Journal of Industrial and Management Optimization

JIMO, 22(10): 5039–5073.
DOI: 10.3934/jimo.2026174
Received: 15 July 2026
Revised: 27 August 2026
Accepted: 08 September 2026
Published: 15 September 2026

https://www.aimspress.com/journal/jimo

Research article

An intelligent algorithm based on reinforcement learning to identify the parameters of solar systems

Xuming Wang[1], Wen Zhang[2,*], Xiaobing Yu[2]

[1] Engineering Training Center, Nanjing University of Information Science & Technology, Nanjing 210044, China
[2] School of Management Science and Engineering, Nanjing University of Information Science & Technology, Nanjing 210044, China

***Correspondence:** E-mail: 202212240021@nuist.edu.cn.

Abstract: Accurately identifying the parameters of solar systems is an essential measure to improve the efficiency of turning solar energy into electrical energy. An intelligent algorithm based on reinforcement learning was developed to identify photovoltaic parameters. The proposed algorithm includes three novelties: i) two new search methods are proposed, namely local search and global search, in which exploitation and exploration of the algorithm are balanced; ii) the ranking strategy is used to determine appropriate search methods for each individual; and iii) the main parameters of two search methods are adjusted by the reinforcement learning technique. To verify the effectiveness of the proposed algorithm in identifying parameters from solar systems, the proposed algorithm was utilized on five photovoltaic models and compared the outcomes with those obtained by other algorithms. The test cases include the single-diode and double-diode models and three photovoltaic panel module models. Performance was assessed using the root mean square error between measured and estimated current values. Ablation experiments examined the contributions of the search strategies, ranking mechanism, and reinforcement learning–based parameter adjustment, while additional comparisons evaluated convergence, stability, and computational time. The final experimental results verified the excellent performance of the proposed algorithm in identifying parameters from solar systems, proving its usefulness in the field of artificial intelligence.

Keywords: solar systems; parameter identification; reinforcement learning; optimization algorithm; photovoltaic models; hybrid heuristic algorithm

Mathematics Subject Classification: 90C29

1. Introduction

In the study of solar systems, parameter identification is a crucial task with significant implications. Parameter identification permits gaining insight into the performance and efficiency of solar systems. The operational status and performance of solar systems can be evaluated by extracting key parameters. In addition, the energy conversion efficiency and output power of solar systems can also be calculated. These data are crucial for improving solar systems, optimizing system structure, and enhancing the generation capacity and stability of systems. Besides, through regular identification and analysis of system parameters, faults or abnormalities can be detected in a timely manner, such as component aging, pollution, and blockages. Timely maintenance or replacement measures based on discovered faults help to avoid performance degradation or even complete failure of solar systems. Furthermore, optimizing parameter identification can also benefit the economy and the environment. By accurately identifying solar system parameters, the energy conversion efficiency of PV cells can be improved. This helps to improve the overall efficiency of the energy system while reducing dependence on traditional energy sources and promoting sustainable development in the field of renewable energy. From an economic perspective, this optimization is expected to reduce manufacturing and operating costs, while also reducing the cost burden associated with energy production. In terms of the environment, solar systems are a clean energy technology that can reduce dependence on traditional energy and effectively reduce adverse impacts on the environment.

A variety of methods and techniques can be used for parameter extraction of solar systems. Field testing and monitoring [1] is one of the most direct and common methods. Sensors and measuring devices are installed to monitor the current, voltage, temperature, and other parameters of photovoltaic (PV) components. Then, the measured data is recorded and analyzed. Field testing and monitoring can provide real-time system parameter information to support the performance and efficiency of solar systems. However, conducting field testing and monitoring takes considerable time and resources, and the precision of parameter extraction is greatly affected by environmental conditions.

Mathematical modeling and computer simulation methods [2] can predict and estimate key parameters of solar systems. By establishing the mathematical model of solar systems, their performance under different working conditions can be simulated, such as output current, voltage, and power generation. This approach can optimize systems and parameter configuration. However, mathematical modeling may lead to errors due to model simplification or unconsidered physical effects. For large solar systems, it may require a lot of computational resources and time to build and solve mathematical models.

Methods based on intelligent algorithms [3,4] have the capability to handle substantial data volumes and extract solar system parameters. These models enable improved accuracy and efficiency through automation, while also having the flexibility to be customized and optimized for different solar systems and application scenarios. Compared with the other methods mentioned above, methods based on intelligent algorithms not only reduce workload and cost associated with actual measurements but also address model uncertainty by extensively searching and optimizing across a vast array of samples.

Intelligent algorithms mainly include swarm intelligence (SI) algorithms and evolutionary algorithms. SI algorithms simulate the cooperative behavior observed in biological groups, such as insects or birds, and seek optimal solutions through information sharing among individuals. SI algorithms can quickly find a more optimal solution by conducting a local search in the vicinity of the current optimal solution. This local search ability enables algorithms to quickly converge to the vicinity of the optimal solution during iteration, thereby improving search efficiency. Some algorithms based on SI have been proposed to solve parameter extraction of solar systems [5,6]. However, these methods also present some problems. For example, excessive local search ability may cause algorithms to overly focus on the sub-optimal solution during the search process, thereby preventing them from exploring optimal solutions.

Evolutionary algorithms emulate the biological evolution process in order to discover optimal solutions. During the iterative process, individuals in a population generate new individuals by continuously performing crossover and mutation operations, and a selection strategy is used to decide which individuals will survive. In this way, the evolutionary algorithm is able to conduct a global search in the solution space to find an optimal solution. Evolutionary algorithms represented by the differential evolution (DE) algorithm show advantages through unique global search mechanisms [7]. DE discovers potential global optimal solutions by exploring the entire solution space; however, emphasizing this behavior may also deteriorate the exploitation of the algorithm. Thus, it is necessary to introduce a balancing mechanism, thereby enhancing its exploration ability, convergence rate, and robustness. Meanwhile, the parameter configuration of these algorithms directly impacts their search efficiency and quality. Consequently, parameter adjustment is a significant step in the practical application of these algorithms. Through a reasonable adjustment of these parameters, algorithms can better adapt to the problem features and increase search efficiency. Some adaptive methods are also used for parameter adjustment. For example, a self-adaptive weight is used to adjust the search process when optimizing solar systems [8]. Despite the effectiveness of adaptive parameter control mechanisms, most existing approaches primarily focus on adjusting algorithmic parameters while keeping the underlying search strategies unchanged. However, parameter configuration and search strategy jointly determine the search behavior of evolutionary algorithms, and adapting one without the other may limit the algorithm's ability to respond to dynamically changing search environments.

As machine learning continues to be developed, reinforcement learning (RL) has been proposed as a model-free approach [9]. RL learns the optimal behavioral strategy through the interactions between the agent and the environment; with appropriate reward mechanisms and strategy selection, RL can better tune parameters. Moreover, by interacting with the environment, an agent receives reward feedback, allowing RL to gradually discover which parameter settings produce better results. These automatic parameter optimization methods can effectively reduce the workload of parameter adjustment and find the appropriate parameters in complex problems. Several studies have explored the integration of RL into the optimization of solar energy systems; for instance, Hu et al. employed RL to adaptively adjust the parameters of a DE algorithm [10]. Evolutionary optimization performance is governed by two closely coupled components: the search strategy and parameter configuration. Focusing exclusively on parameter adaptation while leaving the search strategy unchanged may limit the algorithm's ability to effectively balance exploration and exploitation. Therefore, a more effective optimization framework should jointly adapt the search strategy and parameter configuration to accommodate the dynamically changing optimization environment.

