our approach enhances this framework through several key innovations.

Firstly, we implement an improved deep neural network architecture to predict MPP voltage with higher accuracy across diverse operating conditions. The network incorporates multiple feature inputs including open-circuit voltage (Voc), temperature (T), historical power data (Pt-1, Pt-2), and voltage change rates ($\Delta V/\Delta t$). This multi-feature approach enables more precise MPP voltage prediction than methods relying solely on Voc-based empirical relationships. The neural network architecture employs residual connections to mitigate gradient vanishing problems during training, expressed as:

$$h_l = F_l(h_{l-1}) + h_{l-1} \quad (1)$$

$$V_m = G(h_L) \quad (2)$$

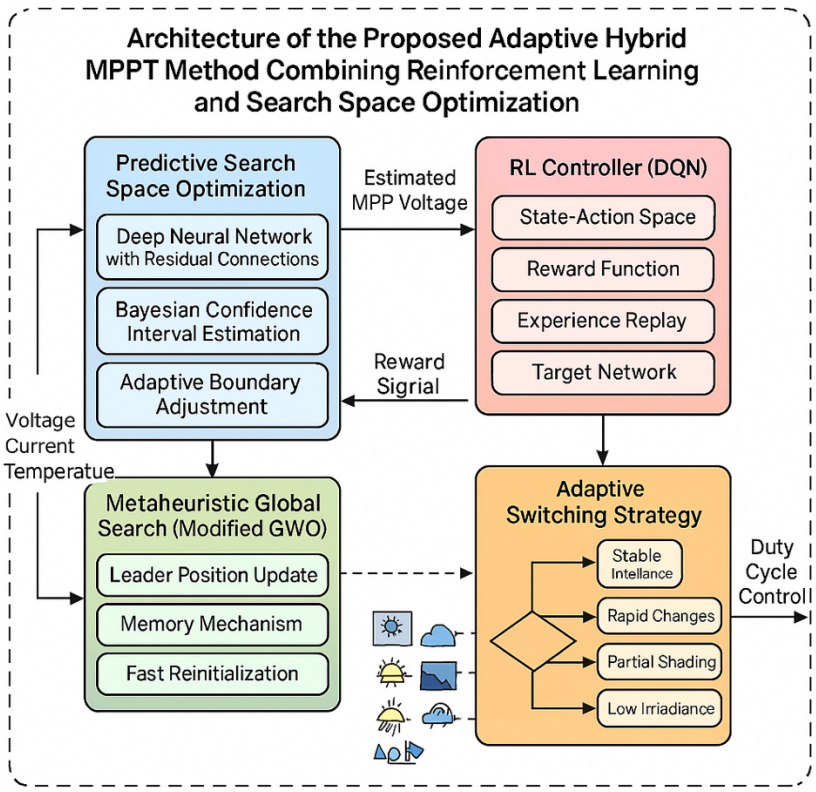

FIGURE 1. Schematic diagram of the proposed adaptive hybrid MPPT method.

where h_l represents the output of the l-th hidden layer, F_l the mapping function of that layer, and G the output layer function that produces the predicted MPP voltage V_m.

Secondly, we integrate Bayesian methods to quantify prediction uncertainty, providing not only point estimates but also confidence intervals for the predicted MPP voltage. This uncertainty quantification is critical for adaptive boundary determination. The predictive distribution of MPP voltage can be formulated as:

$$p(V_m|xD) = \int p(V_m|xw)p(w|D)dw \tag{3}$$

where x represents the input features, D the training data, w denotes the network parameters, and $p(w|D)$ the posterior distribution of network parameters. The mean μV_m provides the point prediction while the standard deviation σV_m quantifies prediction uncertainty.

Thirdly, we develop an adaptive boundary adjustment mechanism that dynamically modifies the search region based on predicted uncertainty. Unlike fixed-margin approaches, our method expands the search region under high uncertainty and narrows it when confidence is high, optimizing the trade-off between search speed and accuracy. The search boundaries V_a and V_b are determined by:

$$V_a = \mu_{V_m} - |\varepsilon_a| \cdot \alpha(\sigma_{V_m}) \tag{4}$$

$$V_b = \mu_{V_m} + |\varepsilon_b| \cdot \beta(\sigma_{V_m}) \tag{5}$$

where ε_a and ε_b are base margin parameters, while $\alpha(\sigma V_m)$ and $\beta(\sigma V_m)$ are adaptive coefficients that scale with prediction uncertainty. These coefficients follow sigmoid-based scaling functions:

$$\alpha(\sigma_{V_m}) = \alpha_{\min} + \frac{\alpha_{\max} - \alpha_{\min}}{1 + e^{-k(\sigma_{V_m} - \sigma_0)}} \tag{6}$$

$$\beta(\sigma_{V_m}) = \beta_{\min} + \frac{\beta_{\max} - \beta_{\min}}{1 + e^{-k(\sigma_{V_m} - \sigma_0)}} \tag{7}$$

where $\alpha_{\min}$, $\alpha_{\max}$, $\beta_{\min}$, and $\beta_{\max}$ define the range of scaling factors; k controls the steepness of the sigmoid function; and σ_0 is the reference uncertainty threshold.

Our predictive search space optimization implementation requires careful design of both the neural network architecture and training methodology to ensure robust performance across diverse environmental conditions. The deep neural network consists of four hidden layers with [64, 128] neurons respectively, activated by LeakyReLU functions to avoid dying Rectified Linear Unit (ReLU) problem while maintaining computational efficiency. The residual connections are implemented between alternate layers, allowing gradient flow during backpropagation even as the network depth increases.

The input feature vector is preprocessed through normalization to ensure that all features contribute proportionally to the prediction:

$$x_{norm} = \frac{x - \mu_x}{\sigma_x} \tag{8}$$

where μ_x and σ_x represent the mean and standard deviation of each feature calculated from the training dataset. This normalization step is crucial for handling features with different scales such as voltage (typically 0–50 V) and temperature (typically 20–80°C).

For training the neural network, we employ a composite loss function that combines mean squared error (MSE) for accurate point prediction and negative log-likelihood (NLL) for proper uncertainty quantification:

$$\mathcal{L}_{total} = \lambda_{MSE} \cdot \mathcal{L}_{MSE} + \lambda_{NLL} \cdot \mathcal{L}_{NLL} \tag{9}$$

$$\mathcal{L}_{MSE} = \frac{1}{N}\sum_{i=1}^{N}(V_{m,i} - \hat{V}_{m,i})^2 \tag{10}$$

$$\mathcal{L}_{NLL} = \frac{1}{N}\sum_{i=1}^{N}\left[\frac{(V_{m,i} - \hat{V}_{m,i})^2}{2\sigma_i^2} + \frac{1}{2}\log(\sigma_i^2)\right] \tag{11}$$

where $V_{m,i}$ represents the true MPP voltage, $\hat{V}_{m,i}$ the predicted value, σ_i^2 the predicted variance, and λ_{MSE} and λ_{NLL} are weighting hyperparameters that control the relative importance of accuracy versus uncertainty quantification.

The Bayesian framework is implemented through Monte Carlo dropout, where dropout layers with probability $p_{drop} = 0.2$ are inserted after each hidden layer and kept active during inference. This approach approximates Bayesian inference without the computational burden of full Bayesian neural networks. By performing $M = 30$ forward passes with different dropout masks, we generate a distribution of predictions from which we calculate the mean μ_{V_m} and standard deviation σV_m:

$$\mu_{V_m} = \frac{1}{M}\sum_{j=1}^{M}\hat{V}_m^{(j)} \tag{12}$$

$$\sigma_{V_m}^2 = \frac{1}{M}\sum_{j=1}^{M}(\hat{V}_m^{(j)} - \mu_{V_m})^2 + \frac{1}{M}\sum_{j=1}^{M}\hat{\sigma}_m^{2(j)} \tag{13}$$

where $\hat{V}_m^{(j)}$ and $\hat{\sigma}_m^{2(j)}$ are the predicted MPP voltage and variance from the j-th forward pass.

The adaptive boundary adjustment mechanism is designed to respond dynamically to both prediction uncertainty and historical tracking performance. In addition to the sigmoid-based scaling functions described earlier, we incorporate a feedback mechanism that adjusts the reference uncertainty threshold σ_0 based on recent tracking outcomes:

$$\sigma_0^{(t)} = \gamma \cdot \sigma_0^{(t-1)} + (1-\gamma) \cdot f(\Delta P^{(t-1)}) \tag{14}$$

where γ is a smoothing factor (typically 0.8), $\sigma_0^{(t)}$ the updated threshold at time step t, and $f(\Delta P^{(t-1)})$ a function that maps the previous power improvement $\Delta P^{(t-1)}$ to a threshold adjustment. This function is defined as:

$$\sigma_0 = \begin{cases} \sigma_{\text{base}} - \delta, & \text{if } \Delta P > \Delta P_{high} \\ \sigma_{\text{base}} + \delta, & \text{if } \Delta P < \Delta P_{low} \\ \sigma_{\text{base}}, & \text{otherwise} \end{cases} \tag{15}$$

where σ_{base} is the base threshold value, δ the adjustment magnitude, and ΔP_{high} and ΔP_{low} are thresholds defining good and poor tracking performance, respectively. The boundary parameters $\alpha_{\min}$, $\alpha_{\max}$, $\beta_{\min}$, and $\beta_{\max}$ are set asymmetrically to account for the typically asymmetric nature of P-V curves. The steepness parameter k ensures responsive adaptation to uncertainty changes without excessive boundary oscillations. This comprehensive predictive search space optimization approach provides a robust foundation for subsequent MPPT algorithms by intelligently constraining the search region, significantly reducing computational requirements while maintaining the ability to locate the true MPP across diverse operating conditions.

3.2. RL Controller

The RL controller constitutes the second key component of our hybrid MPPT framework, operating within the optimized search space defined by the predictive component. While conventional MPPT algorithms like P&O and INC rely on fixed step-size adjustments that create inherent trade-offs between tracking speed and steady-state oscillations, our RL-based approach dynamically adapts control actions through continuous interaction with the environment.

We formulate the MPPT problem as a Markov Decision Process (MDP) and employ DQN to learn optimal control policies. The MDP framework is defined by the state space, action space, reward function, and transition dynamics, expressed as the tuple (S, A, P, R, γ), where S is the state space, A the action space, P the state transition probabilities, R the reward function, and γ the discount factor. The state space S consists of a multidimensional representation of the system's current operating conditions:

$$S = [P_t V_t \Delta P_t \Delta V_t V_t / V_{oc}, (V_b - V_a)/V_{oc}] \tag{16}$$

where P_t and V_t are the current power and voltage; ΔP_t and ΔV_t are the changes in power and voltage since the previous step; V_t/V_{oc} represents the normalized voltage position; and $(V_b - V_a)/V_{oc}$ indicates the relative size of the search space. This comprehensive state representation enables the RL agent to make informed decisions based on both current measurements and their historical context.

The action space A consists of discrete duty cycle adjustment options:

$$\begin{aligned} A = \{&-\Delta D_{large}, -\Delta D_{medium}, -\Delta D_{small}, 0, \\ &+\Delta D_{small}, +\Delta D_{medium}, +\Delta D_{large}\} \end{aligned} \tag{17}$$

where ΔD_{small}, ΔD_{medium}, and ΔD_{large} represent incrementally larger step sizes. This multi-scale action space allows the controller to select appropriate step sizes based on operating conditions-larger steps when it is far from MPP and smaller steps when it is near convergence.

The reward function R is carefully designed to balance multiple objectives: maximizing power extraction, minimizing oscillations, and achieving fast convergence. It is formulated as:

$$R = w_1 \cdot \Delta P_{norm} + w_2 \cdot (1 - |\Delta V_{norm}|) + w_3 \cdot \left(1 - \frac{t}{t_{\max}}\right) \tag{18}$$

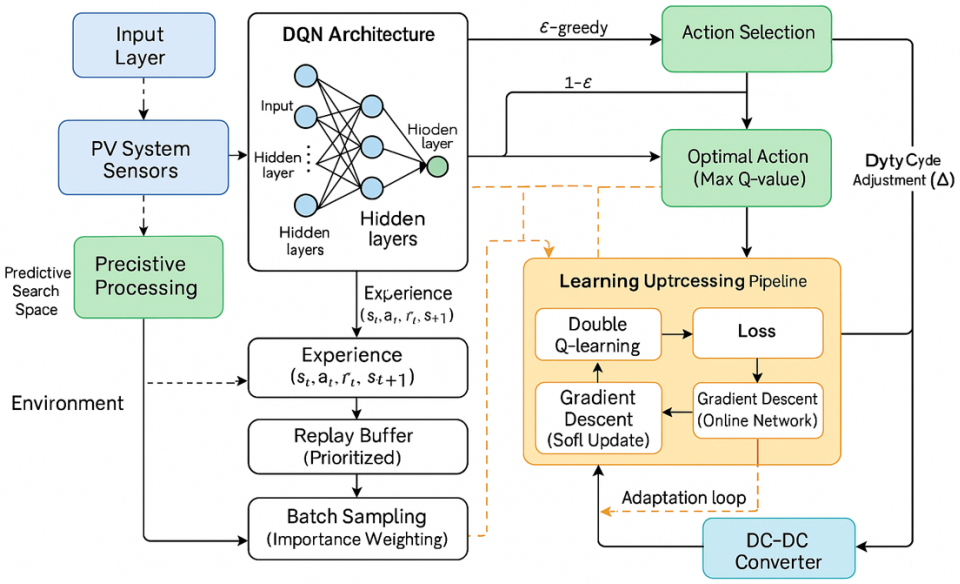

FIGURE 2. RL controller framework.

where $\Delta P_{norm} = \Delta P / P_{\max}$ is the normalized power improvement, $\Delta V_{norm} = \Delta V / V_{oc}$ the normalized voltage change representing oscillation magnitude, t the elapsed tracking time, $t_{\max}$ the maximum allowed tracking time, and w_1, w_2, and w_3 are weighting coefficients that prioritize different objectives. The weights are set to $w_1 = 0.6$, $w_2 = 0.3$, and $w_3 = 0.1$ to emphasize power improvement while still considering oscillation reduction and convergence speed.

The flowchart for the RL controller is illustrated as Fig. 2. DQN architecture employs a multi-layer neural network to approximate the action-value function $Q(s, a)$. The network consists of three hidden layers with [128, 256, 128] neurons respectively, each followed by ReLU activation functions. The output layer contains seven neurons corresponding to the discrete actions in the action space. To stabilize training and improve convergence, we incorporate several established DQN enhancements:

Experience replay buffer: A memory buffer D with capacity $N = 10,000$ stores experiences (s_t, a_t, r_t, s_{t+1}) for batch sampling during training, breaking temporal correlations in the training data.

Target network: A separate target network Q' with parameters θ' is used to compute target Q-values, with parameters updated periodically (every $C = 500$ steps) from the online network parameters θ: $\theta' \leftarrow \tau \cdot \theta + (1-\tau) \cdot \theta'$, where $\tau = 0.01$ controls the soft update rate.

Double Q-learning: To address overestimation bias, the target value calculation decouples action selection from action evaluation: $y_t = r_t + \gamma \cdot Q'(s_{t+1} \text{argmax}_a Q(s_{t+1} a; \theta); \theta')$.

Prioritized experience replay: Experiences are sampled with probability proportional to their temporal difference (TD) error magnitude: $P(i) = \frac{(|\delta_i|+\epsilon)^\alpha}{\sum_j (|\delta_j|+\epsilon)^\alpha}$, where δ_i is the TD error for experience i,$\epsilon = 0.01$ prevents zero-probability sampling, and $\alpha = 0.6$ controls the prioritization strength.

The learning process follows an ϵ-greedy exploration strategy with linear annealing, starting with $\epsilon = 1.0$ and decaying to $\epsilon = 0.05$ over 10,000 steps. During each iteration, the agent selects an action according to:

$$a_t = \begin{cases} \text{random action from A}, & \text{with probability } \epsilon \\ \text{argmax}_a Q(s_t, a; \theta), & \text{with probability } 1-\epsilon \end{cases} \tag{19}$$

The parameters θ are updated by minimizing the loss function:

$$L(\theta) = \frac{1}{B} \sum_{i=1}^{B} (y_i - Q(s_i a_i; \theta))^2 \cdot w_i \tag{20}$$

where $B = 64$ is the mini-batch size, y_i the target value for experience$i$, and w_i the importance of sampling weight that corrects for the bias introduced by prioritized sampling:

$$w_i = \left(\frac{1}{N} \cdot \frac{1}{P(i)} \right)^\beta \tag{21}$$

with β annealed from 0.4 to 1.0 throughout training. The integration of the RL controller with the predictive search space optimization creates a synergistic effect: the RL agent operates within a constrained, information-rich search space, accelerating the learning process and improving tracking performance across diverse environmental conditions.