Based on the above observation, an intelligent hybrid heuristic algorithm based on reinforcement learning (HHARL) is proposed for the accurate identification of solar system parameters. In HHARL, two search strategies, local search (LS) and global search (GS), are proposed to effectively balance the exploitation and exploration ability of the algorithm. A ranking strategy is developed to reasonably guide individuals to choose appropriate search strategies. In the LS strategy, the surrounding solution space is explored by making small changes and adjustments to the current optimal solution, with the aim of finding a better solution within the local scope. In the GS strategy, the focus of the algorithm is to search the entire solution space to locate a global optimal solution. To better leverage the advantages of these two search strategies, the main parameters of the two search methods are regulated by RL. The main contributions of this paper are as follows: An intelligent hybrid heuristic algorithm based on reinforcement learning (HHARL) is proposed for a more accurate identification of solar system parameters; local search (LS) and global search (GS) strategies adjusted by the ranking strategy are proposed to balance the exploitation and exploration capabilities of the algorithm; the main parameters that control the degree of variation and crossover of individuals in a population are adjusted by the RL technique; by comparing HHARL with other advanced algorithms and common optimization algorithms, it is shown that HHARL can effectively and reliably identify solar system parameters.

The remainder of this paper is organized as follows: Section 2 introduces PV models. Section 3 describes the improvement strategies and processes of HHARL in detail. Section 4 summarizes the results of the algorithm for PV model parameter identification. Section 5 provides a conclusion.

2. Photovoltaic models

Performance evaluation and mathematical modeling play a crucial role in the design and optimization of solar systems, as they assess the efficiency of PV cells—devices that convert solar energy into electrical energy. Two commonly used mathematical models for PV cells are the single diode model (SDM) and double diode model (DDM) [11]. Based on PV cell models, the PV panel module model takes into account how cells are connected within a cell assembly and how they influence each other. These mathematical models are vital for optimizing the design of PV cells and performing system-level modeling and simulation. By carefully selecting appropriate models and accurately extracting model parameters, the behavior of PV cells can be comprehended. This, in turn, could increase the efficiency of solar systems through optimized design strategies. The physical meaning of the parameters involved in all models is summarized in Table 1.

Table 1. Physical meaning of parameters used in PV models.

Param.	Meaning	Param.	Meaning
I	PV cell output current	n_2	Recombination diode ideal factor
I_D	Diode current	k	Boltzmann constant (1.3806503× 10^{-23} J/K)
I_{D1}	First diode current	q	Electron charge (1.60217646 × 10^{-19} C)
I_{D2}	Second diode current	T	Temperature of junction
I_{ph}	Introduced current	N	Number of experimental data
I_{sh}	Shunt resistor current	N_p	Number of solar cells in parallel
I_{sd}	Reverse saturation current of a diode	N_s	Number of solar cells in series
I_{sd1}	Diffusion current	V	PV cell output voltage
I_{sd2}	Saturation current	V_t	Junction thermal voltage
n	Emission coefficient	R_s	Series resistance
n_1	Diffusion diode ideal factor	R_{sh}	Shunt resistance

2.1. SDM

SDM is one of the simplest mathematical models for PV cells, and its equivalent circuit diagram is shown in Figure 1. It is based on a current-voltage characteristic curve and assumes that a PV cell is an ideal diode [12]. In accordance with Kirchhoff's current law, Eqs. (1) to (3) are employed to describe the current-voltage relationship in the SDM:

$$I = I_{ph} - I_D - I_{sh} \tag{1}$$

$$I_D = I_{\text{sd}}\left[\exp\left(\frac{V+IR_{\text{s}}}{nV_t}\right) - 1\right] \tag{2}$$

$$I_{sh} = \frac{V+IR_{\text{s}}}{R_{\text{sh}}} \tag{3}$$

The junction thermal voltage V_t is the potential generated by the temperature difference between two points in a closed circuit and is calculated by Eq. (4):

$$V_t = \frac{k\times T}{q} \tag{4}$$

The equivalent output current equation of the SDM can be calculated by substituting Eqs. (2)–(4) into Eq. (1), resulting in Eq. (5):

$$I = I_{\text{ph}} - I_{\text{sd}}\left[\exp\left(\frac{V+IR_{\text{s}}}{nV_t}\right) - 1\right] - \frac{V+IR_{\text{s}}}{R_{\text{sh}}} \tag{5}$$

There are five parameters that need to be identified in SDM, namely $I_{\text{ph}}, I_{\text{sd}}, R_{\text{s}}, R_{\text{sh}}$, and *n*.

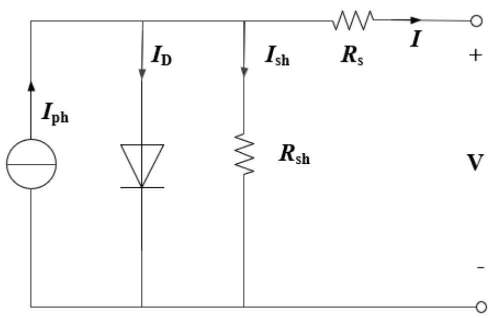


Figure 1. Equivalent circuit diagram of SDM.

2.2. DDM

SDM considers only forward-on characteristics of the diode, so the precision of the model decreases when a diode is reverse-biased. In contrast, DDM takes into account not only forward and reverse cutoff cases but also the effect of temperature on the diode, so it can describe the diode ground characteristics more accurately [11]. The equivalent circuit diagram of DDM is shown in Figure 2. Eqs. (6)–(7) are used to describe the current-voltage relationship in DDM:

$$I = I_{ph} - I_{D1} - I_{D2} - I_{sh} \tag{6}$$

$$I_t = \quad I_{ph} - I_{sd_1}\left[\exp\left(\frac{V+IR_{\text{s}}}{n_1V_{\text{t}}}\right) - 1\right] - I_{\text{sd}_2}\left[\exp\left(\frac{V+IR_{\text{s}}}{n_2V_{\text{t}}}\right) - 1\right] - \frac{V+IR_{\text{s}}}{R_{\text{sh}}} \tag{7}$$

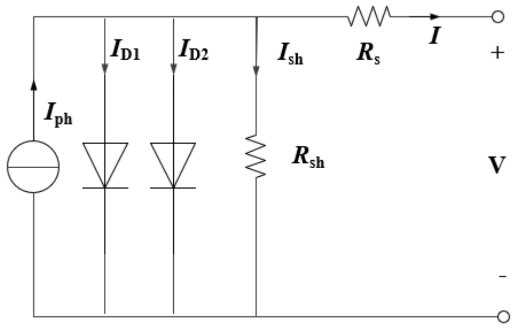


Figure 2. Equivalent circuit diagram of DDM.

It can be seen that there are seven parameters needed to be identified in DDM, namely $I_{\mathrm{ph}}, I_{\mathrm{sd1}}, I_{\mathrm{sd2}}, R_{\mathrm{s}}, R_{\mathrm{sh}}, n_1$, and n_2.