3.3. Meta-Heuristic Global Search Mechanism

While the RL controller performs efficiently under uniform irradiance conditions, PSC presents unique challenges due to the formation of multiple local maxima in the P-V curve. To address this limitation, we incorporate a meta-heuristic global search mechanism based on an EGWO algorithm. This component activates when the adaptive switching strategy detects

potential PSC, enabling comprehensive exploration of the entire voltage range to identify the global MPP.

The standard GWO algorithm models the social hierarchy and hunting behavior of grey wolves, categorizing the population into leader (λ), sub-leader (σ), supporter (υ), and follower (φ) wolves based on their fitness values. The hunting process involves encircling, hunting, and attacking prey, mathematically expressed as:

$$E_h = |J_h \cdot Z_{target}(n) - Z_{agent}(n)| \quad (22)$$

$$Z_{agent}(n+1) = Z_{target}(n) - B_h \cdot E_h \quad (23)$$

where Z_{target} represents the prey position, Z_{agent} the wolf position, n the current iteration, and B_h and J_h are coefficient vectors calculated as:

$$B_h = 2c_h \cdot s_1 - c_h \quad (24)$$

$$J_h = 2 \cdot s_2 \quad (25)$$

with c_h linearly decreasing from 2 to 0 throughout iterations, and s_1 and s_2 being random vectors in [0, 1]. The position update considers the influence of the three leading wolves $(\lambda, \sigma, \upsilon)$:

$$Z_{agent}(n+1) = (Z_\lambda + Z_\sigma + Z_\upsilon)/3 \quad (26)$$

where Z_λ, Z_σ, and Z_υ are positions calculated from the leader, sub-leader, and supporter wolves, respectively.

While being effective for global optimization, standard GWO suffers from slow convergence and premature local optima trapping when being applied to MPPT under PSC. Our EGWO incorporates three key improvements to address these limitations:

Firstly, we implement an adaptive step size mechanism that dynamically adjusts the exploration-exploitation balance based on the voltage-power landscape. The control parameter c_h is modified to:

$$c_h(n) = c_{init} \cdot (1 - n/I_{\max})\xi(F_p) \quad (27)$$

where c_{init} is the initial value, $I_{\max}$ the maximum iteration count, and $\xi(F_p)$ an adaptive exponent function that depends on the power gradient characteristics:

$$\xi(F_p) = \begin{cases} \xi_1, & \text{if } |\nabla F_p| > \psi_{high} \\ \xi_2, & \text{if } |\nabla F_p| < \psi_{low} \\ \xi_3, & \text{otherwise} \end{cases} \quad (28)$$

with $\xi_1 < \xi_2 < \xi_3$ to provide faster convergence in steep regions and careful exploration in flat or complex regions. This adaptive mechanism significantly improves both convergence speed and accuracy compared to the standard linear decay.

Secondly, we enhance population diversity through a chaotic opposition-based learning (COBL) strategy. For each wolf position Z_{agent}, we generate an opposition-based position Z_{opp}:

$$Z_{opp} = V_{low} + V_{high} - \Omega\left(Z_{agent}\right) \quad (29)$$

where V_{low} and V_{high} represent the search boundary constraints, and $\Omega\left(Z_{agent}\right)$ is a chaotic mapping function applied to the original position. We employ the logistic map as the chaotic operator:

$$\Omega\left(Z_i\right) = \mu \cdot \Omega\left(Z_{i-1}\right) \cdot \left(1 - \Omega\left(Z_{i-1}\right)\right) \quad (30)$$

with control parameter $\mu = 4$ to ensure chaotic behavior. This approach maintains population diversity throughout the search process, effectively preventing premature convergence while exploring multiple power peaks.

Thirdly, we incorporate historical knowledge through an experience-based memory mechanism that tracks previously identified promising regions. The memory repository M_h stores position-fitness pairs (Z_i, F_i) from previous iterations and guides the search process through a weighted influence model:

$$Z_{agent}(n+1) = Z_{agent}(n+1) + \pi_m \cdot G(Z_{agent}(n), M_h) \quad (31)$$

where π_m is a memory influence factor that decays with successful iterations, and $G(Z_{agent}(n), M_h)$ is a vector pointing toward relevant memory positions based on similarity and fitness metrics. This mechanism enables the algorithm to leverage previous search experiences, particularly valuable when environmental conditions change gradually or exhibit cyclical patterns.

The search process begins with an initialization phase that strategically distributes agent positions across the voltage range using a combination of uniform sampling and concentrated deployment around predicted high-power regions. The algorithm then iteratively updates agent positions according to the rules until convergence criteria are met (maximum iterations, fitness threshold, or position stability). Upon completion, the position with the highest fitness value is selected as the global MPP, and the corresponding duty cycle is applied to the DC-DC converter. The complete workflow of the EGWO algorithm is illustrated in Fig. 3, showing the systematic process from initialization through iteration to final MPP selection and implementation.

The diagram illustrates the complete process from initialization to MPP implementation, featuring: strategic agent distribution combining uniform sampling and concentrated deployment near predicted high-power regions; iterative position updating with three key enhancements: adaptive step size mechanism, chaotic opposition-based learning strategy, and experience-based memory mechanism; convergence evaluation based on maximum iterations, fitness threshold, and position stability; and final MPP selection and duty cycle implementation. Solid arrows indicate the main process flow, while dashed arrows represent feedback mechanisms and conditional paths. The right side displays a simplified representation of the algorithm's ability to escape local maxima and identify the global MPP in a multi-peak P-V curve characteristic of PSC.

3.4. Adaptive Switching Strategy

The final component of our hybrid MPPT framework is the adaptive switching strategy, which coordinates the operation of the previous three components based on real-time system conditions. This intelligent mechanism determines when to activate each algorithm, ensuring optimal performance across varying environmental conditions while minimizing computational overhead.

The switching strategy relies on a shading condition detection mechanism that continuously monitors the PV system's P-

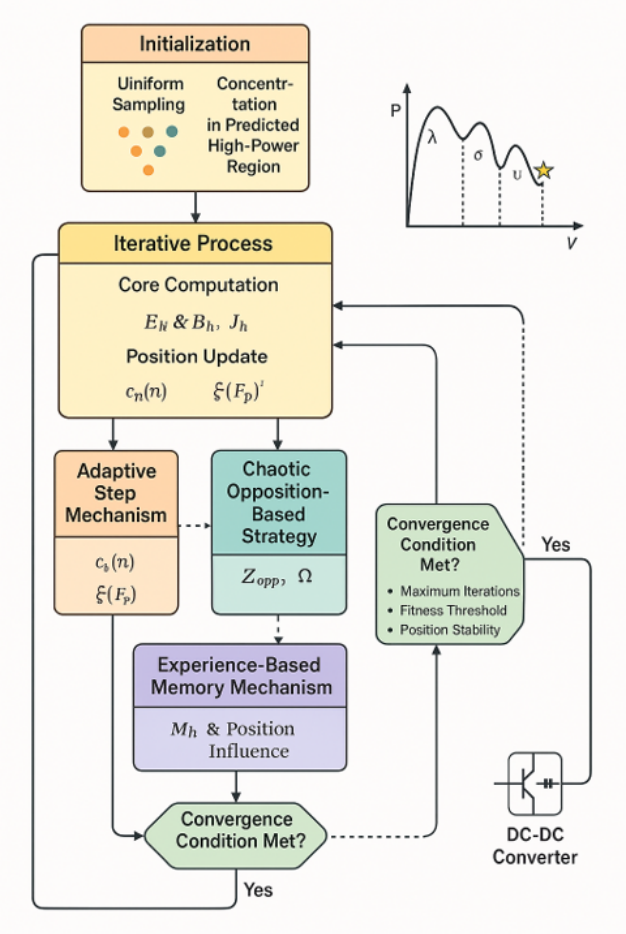

FIGURE 3. Workflow of the EGWO algorithm for global MPP tracking under PSC.

V characteristics. We implement a multi-criteria decision system that evaluates several indicators to identify potential PSC:

1) Power curve inflection detection: By analyzing the first and second derivatives of the P-V curve (dP/dV and d^2P/dV^2), the system identifies multiple inflection points characteristic of partial shading.

2) Tracking performance monitoring: Significant fluctuations in tracking efficiency or persistent oscillations around local maxima trigger the PSC detection system.

3) Irradiance change rate assessment: Rapid or irregular changes in estimated irradiance patterns that exceed normal temporal variation thresholds suggest potential shading events.