2.3. PV panel module model

Both SDM and DDM ignore some influencing factors or make some approximate assumptions, so they are relatively simplified PV models. PV panel module models consider the internal structure and operation of solar systems, including multiple current and voltage components, allowing for a more detailed and accurate description of solar systems. The equivalent circuit diagram of the PV panel module model is shown in Figure 3. Eq. (8) is used to describe the current-voltage relationship in the PV panel module model:

$$I = I_{ph} - I_{sd}\left[\exp\left(\frac{V+IN_sR_s}{n_1N_sV_t}\right) - 1\right] - \frac{V+IN_sR_s}{R_{sh}N_s} \tag{8}$$

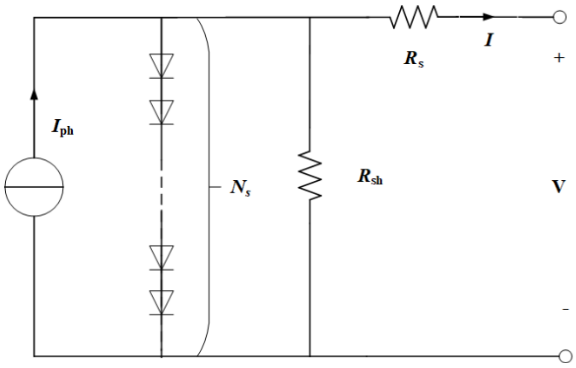


Figure 3. Equivalent circuit diagram of the PV panel module model.

Five parameters need to be identified in the PV panel module model, namely $I_{\mathrm{ph}}, I_{\mathrm{sd}}, R_{\mathrm{s}}, R_{\mathrm{sh}}$, and *n*.

2.4. PV model optimization objective

Similar to other research in this field [13], the accuracy of algorithms is evaluated by quantifying the difference between the estimated current and measured current in experimental settings. In this paper, the absolute error is represented by the function $f(V, I, x)$, where *x* represents the value of the parameter to be identified in PV models. Eqs. (9)–(11) give the absolute error function expressions of the SDM, DDM, and PV panel module models, respectively.

$$f(V,I,x) = I_{\mathrm{ph}} - I_{\mathrm{sd}}\left[\exp\left(\frac{V+IR_{\mathrm{s}}}{nV_{\mathrm{t}}}\right) - 1\right] - \frac{V+IR_{\mathrm{s}}}{R_{\mathrm{sh}}} - I \tag{9}$$

$$f(V,I,x) = I_{ph} - I_{sd_1}\left[\exp\left(\frac{V+IR_s}{n_1V_t}\right) - 1\right] - I_{sd_2}\left[\exp\left(\frac{V+IR_s}{n_2V_t}\right) - 1\right] - \frac{V+IR_s}{R_{sh}} - I \tag{10}$$

$$f(V,I,x) = I_{\mathrm{ph}} - I_{\mathrm{sd}}\left[\exp\left(\frac{V+IR_{\mathrm{s}}N_{\mathrm{s}}}{nN_{\mathrm{s}}V_{\mathrm{t}}}\right) - 1\right] - \frac{V+IR_{\mathrm{s}}N_{\mathrm{s}}}{R_{\mathrm{sh}}N_{\mathrm{s}}} - I \tag{11}$$

The root mean square error (RMSE) is a widely used evaluation metric for assessing the accuracy and error magnitude of prediction models. It is calculated by taking the square root of the mean square deviation between simulated and measured values. Because RMSE calculates the square of the error, it is more sensitive to big errors. This means that if there is a significant prediction mistake, the RMSE is more clearly reflected in the assessment rather than merely taking an absolute value. RMSE presents advantages over other metrics, such as the mean absolute error (MAE) and the mean squared error (MSE), in that it is more sensitive to outliers and presents some mathematical convenience. To summarize, RMSE is used as the objective function in this analysis, as in the study [13]. The calculation formula is presented in Eq. (12):

$$RMSE(x) = \sqrt{\frac{1}{N}\sum_{i=1}^{N} f_i(V,I,x)^2} \tag{12}$$

3. An intelligent algorithm based on reinforcement learning

In this section, we will elaborate on the proposed intelligent algorithm for identifying solar system parameters. First, we explain the basic principles of RL. Then, HHARL is introduced from search strategy, using RL technology to adjust parameters, algorithm flow, and computational complexity.

3.1. Reinforcement learning

RL can be traced back to the dynamic programming theory of the 1950s and 1960s, proposed by Richard Bellman [9]. RL is a machine-learning approach to allow agents to make optimal decisions through interactions with the environment. It simulates the way humans learn, gaining knowledge and experience through trial and error. In RL, agents can observe the state of environments and take actions, which in turn elicit responses and provide a reward signal indicating the quality of the actions. The objective of the agent is to maximize the cumulative reward through interactions with environments. Thus, the agent tries different actions and adapts its strategy based on the feedback from the environment for higher rewards. It decides which actions to take in different states by learning a policy or a mapping from states to actions. In addition, RL provides a flexible learning framework capable of handling complex decision-making problems while adapting to changing environments. Through interaction with the environment, RL enables the agent to learn from past experiences and gradually improves its decision-making ability. Nowadays, RL has found diverse applications across various fields.

As a value-based algorithm in RL, Q-learning was proposed by Christopher Watkins in 1989 [10]. Q-learning can learn optimal policies by interacting with the environment. The fundamental concept of Q-learning revolves around acquiring a state-action value function, which estimates the expected cumulative reward achievable by taking specific actions in given states. The state-action value represents the cumulative reward that the agent can earn in a series of future decisions starting from the current state after making an action. The learning process of Q-learning is based on Eq. (13):

$$NewQ(s,\ a) = Q\ (s,a) + \alpha \times [r + \gamma \times maxQ'(s',a') - Q(s,a)] \tag{13}$$

Where $NewQ(s,\ a)$ is a new Q-value based on state and action, $Q\ (s,a)$ is the current Q-value, α is the learning efficiency factor, r is the reward value based on status and action, γ is the discount factor, s' is a new state after performing action a, and a' is the next action selected under the new state s'. The learning process of Q-learning gradually approximates the optimal strategy by constantly updating the state-action value function. At each moment, the agent chooses an action based on its current state and observes feedback from the environment, including instant rewards and new states. Then, the corresponding state-action value is updated by Eq. (13) to incorporate new experiences into the learning process. The diagram of Q-learning is shown in Figure 4. The state reward in Figure 4 represents the instant reward received after executing action a in a specific state s. The state reward can be positive, negative, or zero, depending on the design of the environment and the nature of the task. The goal of the agent is to learn so that the actions taken in each state maximize the cumulative reward, not just the immediate reward in the current state.

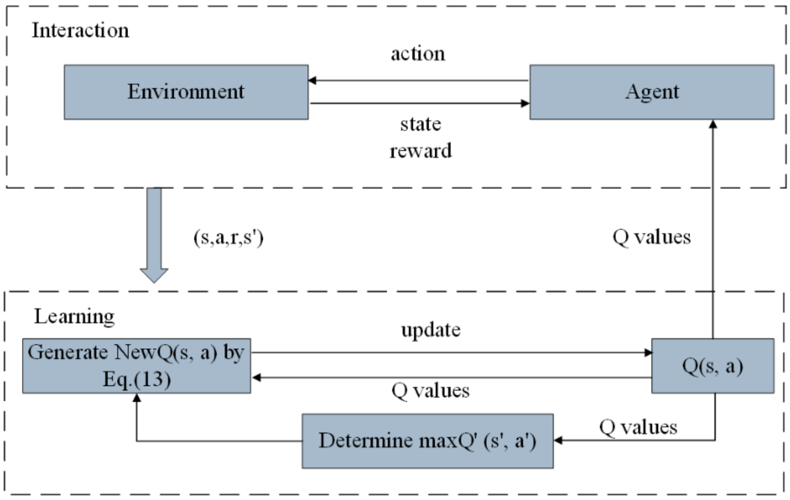


Figure 4. The diagram of Q-learning.