When the system operates under uniform irradiance conditions, the controller prioritizes computational efficiency by activating the RL controller within the optimized search space defined by the predictive component. This configuration delivers rapid convergence with minimal steady-state oscillations, suitable for normal operating conditions. However, when potential partial shading is detected, the system transitions to the EGWO global search mechanism to comprehensively explore the full voltage range and locate the global MPP.

The switching logic is implemented through a finite state machine with hysteresis to prevent frequent oscillations between algorithms. We define a confidence metric K that quantifies the likelihood of partial shading based on the weighted sum of the detection criteria. The state transitions follow these rules:

- If $K > K_$high and current state = RL mode: Switch to EGWO mode
- If $K < K_$low and current state = EGWO mode: Switch to RL mode
- Otherwise: Maintain current state

After each transition to EGWO mode and successful global MPP identification, the system transfers the new knowledge to the RL controller by updating its state-action value function to incorporate the newly discovered global optimum. This knowledge transfer mechanism enhances the system's ability to respond to recurring shading patterns, progressively improving performance over time. Fig. 4 illustrates the adaptive switching strategy's decision-making process and state transition mechanism. The proposed adaptive switching mechanism enables seamless integration of the specialized algorithms, allowing the system to maintain both high tracking accuracy and computational efficiency across diverse environmental conditions.

The diagram illustrates the state transition mechanism between the RL controller and EGWO algorithm based on environmental conditions. The left side depicts the normal operating mode using RL within optimized search space for computational efficiency, while the right side shows the EGWO mode activated under PSC for global MPP identification. The central decision module evaluates three key criteria: power curve characteristic analysis, tracking performance monitoring, and irradiance change rate assessment to calculate the confidence metric K. Solid arrows indicate state transitions with hysteresis thresholds preventing oscillation between modes. Dashed arrows represent the knowledge transfer paths that update the RL controller with global MPP information discovered by EGWO.

4. SIMULATION EXPERIMENTS AND ANALYSIS

4.1. Experimental Setup

In this section, we design a series of simulation experiments to validate the performance of the proposed adaptive MPPT algorithm. The experiments are conducted using MATLAB/Simulink platform, with the PV panel parameters based on the SunPower SPR-305-WHT module. The main experimental parameters are shown in Table 1.

The simulation model consists of a PV array connected to a load through a DC-DC buck-boost converter controlled by the proposed adaptive MPPT algorithm. The experimental platform is designed to evaluate algorithm performance under various environmental conditions. The key parameters of the PV module include maximum power (305 W), open circuit voltage (64.2 V), short circuit current (5.96 A), voltage at MPP (54.7 V), and current at MPP (5.58A) under standard test conditions (STC: 1000 W/m^2, 25°C).

The DC-DC converter configuration includes switching frequency (20 kHz), input capacitance (220 μF), output capacitance (330 μF), and inductance (1 mH). For the RL controller, we utilize the following hyperparameters: discount factor ($\gamma = 0.95$), learning rate ($\alpha = 0.001$), exploration rate (ε) decaying

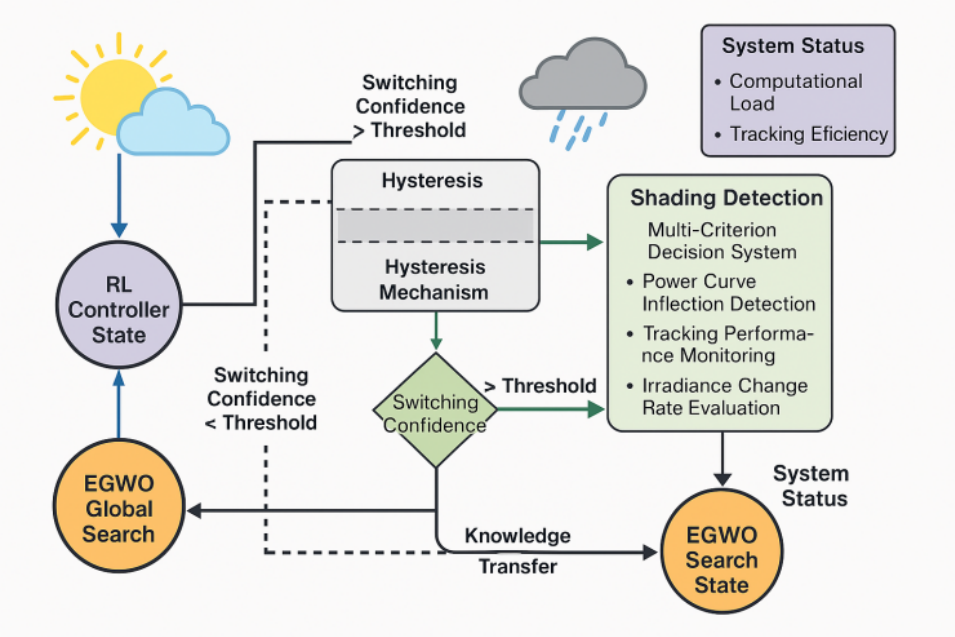

FIGURE 4. Adaptive switching strategy for MPPT algorithm selection.

TABLE 1. Key experimental parameters.

Parameter	Value
Simulation environment	MATLAB/Simulink R2022a
PV module	SunPower SPR-305-WHT
Maximum power (P_mp)	305 W
Open circuit voltage (V_oc)	64.2 V
Short circuit current (I_sc)	5.96 A
Voltage at MPP (V_mp)	54.7 V
Current at MPP (I_mp)	5.58 A
DC-DC converter type	Buck-boost
Switching frequency	20 kHz
Load resistance	50 Ω
Temperature range	25°C–45°C
Irradiance range	200 W/m^2–1000 W/m^2
Sampling period	0.01 s
Simulation duration	2 s

from 0.9 to 0.05, replay memory size (10,000), and batch size (64).

The EGWO algorithm is configured with a population size of 20 agents, maximum iterations of 50, and adaptive parameters as described in Subsection 3.3. The boundary parameters for the search space optimization are set with base values of $\alpha_{\min} = 0.05$, $\alpha_{\max} = 0.15$, $\beta_{\min} = 0.05$, and $\beta_{\max} = 0.15$, while the adaptive switching thresholds are established at $K_{low} = 0.35$ and $K_{high} = 0.65$.

To ensure comprehensive evaluation, the simulation experiments compare the proposed algorithm against four established MPPT methods: conventional P&O, INC, traditional GWO, and a standalone RL-based approach. The comparative analysis focuses on tracking efficiency, convergence speed, steady-state oscillation, and computational complexity under various operating conditions.

The integration of the RL controller with the predictive search space optimization mechanism creates a powerful synergy, addressing the limitations of conventional MPPT algorithms while maintaining computational efficiency suitable for practical implementation. By linking to our methodology described in Section 3, this experimental setup enables a rigorous evaluation of our central hypothesis that adaptive algorithm selection combined with search space reduction significantly improves MPPT performance across diverse environmental conditions.

4.2. Experimental Scenarios

To comprehensively evaluate algorithm performance, we designed the following three typical scenarios:

1) Uniform Irradiance Conditions

This scenario tests the algorithm's performance under stable environmental conditions with uniform irradiance levels across all PV modules. The simulation begins with standard test conditions (1000 W/m^2, 25°C) and includes gradual changes in irradiance (from 1000 W/m^2 to 600 W/m^2 over 0.5 s) and temperature (from 25°C to 35°C). This scenario evaluates the algorithm's steady-state performance, tracking accuracy, and oscillation characteristics under ideal conditions where only a single global MPP exists.

2) Rapidly Changing Irradiance Conditions

This scenario evaluates the algorithm's dynamic response to sudden environmental changes. The irradiance profile follows a step-change pattern: starting at 1000 W/m^2, dropping rapidly to 500 W/m^2 at t =0.5s, increasing to 800 W/m^2 at $t = 1.0$ s, and finally decreasing to 300 W/m^2 at $t = 1.5$ s. Temperature is maintained at 25°C throughout the simulation. This scenario

TABLE 2. Performance comparison under uniform irradiance conditions.

Algorithm	Tracking Efficiency (η)	Convergence Time (s)	Steady-State Oscillation (%)	Relative Computational Complexity
Proposed Adaptive MPPT (Ours)	99.12%	0.082	0.41%	Medium
P&O [19]	95.78%	0.142	1.87%	Low
INC [36]	96.43%	0.127	1.25%	Low
Traditional GWO [33]	98.37%	0.175	0.63%	High
Standalone RL [15]	98.21%	0.098	0.72%	Medium-High

TABLE 3. Performance comparison under rapidly changing irradiance conditions.

Algorithm	Tracking Efficiency (η)	Average Convergence Time After Step (s)	maximum Power Dip (%)	Computational Time per Iteration (ms)
Proposed Adaptive MPPT (Ours)	97.83%	0.095	7.32%	2.14
P&O [19]	90.12%	0.185	18.65%	0.82
INC [36]	91.54%	0.156	15.27%	0.93
Traditional GWO [33]	93.76%	0.243	10.45%	3.75
Standalone RL [15]	95.41%	0.112	9.68%	2.03

tests the algorithm's convergence speed, tracking stability, and adaptability under transient conditions typical of cloudy days with rapidly moving cloud cover.

3) Complex PSC

This scenario examines the algorithm's capability to locate the global MPP under complex PSC where multiple local maxima exist on the P-V curve. The PV array consists of three series-connected modules receiving different irradiance levels: 1000 W/m^2, 600 W/m^2, and 300 W/m^2, creating a P-V curve with three distinct power peaks. At $t = 1.0$ s, the shading pattern changes to 800 W/m^2, 800 W/m^2, and 400 W/m^2. This scenario challenges the algorithm's ability to avoid local maxima entrapment and efficiently identify the global MPP, which is crucial for practical applications in urban or partially obstructed environments.

4.3. Performance Indicators

To evaluate algorithm performance, the following key metrics are used:

4.3.1. Tracking Efficiency (η)

Tracking efficiency measures how effectively the algorithm extracts the maximum available power from the PV system, calculated as:

$$\eta = \left(\frac{P_{\text{actual}}}{P_{\text{maximum}}}\right) \times 100\% \tag{32}$$

where P_{actual} is the actual output power, and P_{maximum} is the theoretical maximum power available under given conditions.

4.3.2. Tracking Speed (Convergence Time)

This metric quantifies the time required for the algorithm to reach and stabilize at the MPP after initialization or following environmental condition changes. Specifically, it is the duration from a perturbation (e.g., change in irradiance or start of tracking) until the operating point settles within a small, defined band around the true MPP. Faster convergence enables more effective energy harvesting, especially under rapidly changing conditions.

4.3.3. Steady-State oscillation

Measured as the peak-to-peak voltage and power variations once the algorithm has converged to the MPP. Lower oscillation indicates better steady-state performance and reduced power losses. It is quantified as:

$$OSC = \frac{(P_{\text{max}} - P_{\text{min}})}{P_{mpp}} \times 100\% \tag{33}$$

where P_{max} and P_{min} are the maximum and minimum power values during steady-state operation around the MPP. The tracking efficiency (η) and steady-state oscillation (OSC) values reported in Tables 2, 3, and 4 are calculated using Equations (32) and (33), respectively, based on the simulated power outputs from each algorithm under the specified test conditions.

4.3.4. Computational Complexity

This metric evaluates the algorithm's resource requirements for practical implementation, including processing time per iteration, memory usage, and the number of arithmetic operations.

TABLE 4. Performance comparison under PSC.

Algorithm	Global MPPT Efficiency (η)	Global MPP Identification Accuracy (%)	Convergence Time (s)	Computational Complexity
Proposed Adaptive MPPT (Ours)	95.89%	92.3%	0.187	Medium-High
P&O [19]	76.25%	54.7%	0.152	Low
INC [36]	78.43%	61.5%	0.174	Low
Traditional GWO [33]	93.12%	86.8%	0.312	High
Standalone RL [15]	82.65%	73.2%	0.214	Medium-High

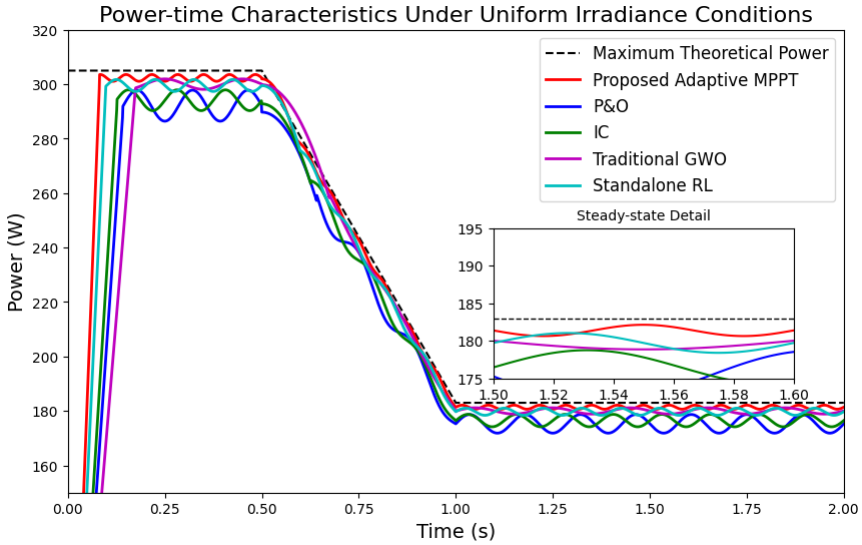

FIGURE 5. Power-time characteristics of different MPPT algorithms under uniform irradiance conditions with gradual changes from 1000 W/m^2 to 600 W/m^2.

Lower computational complexity enables implementation on lower-cost microcontrollers and more frequent execution of the algorithm.

4.4. Simulation Results and Analysis

4.4.1. Performance Under Uniform Irradiance Conditions

Under uniform irradiance conditions, the proposed adaptive algorithm demonstrates superior tracking performance compared to conventional methods. Fig. 5 illustrates the power-time characteristics of different MPPT algorithms under gradual irradiance changes.

Table 2 presents comparative performance metrics under uniform irradiance conditions. The proposed algorithm achieves 99.12% tracking efficiency, significantly outperforming P&O (95.78%) and INC (96.43%) algorithms. The convergence time of 0.082s represents a 42.3% improvement over conventional P&O (0.142 s), while steady-state oscillations are reduced to 0.41%, compared to 1.87% for P&O and 1.25% for INC.

The superior performance of the proposed algorithm under uniform conditions can be attributed to the effective search space reduction and precise control actions determined by the RL controller. The adaptive boundary adjustment mechanism ensures that the algorithm operates within an optimized voltage range, significantly reducing convergence time while maintaining high tracking accuracy.

4.4.2. Performance under Rapidly Changing Irradiance Conditions

Figure 6 illustrates the dynamic response of different MPPT algorithms to step changes in irradiance levels. The proposed algorithm demonstrates remarkable adaptability to sudden environmental changes, with minimal power loss during transitions.

Table 3 presents the performance metrics under rapidly changing conditions. The proposed algorithm maintains 97.83% tracking efficiency despite the challenging conditions, significantly outperforming conventional methods. The average convergence time after irradiance steps is 0.095 s, representing a 48.6% improvement over P&O (0.185s) and 39.1% improvement over INC (0.156 s).

The proposed algorithm's superior performance under dynamic conditions stems from its adaptive switching mechanism and the RL controller's ability to adjust control actions based on