3.2. An intelligent algorithm

HHARL mainly uses LS and GS to balance the exploitation and exploration ability of the algorithm. To fully utilize the advantages of the two search strategies, RL is used to fine-tune the key parameters, considering the key advantages offered by Q-learning [10]: (1) it has a simple and streamlined structure; (2) it does not rely on prior knowledge; and (3) it enables continuous learning and adaptation during task execution, without requiring the task to be completed beforehand. Based on the above analysis, the Q table provides a clear state-action mapping in Q-learning. For small state spaces, the Q table can efficiently store and update Q values. In this case, the agent can make a quick decision by looking up the values in the table. So, this paper adopted Q table technology. Furthermore, a ranking strategy is established to guide individuals in selecting the appropriate search strategy. Detailed descriptions of these enhancements are provided below. Also, the flow of HHARL is given, and the computational complexity analysis is explained.

3.2.1. Two search methods

Exploration and exploitation are two important concepts in intelligent algorithms that determine how the algorithm makes trade-offs and decisions between different options with limited resources. Exploration refers to the active probing of unknown regions or selected spaces to obtain more information and knowledge. In the exploration phase, the algorithm tries different options or actions to discover potentially beneficial information. Exploitation refers to maximizing the benefits of a given choice space and making decisions based on existing knowledge and experience. In the exploitation phase, the algorithm uses known information to select an option that has been validated or is considered the best.

However, most SI and evolutionary algorithms tend to emphasize either exploration or exploitation, making it challenging to effectively coordinate these two complementary search behaviors. For example, SI algorithms typically rely on the best-performing individuals to guide population evolution [14,15], which may strengthen exploitation but increase the risk of premature convergence. In contrast, evolutionary algorithms often introduce mutation operators to enhance population diversity and exploration [7], potentially at the expense of local refinement. Such an imbalance between exploration and exploitation can substantially restrict the search capability and robustness of optimization algorithms. Therefore, designing effective mechanisms to coordinate and dynamically balance exploration and exploitation is of fundamental importance for achieving superior optimization performance.

HHARL uses two search methods, namely LS and GS. Among the many SI algorithms, the teaching-learning-based optimization algorithm (TLBO) has the advantages of good robustness, few parameters, and easy implementation [16]. The strategy of learning from teachers in TLBO is a great local search strategy. LS is designed to enhance the search for neighborhoods surrounding the best individuals within a population. LS can be described as Eq. (14):

$$X_i^{t+1} = X_i^t + F_i^t * \left(X_{cb} - \frac{1}{NP}\sum_{i=1}^{NP} X_i\right) \tag{14}$$

where X_i^t is the *ith* individual of *t* iteration, F_i^t is a random number between 0 and 1, *NP* is the number of individuals in the population, and X_{cb} represents the best individual in the current population. By increasing the value of *F*, the global search capability of the algorithm is augmented. This facilitates the algorithms to effectively break free from local optimal solutions and explore the solution space extensively. On the contrary, a smaller *F* is beneficial for local search, thereby maintaining the stability and convergence of the algorithm search.

Among evolutionary algorithms, the DE algorithm is known for its simplicity and robust global search capability [7]. DE searches the solution space by introducing differential operations and mutation strategies. The utilization of these strategies enhances the global search capability of the algorithms, allowing them to effectively explore and identify optimal solutions across a wide solution space. On this basis, a modified global search strategy, GS, is proposed. GS explores the information of other individuals so that it gains more opportunities to break free from the local optima, continue to search, and ultimately discover improved solutions. GS can be represented by Eqs. (15)–(16):

$$X_i^{t+1} = X_{tb}^t + F_i^t \times (X_{r2}^t - X_{r3}^t) \tag{15}$$

$$X_i^{t+1,j} = \begin{cases} X_i^{t+1,j}, if\ rand \le C_i\ or\ j = j_{rand} \\ X_i^{t,j},\ otherwise \end{cases} \tag{16}$$

where X_{tb}^{t} is the individual with the best fitness value among the three individuals randomly selected from the population in the t iteration, and X_{r2}^{t} and X_{r3}^{t} are the other two individuals. Moreover, C_i determines the probability of the ith individual selecting offspring variant individuals in the mutation operation. A smaller C can fully utilize the information of the parent, which benefits from converging to a more optimal solution in local search. A larger C makes the algorithm rely more on variant individuals, which leads to an expanded search space and increases the likelihood of global search. Besides, j stands for the dimension of the problem, and j_{rand} is a randomly selected index used to select a dimension of the individual for adjustment in the mutation operation, with values ranging from 1 to the total dimension of the problem.

Determining the appropriate search strategy for each individual is critical to maintaining a desirable balance between exploration and exploitation. Accordingly, a ranking-based mechanism is developed to quantify the search preference of each individual through the parameter (R). Based on the resulting value of (R), each individual adaptively selects either the LS or GS strategy, as defined in Eq. (17).

$$R(i) = \left(\frac{NP-i}{NP}\right)^2 \tag{17}$$

where i is the ith individual, and NP is the number of individuals in the population. In the ranking strategy, individuals with better fitness values are regarded as excellent individuals, so they have larger R-values. On the contrary, those who rank low are considered poor individuals and have little useful information in their current search direction, which is expected to increase the probability of them choosing the GS method to search other regions of the solution space. Therefore, search method selection based on ranking strategy is summarized in Eq. (18):

$$Method(i) = \begin{cases} LS, if\ r < R(i) \\ GS,\ otherwise \end{cases} \tag{18}$$

where r is a random number between 0 and 1, and $R(i)$ also ranges between 0 and 1. It can be seen that individuals with larger R values have more opportunities to conduct LS, which is conducive to utilizing their favorable information, and on the contrary, to conduct GS to jump out of locally optimal solutions.

3.2.2. Parameter adjustment based on RL

Conventional RL-based algorithms typically focus on adapting a single algorithmic parameter, which limits their ability to respond comprehensively to changes in the search environment [10]. In contrast, the proposed algorithm simultaneously adapts multiple key parameters in a coordinated manner, thereby providing more flexible control over the search dynamics and improving the overall adaptability and optimization performance.

In Q-learning, a Q table is a two-dimensional table that stores the Q value of state-action pairs. As shown in Figure 5, rows of the Q table represent states, and columns represent actions; each cell stores the Q value of the corresponding state and action. Incremental values of adjustment parameters are determined based on the Q table. In this paper, two states are set to provide a basis for making decisions: (1) the fitness value of the offspring is better than that of the parent, in which case the reward is set to 1; (2) the fitness value of the parent is better than that of the offspring, in which case the reward is set to 0. Besides, three actions for adjusting parameters are set, namely (1) f=-0.1, f_c =+0.1; (2) f=+0.1, f_c =+0.1; (3) f=+0, f_c =+0. The increment f is used to adjust F, and f_c is used to adjust C. The agent selects the optimal action based on the Q value. By learning and updating the Q value in the Q table, the agent gradually optimizes its search strategy, so that the optimal action is chosen in different states to obtain the maximum cumulative reward. The selection probability of action is calculated by Eq. (19):

$$\pi(s,a) = \frac{e^{Q(s_i,a_j)/T}}{\sum_{j=1}^{n} e^{Q(s_i,a_j)/T}} \tag{19}$$

where $Q(s_i, a_j)$ is the value of the Q table, and T is the control parameter. By adjusting the increment value based on the action, F and C are given by Eqs. (20)–(21).

$$F = F + f \tag{20}$$

$$C = C + f_c \tag{21}$$

In the process of HHARL iteration, if the offspring is better than the parent, it means that the search direction at this time is toward the direction of an optimal solution. Therefore, it is necessary to reduce the value of F to strengthen the current local search and reduce the influ ence of other individuals involved in the search, so that individuals rely more on their informa tion. In contrast, if the parent is better than the offspring, it means that the individual does not search for the optimal solution in the process of variation. In this case, the variation of an in dividual is enhanced by increasing the value of F, making more use of information from other individuals in the population. In addition, as can be seen from Eq. (16), a larger C value mea ns a greater probability of crossing with mutated individuals, so to retain more information ab out the offspring, HHARL prefers a larger C value.