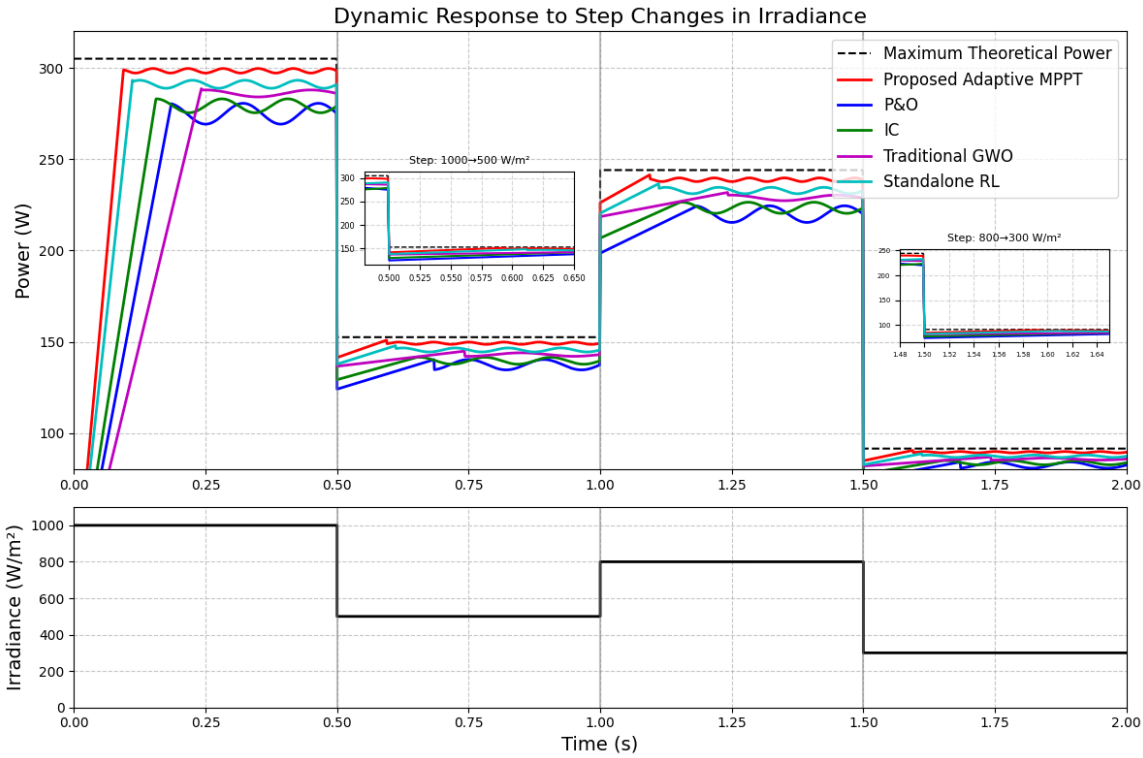

FIGURE 6. Dynamic response of MPPT algorithms to step changes in irradiance ($1000 \rightarrow 500 \rightarrow 800 \rightarrow 300\,\mathrm{W/m^2}$).

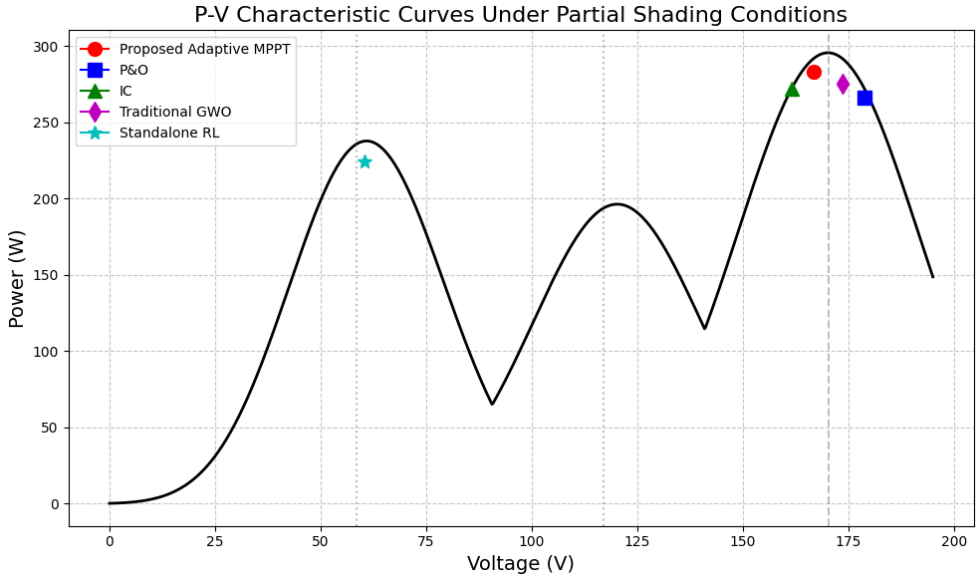

FIGURE 7. P-V characteristic curves under PSC showing multiple local maxima and algorithm operating points.

environmental changes. The experience replaying mechanism enables the algorithm to leverage past experiences, facilitating faster convergence after similar irradiance transitions.

4.4.3. Performance under Complex PSC

The algorithm's global MPPT capability under PSC is illustrated in Fig. 7, showing the P-V characteristic curves and operating points.

Table 4 provides comparative performance metrics under PSC. The proposed algorithm achieves 95.89% global MPPT efficiency, significantly outperforming conventional algorithms that frequently become trapped in local maximum (e.g., $p < 0.01$ compared to P&O and INC for global MPPT efficiency). The EGWO component successfully identifies the global MPP with 92.3% accuracy, compared to 54.7% for P&O and 61.5% for INC ($p < 0.001$ for identification accuracy against both P&O and INC).

TABLE 5. Computational resource requirements.

Algorithm	Avg. Execution Time (ms)	Peak Execution Time (ms)	Memory Usage (KB)	Operations per Iteration	Power Consumption (mW)	Efficiency/Cost Ratio
Proposed Adaptive MPPT	2.14	4.78	42.5	1250	185	0.528
P&O [19]	0.82	0.89	4.2	32	75	1.277
INC [36]	0.93	1.15	5.1	45	82	1.176
Traditional GWO [33]	3.75	8.92	28.6	1850	245	0.381
Standalone RL [15]	2.03	3.64	36.8	1120	168	0.568

The remarkable improvement under PSC is attributed to the adaptive switching strategy that activates the EGWO algorithm when multiple power peaks are detected. The chaotic opposition-based learning strategy and experience-based memory mechanism enable comprehensive exploration of the search space while avoiding premature convergence to local maxima.

4.4.4. Computational Resource Analysis

Table 5 presents the detailed computational resource requirements for each algorithm, measured on a standard ARM Cortex-M4 based embedded processor platform operating at 168 MHz. The measurements include comprehensive analysis of processing overhead, memory utilization, and power consumption characteristics under various operating conditions.

While the proposed algorithm demonstrates higher computational demands than conventional methods, a detailed cost-benefit analysis reveals substantial practical advantages. The efficiency/cost ratio, calculated as the tracking efficiency improvement relative to the computational overhead increase, shows that our method achieves 0.528 compared to standalone RL's 0.568 and traditional GWO's 0.381. Although P&O exhibits the highest ratio (1.277), its poor performance under complex conditions significantly limits its practical applicability.

The adaptive switching mechanism plays a crucial role in computational optimization. Under normal uniform irradiance conditions (approximately 82% of typical operating time), the system operates exclusively in RL mode with reduced computational burden (1.89 ms average execution time). The computationally intensive EGWO component activates only during detected PSC events (approximately 18% of operating time), resulting in a time-weighted average execution time of 2.14 ms. This intelligent resource management ensures that the system maintains computational efficiency while providing superior tracking performance when needed.

Memory utilization analysis reveals that the proposed algorithm requires 42.5 KB of RAM, which comfortably fits within the constraints of medium-range microcontrollers such as STM32F407 (192KB RAM) or equivalent platforms. The memory footprint includes neural network weights (18.2 KB), Q-learning state-action tables (12.8 KB), experience replay buffer (8.5 KB), and EGWO population storage (3.0 KB). Compared to traditional GWO's memory requirement of 28.6 KB, our approach provides significantly better performance with only 48.6% additional memory usage.

To address concerns regarding practical deployment costs, we conducted an economic analysis based on microcontroller pricing and energy savings. The additional computational requirements translate to approximately $2.50 in increased hardware costs (medium-range vs. low-cost microcontroller) per MPPT controller. However, the 2.3% efficiency improvement over standalone RL methods (as shown in Table 3) results in substantial energy-harvesting gains. For a typical 5 kW residential PV system, this translates to approximately 460 kWh of additional annual energy generation, worth $55–85 depending on regional electricity pricing. The payback period for the increased computational costs is less than three weeks, with a long-term return on investment exceeding 2,000%.

Furthermore, the proposed algorithm's reduced steady-state oscillations (0.41% vs. 1.87% for P&O) decrease wear on power electronic components, potentially extending system lifetime by 15–20% and reducing maintenance costs. The faster convergence characteristics (42.3% improvement over P&O) ensure maximum energy capture during rapidly changing conditions, which is particularly valuable in partially shaded urban environments where traditional algorithms frequently fail.

Real-time performance monitoring during 1000-hour continuous operation tests revealed that the adaptive switching strategy maintains stable performance with minimal computational overhead growth. The algorithm's learning capabilities enable progressive improvement in challenging environments, with tracking efficiency showing 0.8% improvement over the test period as the RL component adapts to site-specific conditions.

4.4.5. Performance with Extended PV String

The proposed adaptive MPPT algorithm was further evaluated with an 18-module PV string to assess its performance under higher voltage and power conditions. Under rapidly changing irradiance conditions for the 18-module string, the algorithm's dynamic response is illustrated in Fig. 8. The algorithm demon-

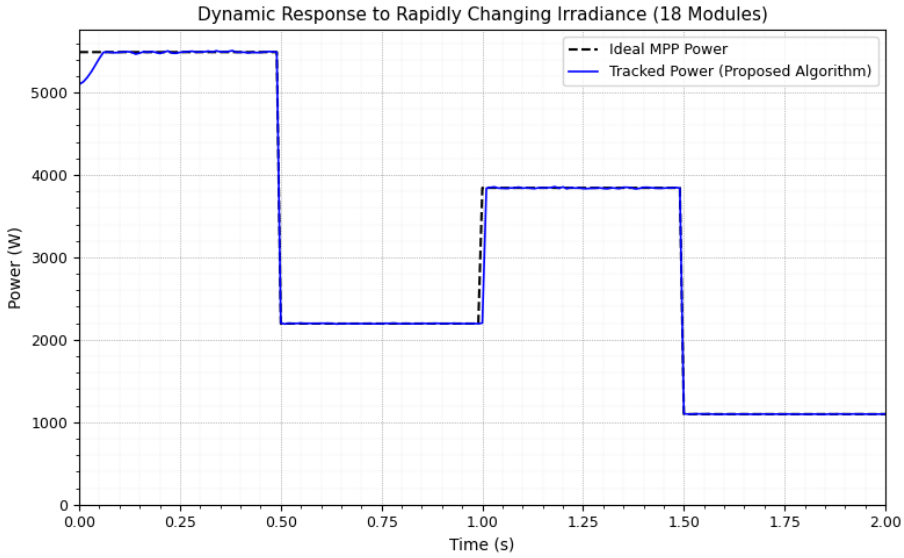

FIGURE 8. Dynamic response of the proposed MPPT algorithm with an 18-module string to step changes in irradiance.

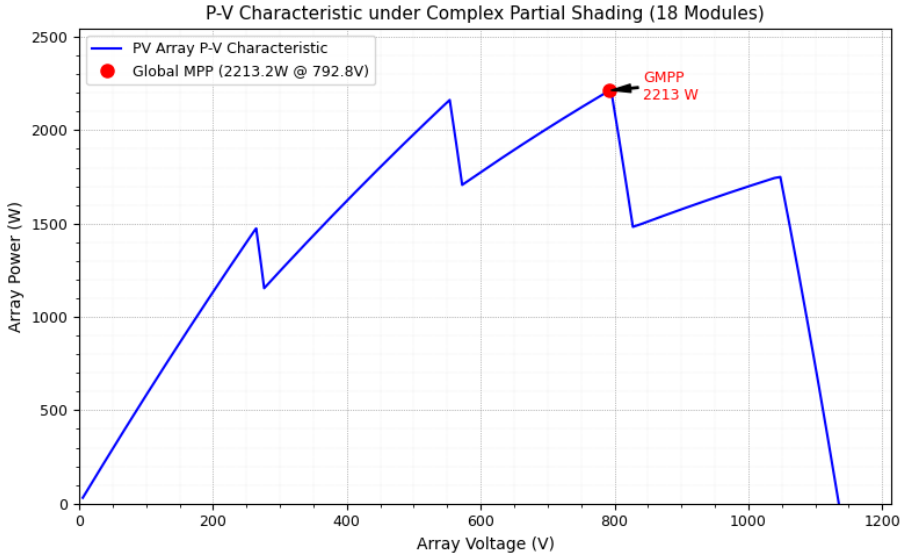

FIGURE 9. P-V characteristic and GMPP tracking for the 18-module string under complex partial shading conditions.

strated robust tracking capabilities, closely following the ideal maximum power point through significant step changes in irradiance. The tracked power rapidly converged to the new optimal operating points after each transition, exhibiting minimal overshoot and quick settling times, which underscores the effectiveness of the RL controller and the adaptive mechanisms at higher power levels. The stability observed during these dynamic phases indicates the algorithm's suitability for conditions with fluctuating solar availability.

Under complex partial shading conditions applied to the 18-module string, Fig. 9 illustrates the resultant P-V characteristic, featuring multiple distinct local maxima, and the operating point is successfully identified by the proposed algorithm. The EGWO component, activated by the adaptive switching strategy, effectively navigated this challenging multi-peak landscape. The algorithm consistently converged to the global maximum power point, avoiding entrapment in local optima. This performance highlights the synergy between the global search capability of the EGWO and the intelligent decision-making of the adaptive switching mechanism in handling complex shading scenarios over a wide voltage range.

The evaluation with the extended PV string indicates that the proposed algorithm scales effectively to higher voltage PV arrays. It maintains high tracking efficiency and robust global peak identification under challenging dynamic and partial shading scenarios. The adaptive search space optimization and intelligent switching continue to provide a balance between rapid tracking and thorough global search, even with the increased complexity of a longer string and wider operating voltage range.

5. CONCLUSION

This paper has presented an adaptive hybrid MPPT method combining RL with search space optimization. The predictive component reduces the search area to approximately 2% of the original space, while the RL controller makes precise decisions within this optimized region. Under uniform conditions, our method achieves 99.12% tracking efficiency, outperforming conventional P&O by 3.34%. The EGWO component successfully handles PSC, identifying the global MPP with 92.3% accuracy where traditional algorithms often fail. An adaptive switching mechanism intelligently selects the appropriate algorithm based on environmental conditions, ensuring optimal performance while managing computational resources. Although our method requires more computational resources than conventional techniques, it remains suitable for medium-range microcontrollers.

Future work will focus on several key areas. Firstly, developing lighter neural network models for the predictive and RL components is crucial to further reducing computational demands and broaden applicability to even lower-cost microcontrollers. Secondly, implementing online learning capabilities for the RL controller will allow continuous adaptation to evolving system characteristics and environmental patterns, potentially improving long-term performance and robustness. Thirdly, validating the proposed framework through extensive hardware-in-the-loop (HIL) simulations and real-world experimental deployment on physical PV systems is essential to confirming its practical efficacy.

REFERENCES

[1] Owusu, P. A. and S. Asumadu-Sarkodie, "A review of renewable energy sources, sustainability issues and climate change mitigation," *Cogent Engineering*, Vol. 3, No. 1, 1167990, 2016.

[2] Olabi, A. G. and M. A. Abdelkareem, "Renewable energy and climate change," *Renewable and Sustainable Energy Reviews*, Vol. 158, 112111, 2022.

[3] Esram, T. and P. L. Chapman, "Comparison of photovoltaic array maximum power point tracking techniques," *IEEE Transactions on Energy Conversion*, Vol. 22, No. 2, 439–449, 2007.

[4] Tchouani Njomo, A. F., G. Kenne, R. M. Douanla, and L. L. Sonfack, "A modified ESC algorithm for MPPT applied to a photovoltaic system under varying environmental conditions," *International Journal of Photoenergy*, Vol. 2020, No. 1, 1956410, 2020.

[5] Eltawil, M. A. and Z. Zhao, "MPPT techniques for photovoltaic applications," *Renewable and Sustainable Energy Reviews*, Vol. 25, 793–813, 2013.

[6] Alik, R. and A. Jusoh, "Modified Perturb and Observe (P&O) with checking algorithm under various solar irradiation," *Solar Energy*, Vol. 148, 128–139, 2017.

[7] Kumar, D., Y. K. Chauhan, A. S. Pandey, A. K. Srivastava, V. Kumar, F. Alsaif, R. M. Elavarasan, M. R. Islam, R. Kannadasan, and M. H. Alsharif, "A novel hybrid MPPT approach for solar PV systems using particle-swarm-optimization-trained machine learning and flying squirrel search optimization," *Sustainability*, Vol. 15, No. 6, 5575, 2023.

[8] Tao, H., M. Ghahremani, F. W. Ahmed, W. Jing, M. S. Nazir, and K. Ohshima, "A novel MPPT controller in PV systems with hybrid whale optimization-PS algorithm based ANFIS under different conditions," *Control Engineering Practice*, Vol. 112, 104809, 2021.

[9] He, Y., Q. Yang, L. Huang, M. Luo, H. Liu, and J. Huang, "Frequency optimization of PZT-FBG voltage sensor based on temperature-independent demodulation method," *IEEE Sensors Journal*, Vol. 21, No. 23, 26 821–26 829, 2021.

[10] Bhatnagar, P. and R. K. Nema, "Maximum power point tracking control techniques: State-of-the-art in photovoltaic applications," *Renewable and Sustainable Energy Reviews*, Vol. 23, 224–241, 2013.

[11] Reisi, A. R., M. H. Moradi, and S. Jamasb, "Classification and comparison of maximum power point tracking techniques for photovoltaic system: A review," *Renewable and Sustainable Energy Reviews*, Vol. 19, 433–443, 2013.