After updating F and C, based on the relationship between the fitness values of offspring and parent, appropriate rewards and punishments are given to the selected behavior. Finally, update the Q table according to Eq. (13).

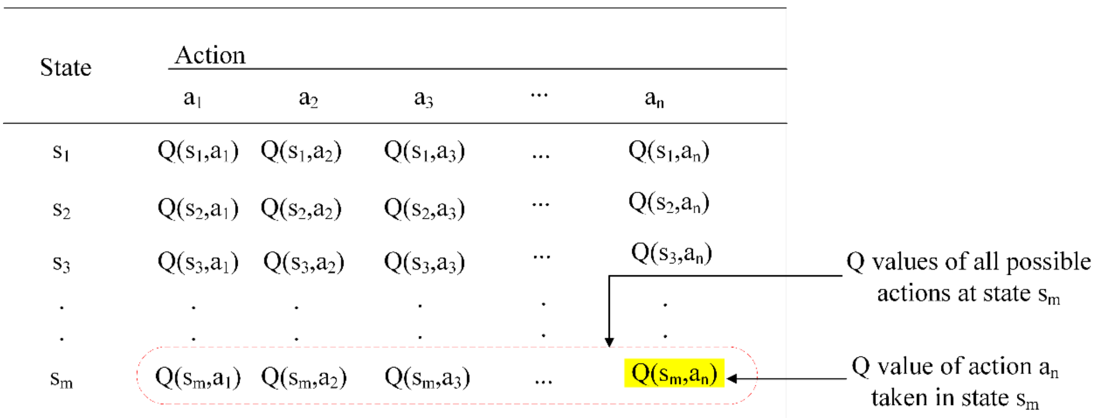


Figure 5. Details of the Q table.

3.2.3. Flow of the proposed algorithm

The entire process of running HHARL is given in the pseudo-code represented by Algorithm 1. First, initialize the population, parameters required by the algorithm, and Q table. Then, the parameters R of the ranking strategy are calculated, F and C are updated, and the corresponding search method is selected by Eq. (18) to generate new individuals. Then, update f and f_c by the selection probability of the Q table, calculate the fitness value of a new individual, and assign a reward value and state to individuals with fitness value. Finally, the Q table is updated, and the best individual and optimal solutions are preserved. The flow diagram for HHARL is shown in Figure 6. The main parameters of HHARL are set as follows: $\alpha = 0.1; \gamma = 0.9$; Softmax temperature $T = 1$; initial Q-table is a zero matrix, F and C range in [0,1], and their initial values are random values in [0,1].

3.2.4. Computational complexity

In HHARL, the number of individuals in population is defined as *NP*. *D* represents the dimension of the problem. According to the pseudo-code of HHARL, the computational complexity analysis can be obtained. First, initialize the population, which requires $o(NP \times D)$. Then, evaluating the agent and calculating the ranking parameter R require $o(NP + NP \times \log(NP) + NP)$. Updating F and C also needs $o(NP)$. Generating new individuals based on LS or GS involves creating *NP* individuals with D dimensions, so this step needs $o(NP \times D)$. Next, evaluating the fitness of each new individual and selecting the better individual requires $o(NP)$. Finally, updating the Q table requires $o(NP)$. In summary, the total computational complexity for the proposed HHARL is

$$\begin{aligned} &o(NP \times D + NP + NP \times \log(NP) + NP + NP + NP \times D + NP + NP) \\ &= o(2 \times NP \times D + 5 \times NP + NP \times \log(NP)) \end{aligned}$$

for each iteration. The dimension D is often more than 5. For each iteration, the complexity is equal to $o(NP \times D)$, according to the time complexity calculation criteria.

Algorithm 1: HHARL

Initialize the population and Q table;
Calculate the fitness value of each individual, find the optimal individual, and calculate the *Mean*;
While *NFE<Max_NFE* %*NFE*: number of fitness evaluations
for *i*=1 to *NP*; % *NP*: number of individuals in population
Calculate ranking parameter *R* for each individual by Eq. (17);
end for
for *i*=1 to *NP*;
Update *F* and *C* by Eqs. (20)–(21);
if *rand<R*
Generate new individuals based on LS by Eq. (14);
else if
Generate new individuals based on GS by Eqs. (15)–(16);
end if
Calculate the action selection probability of each individual with Eq. (19);
Choose action to adjust *F* and *C* by Eqs.(20–21);
Update the action of each agent;
Calculate the fitness value of the new individual and retain the better individual;
NFE=NFE+1;
if $f(X_i^{t+1})<f(X_i^t)$
Remain X_i^{t+1} and $f(X_i^{t+1})$;
Reward=1, state=1;
else if
Remain X_i^t and $f(X_i^t)$;
Reward=0, state=2;
end if
Update Q table by Eq. (13);
Retain the best individual and the best solution;
end for
end while

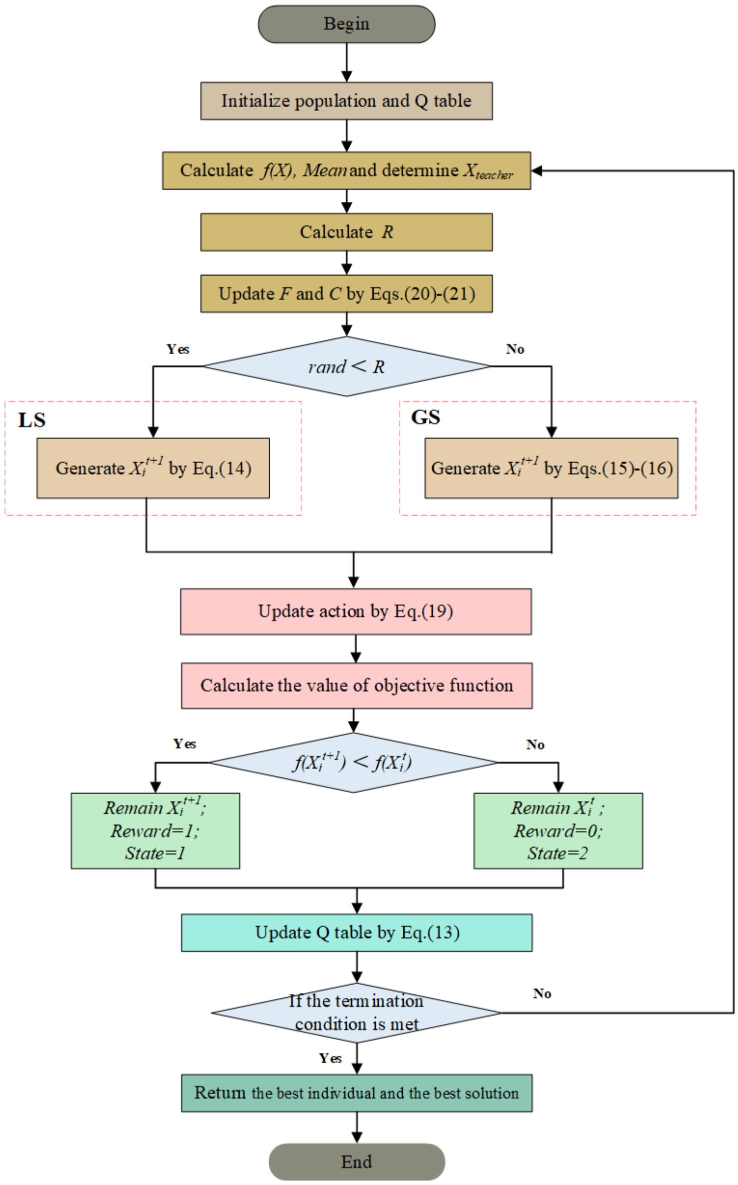


Figure 6. Flow diagram of HHARL.

4. Results and discussion

The PV models studied by SDM, DDM and three PV panel module models were the R.T.C France cell, Photowatt-PWP201 module, mono-crystalline STM6-40/36 module, and poly-crystalline STP6-120/36 module. The performance parameters of these cells are summarized in Table 2. Data for SDM and DDM were obtained from [17]. Data for Photowatt-PWP201, a poly-crystalline silicon cell model consisting of 36 series cells, were measured at an external irradiance of 1000 W/m^2 and a temperature of 45 °C. Data for STM6-40/36 and STP6-120/36 were also measured under an external irradiance of 1000 W/m^2 and temperatures of 51 and 55 °C, respectively [17,18]. These measured data, provided by the manufacturer, were considered as one of the criteria for evaluating the accuracy of photovoltaic parameter extraction. Table 3 summarizes feature descriptions of the five models.

The equivalent circuit models and objective function of photovoltaic cells were introduced in the previous paper. As observed, I-V curve data for PV cells are first collected in experiments using appropriate measurement devices and instruments by the manufacturer. Then, RMSE is calculated according to Eq. (12) using the data obtained from the PV models. Next, according to the optimization process of HHARL, the parameters of PV models are obtained. The purpose of HHARL is to obtain a smaller RMSE by iteratively approximating the best of the identification parameters. Finally, the results of extracted parameters are analyzed and compared with parameters identified by other algorithms. In order to compare the fairness of the results, the parameters use the same boundary constraints. These boundary values are all derived from previous work [19], and their selection is reliable and convincing, as summarized in Table 4. In the numerical experiment, the values of final recognition parameters are the result of running each model 30 times, with evaluation times of each algorithm on the objective function being 30,000. In this paper, the main criterion for evaluating the effectiveness of parameter identification of each algorithm is the minimum value of RMSE ($RMSE_{min}$). The smaller the $RMSE_{min}$, the more accurate the parameter identification by the optimizers, and the more reliable is the algorithm.

Table 2. Performance parameters of each model [17,18].

Parameter	R.T.C France cell	Photowatt-PWP201	STM6-40/36	STP6-120/36
Type	Poly-crystalline	Poly-crystalline	Mono-crystalline	Poly-crystalline
N_s	1	36	36	36
N_p	1	1	1	1
Measurement environment	1000 W/m^2, 33 ℃	1000 W/m^2, 45 ℃	1000 W/m^2, 51 ℃	1000 W/m^2, 55 ℃

Table 3. Feature descriptions of each model [17,18].

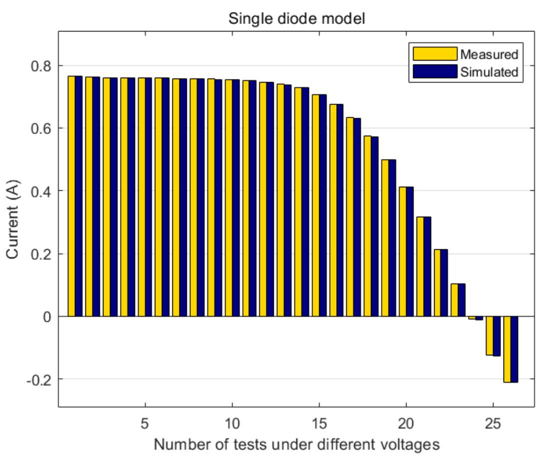
PV model	Feature descriptions
SDM, DDM	57 mm diameter commercial R.T.C. France silicon solar cell
Photowatt-PWP201	No external dimensions are provided by the manufacturer
STM6-40/36	External dimensions (mm): 670×420×25
STP6-120/36	External dimensions (mm):1473×670×30

Table 4. Search scope of the parameters [20].

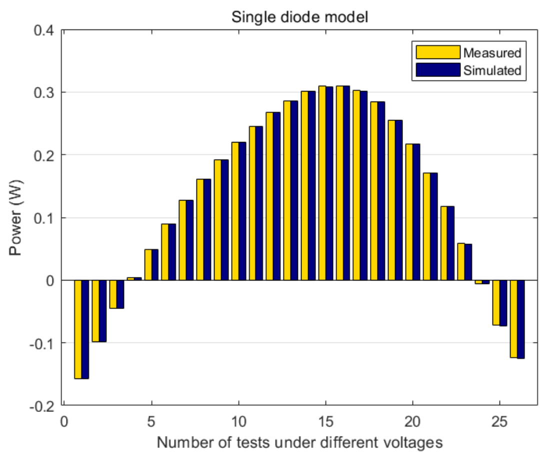
Parameter	R.T.C France cell		Photowatt-PWP201		STM6-40/36		STP6-120/36	
	LB	UB	LB	UB	LB	UB	LB	UB
I_{ph} (A)	0	1	0	2	0	2	0	8
I_{sd1}, I_{sd2} (μA)	0	1	0	50	0	50	0	50
R_s (Ω)	0	0.5	0	2	0	0.36	0	0.36
R_{sh} (Ω)	0	100	0	2000	0	1000	0	1500
n_1, n_2	1	2	1	50	1	60	1	50

4.1. PV parameters identified by HHARL

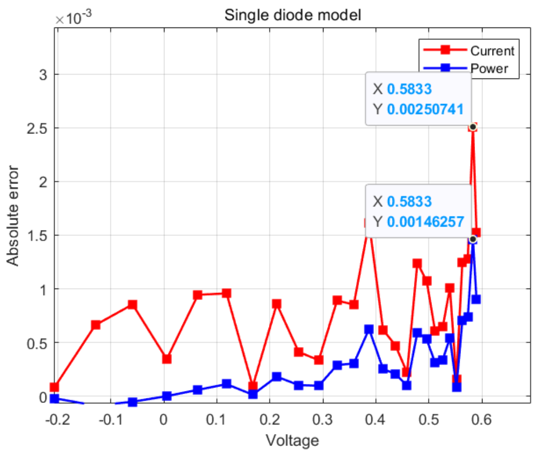
The parameters identified by HHARL regarding the five PV models are summarized in Table 5. Namely, the $RMSE_{min}$ of SDM, DDM, Photowatt-PWP-201, STM6-40/36, and STP6-120/36 are 9.86E-04, 9.82E-04, 0.0024, 0.0017, and 0.0166, respectively. To display the performance of HHARL more intuitively, the *I-V* and *P-V* curves of the five models are shown in Figures 7–11. In the figures, the *X* and *Y* axes represent the voltage and the absolute error value of the current, respectively. As clearly depicted in Figure 7(c) and Figure 8(c), the absolute error of current is 0.0025, and the absolute error of power is 0.0015 in SDM and DDM. Figure 9(c) shows that, in Photowatt-PWP-201, the absolute error of the current does not exceed 0.0048, and the absolute error of the power does not exceed 0.0799. As shown in Figure 10(c), in STM6-40/36, the absolute value of the maximum current error is 0.0061, and the absolute value of the maximum power error is 0.0906. In STP6-120/36, the maximum absolute error of the current is 0.0443, and the absolute error of the power is 0.7126, while the error of most simulation data and experimental data is not higher than 0.4, as seen in Figure 11(c). The above data show that the parameters identified by HHARL in these five models are accurate and reliable.