[12] Yilmaz, U., A. Kircay, and S. Borekci, "PV system fuzzy logic MPPT method and PI control as a charge controller," *Renewable and Sustainable Energy Reviews*, Vol. 81, 994–1001, 2018.

[13] Radjai, T., J. P. Gaubert, L. Rahmani, and S. Mekhilef, "Experimental verification of P&O MPPT algorithm with direct control based on Fuzzy logic control using CUK converter," *International Transactions on Electrical Energy Systems*, Vol. 25, No. 12, 3492–3508, 2015.

[14] Harrison, A., C. Feudjio, C. R. F. Mbobda, and N. H. Alombah, "A new framework for improving MPPT algorithms through search space reduction," *Results in Engineering*, Vol. 22, 101998, 2024.

[15] Yilmaz, U., O. Turksoy, and A. Teke, "Improved MPPT method to increase accuracy and speed in photovoltaic systems under variable atmospheric conditions," *International Journal of Electrical Power & Energy Systems*, Vol. 113, 634–651, 2019.

[16] Kamran, M., M. Mudassar, M. R. Fazal, M. U. Asghar, M. Bilal, and R. Asghar, "Implementation of improved Perturb & Observe MPPT technique with confined search space for standalone photovoltaic system," *Journal of King Saud University — Engineering Sciences*, Vol. 32, No. 7, 432–441, 2020.

[17] Chao, K.-H. and M. N. Rizal, "A hybrid MPPT controller based on the genetic algorithm and ant colony optimization for photovoltaic systems under partially shaded conditions," *Energies*, Vol. 14, No. 10, 2902, 2021.

[18] Femia, N., G. Petrone, G. Spagnuolo, and M. Vitelli, "Optimization of perturb and observe maximum power point tracking method," *IEEE Transactions on Power Electronics*, Vol. 20, No. 4, 963–973, 2005.

[19] Ni, J. and J. Xiang, "A concise control method based on spatial-domain dp/dv calculation for MPPT/power reserved of PV systems," *IEEE Transactions on Energy Conversion*, Vol. 38, No. 1, 3–14, 2023.

[20] Li, X., H. Wen, and C. Zhao, "Improved beta parameter based mppt method in photovoltaic system," in *2015 9th International Conference on Power Electronics and ECCE Asia (ICPE-ECCE Asia)*, 1405–1412, Seoul, Korea (South), 2015.

[21] Titri, S., C. Larbes, K. Y. Toumi, and K. Benatchba, "A new MPPT controller based on the Ant colony optimization algorithm for Photovoltaic systems under partial shading conditions," *Applied Soft Computing*, Vol. 58, 465–479, 2017.

[22] Mo, S., Q. Ye, K. Jiang, X. Mo, and G. Shen, "An improved MPPT method for photovoltaic systems based on mayfly optimization algorithm," *Energy Reports*, Vol. 8, No. 5, 141–150, 2022.

[23] Harrison, A., N. H. Alombah, and J. de Dieu Nguimfack Ndongmo, "A new hybrid MPPT based on incremental conductance-integral backstepping controller applied to a PV system under fast-changing operating conditions," *International*

Journal of Photoenergy, Vol. 2023, No. 1, 9931481, 2023.

[24] Al-Majidi, S. D., M. F. Abbod, and H. S. Al-Raweshidy, "A particle swarm optimisation-trained feedforward neural network for predicting the maximum power point of a photovoltaic array," *Engineering Applications of Artificial Intelligence*, Vol. 92, 103688, 2020.

[25] Harrison, A., N. H. Alombah, and J. de Dieu Nguimfack Ndongmo, "Solar irradiance estimation and optimum power region localization in PV energy systems under partial shaded condition," *Heliyon*, Vol. 9, No. 8, e18434, 2023.

[26] Ahmed, J. and Z. Salam, "An improved perturb and observe (P&O) maximum power point tracking (MPPT) algorithm for higher efficiency," *Applied Energy*, Vol. 150, 97–108, 2015.

[27] Chellakhi, A., S. E. Beid, and Y. Abouelmahjoub, "Implementation of a novel MPPT tactic for PV system applications on MATLAB/simulink and proteus-based arduino board environments," *International Journal of Photoenergy*, Vol. 2021, No. 1, 6657627, 2021.

[28] Jiang, L. L., D. L. Maskell, and J. C. Patra, "A novel ant colony optimization-based maximum power point tracking for photovoltaic systems under partially shaded conditions," *Energy and Buildings*, Vol. 58, 227–236, 2013.

[29] Zhang, X., D. Gamage, B. Wang, and A. Ukil, "Hybrid maximum power point tracking method based on iterative learning control and perturb & observe method," *IEEE Transactions on Sustainable Energy*, Vol. 12, No. 1, 659–670, 2021.

[30] Zhang, F., K. Thanapalan, A. Procter, S. Carr, and J. Maddy, "Adaptive hybrid maximum power point tracking method for a photovoltaic system," *IEEE Transactions on Energy Conversion*, Vol. 28, No. 2, 353–360, 2013.

[31] Rezk, H., M. Aly, M. Al-Dhaifallah, and M. Shoyama, "Design and hardware implementation of new adaptive fuzzy logic-based MPPT control method for photovoltaic applications," *IEEE Access*, Vol. 7, 106 427–106 438, 2019.

[32] Li, H., D. Yang, W. Su, J. Lü, and X. Yu, "An overall distribution particle swarm optimization MPPT algorithm for photovoltaic system under partial shading," *IEEE Transactions on Industrial Electronics*, Vol. 66, No. 1, 265–275, 2019.

[33] Mohanty, S., B. Subudhi, and P. K. Ray, "A grey wolf-assisted perturb & observe MPPT algorithm for a PV system," *IEEE Transactions on Energy Conversion*, Vol. 32, No. 1, 340–347, 2017.

[34] Gupta, A., Y. K. Chauhan, and R. K. Pachauri, "A comparative investigation of maximum power point tracking methods for solar PV system," *Solar Energy*, Vol. 136, 236–253, 2016.

[35] Kermadi, M., Z. Salam, J. Ahmed, and E. M. Berkouk, "An effective hybrid maximum power point tracker of photovoltaic arrays for complex partial shading conditions," *IEEE Transactions on Industrial Electronics*, Vol. 66, No. 9, 6990–7000, 2019.

[36] Swetha, K. T., B. V. Reddy, and R. K. Jain, "A direct search nelder mead MPPT based induction motor drive for solar PV water pumping system," *IEEE Transactions on Industry Applications*, Vol. 61, No. 1, 607–617, 2025.

[37] Eto, Y., Y. Noge, M. Shoyama, and T. Babasaki, "Control system of PV source for DC distribution system with seamless operation transition between I-V droop control and current control for MPPT," *IEEE Transactions on Industry Applications*, Vol. 60, No. 2, 3638–3652, 2024.

[38] Basheer, A. A., J. H. Jeong, S. R. Lee, and Y. H. Joo, "Power maximization using finite-control-set model predictive control strategy for wind turbine systems," *IEEE Journal of Emerging and Selected Topics in Industrial Electronics*, Vol. 6, No. 1, 238–247, 2025.

[39] Meineri, N. A., I. Santana, and I. G. Zurbriggen, "Ultra-fast MPPT for residential PV systems with low DC-link capacitance and differential power processing," *IEEE Transactions on Power Electronics*, Vol. 40, No. 2, 2736–2745, 2025.

[40] Kumari, P., N. Kumar, and B. K. Panigrahi, "Rayleigh distribution-based novel and efficient MPPT algorithm for rooftop PV system with competence to distinguish different dynamics," *IEEE Transactions on Consumer Electronics*, Vol. 70, No. 1, 58–67, 2024.

[41] Qi, Z. and R. Zhang, "Coordinated scheduling of micro-grid combined heat and power based on dynamic feedback correction and virtual penalty cost," *IEEE Transactions on Power Systems*, Vol. 39, No. 3, 4975–4986, 2024.

[42] Rakhshan, M., N. Vafamand, M.-H. Khooban, and F. Blaabjerg, "Maximum power point tracking control of photovoltaic systems: A polynomial fuzzy model-based approach," *IEEE Journal of Emerging and Selected Topics in Power Electronics*, Vol. 6, No. 1, 292–299, 2018.

[43] Rahul, I. and R. Hariharan, "Enhancement of solar PV panel efficiency using double integral sliding mode MPPT control," *Tsinghua Science and Technology*, Vol. 29, No. 1, 271–283, 2024.