(a) SDM-IV



(b) SDM-PV

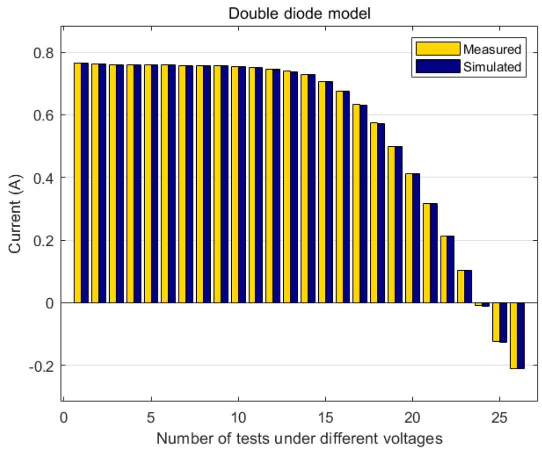


(c) SDM-AE

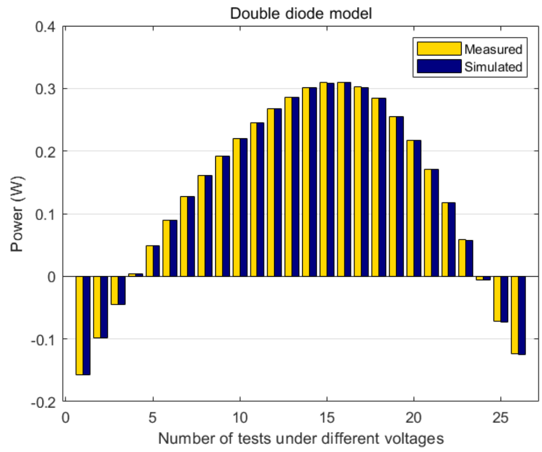
Figure 7. Comparison of experimental data and simulation data regarding SDM by HHARL.

Table 5. Parameters identified by HHARL.

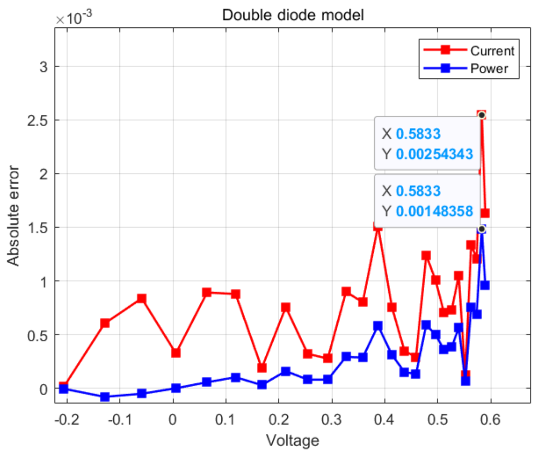
Model	I_{ph}(A)	I_{sd1}(μA)	$R_s(\Omega)$	$R_{sh}(\Omega)$	n_1	I_{sd2}(μA)	n_2	$RMSE_{min}$
SDM	0.7608	0.3230	0.0364	53.7185	1.4812	/	/	9.86E-04
DDM	0.7608	0.7494	0.0367	55.4854	2.0000	0.2260	1.4510	9.825E-04
Photowatt-PWP-201	1.0305	3.4823	1.2013	981.9823	48.6428	/	/	0.0024
STM6-40/36	1.6639	1.7387	0.0043	15.9283	1.5203	/	/	0.0017
STP6-120/36	7.4725	2.3350	0.0046	22.2199	1.2601	/	/	0.0166



(a) DDM-IV

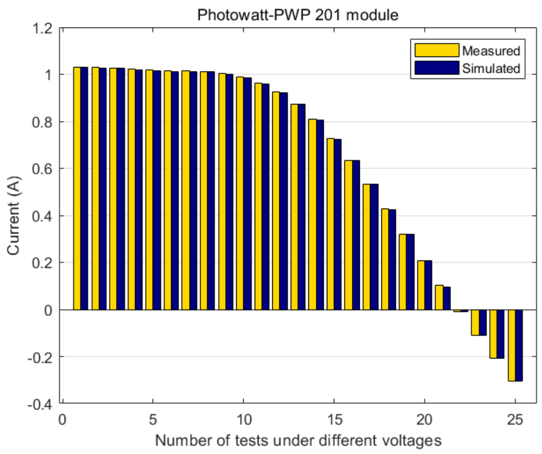


(b) DDM-PV

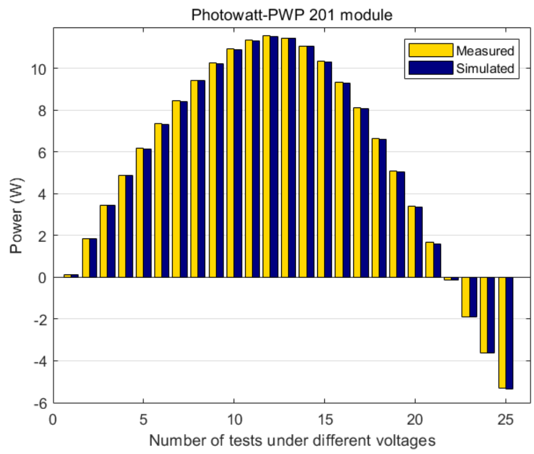


(c) DDM-AE

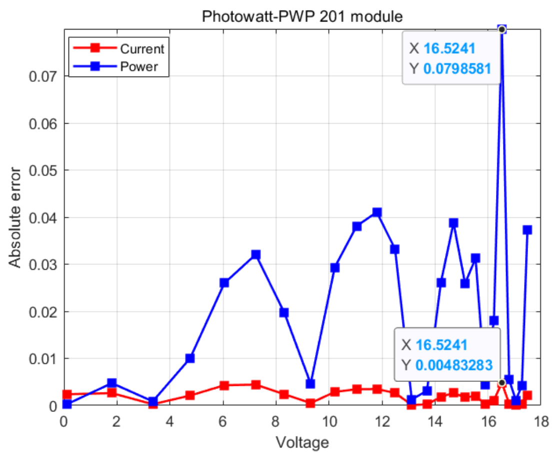
Figure 8. Comparison of experimental data and simulation data regarding DDM by HHARL.



(a) Photowatt-PWP-201-IV

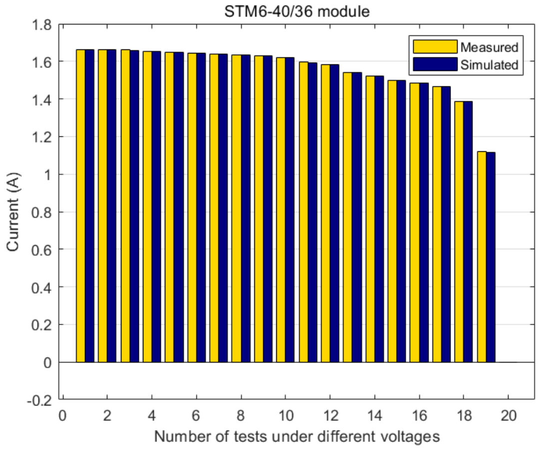


(b) Photowatt-PWP-201-PV

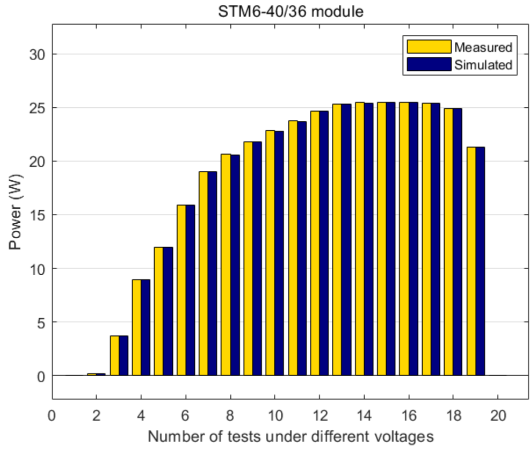


(c) Photowatt-PWP-201-AE

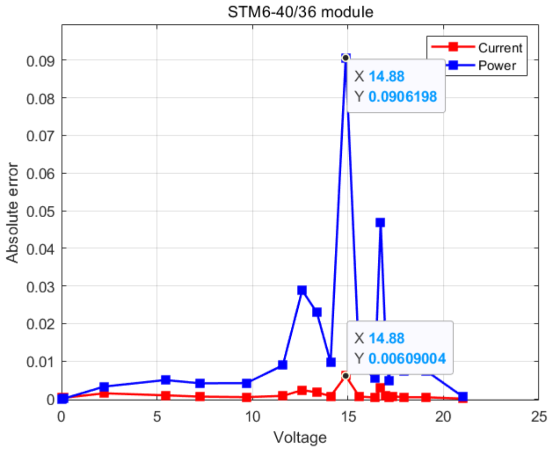
Figure 9. Comparison of experimental data and simulation data regarding Photowatt-PWP-201 by HHARL.



(a) STM6-40/36-IV

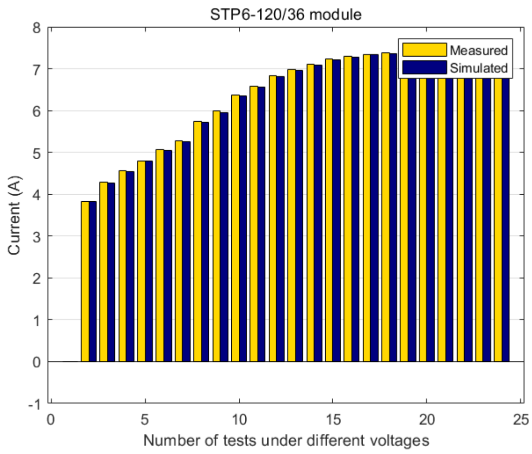


(b) STM6-40/36-PV

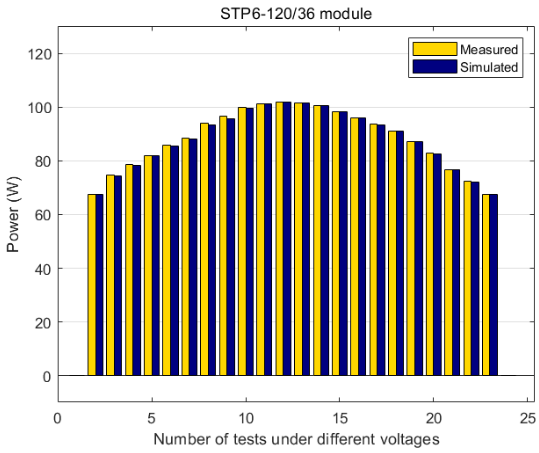


(c) STM6-40/36-AE

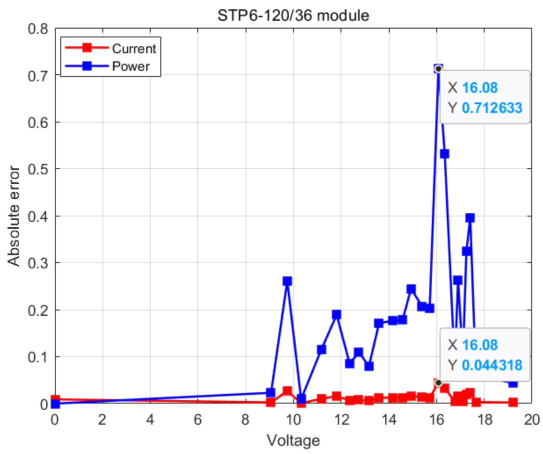
Figure 10. Comparison of experimental data and simulation data regarding STM6-40/36 by HHARL.



(a) STP6-120/36-IV



(b) STP6-120/36-PV



(c) STP6-120/36-AE

Figure 11. Comparison of experimental data and simulation data regarding STP6-120/36 by HHARL.

4.2. Comparison with advanced correlation algorithms

There are many other improved algorithms related to HHARL that perform well in identifying PV model parameters. As such, this paper provides a comparison with the following advanced algorithms: TLBO [21], hybrid adaptive teaching-learning-based optimization and differential evolution (TLBODE) [3], improved teaching-learning-based optimization (ITLBO) [20], differential evolution with reinforcement learning (RLDE) [10], and improved teaching-learning-based optimization algorithm with reinforcement learning (IRLTLBO) [22]. The parameters used by each algorithm in the experiment are summarized in Table 6.

Table 6. Parameters of each algorithm.

Algorithm	Parameter
TLBO	NP=50
TLBODE	NP=50, F=0.5
ITLBO	NP=50
RLDE	NP=50, CR=0.9
IRLTLBO	NP=50
HHARL	NP=50

Table 7 summarizes the four statistical indicators of $RMSE$. In SDM, with the exception of IRLTLBO, all algorithms obtained the same optimal value ($RMSE_{min}$ = 0.0009860), which indicates that these algorithms are able to identify parameters precisely. ITLBO, RLDE, and HHARL obtained the minimum $RMSE_{mean}$ (0.0009860) and the minimum $RMSE_{max}$ (0.0009860). As a result, the parameters obtained through the utilization of these three algorithms exhibit a higher degree of proximity to the measured data in comparison to the remaining algorithms. In addition, ITLBO obtained the smallest $RMSE_{std}$ (2.1533E-17), which indicates that it has the most stable performance, with HHARL ($RMSE_{std}$ = 2.7823E-17) coming in second and RLDE ($RMSE_{std}$ = 1.5967E-16) coming in third. In summary, ITLBO, RLDE, and HHARL are the most efficient algorithms for parameter identification.

In DDM, RLDE and HHARL obtained optimal values ($RMSE_{min}$= 0.0009825), which indicates that they can identify parameters more accurately compared to the other four algorithms. On top of that, HHARL also obtained the smallest $RMSE_{mean}$ (0.0009847) and $RMSE_{max}$ (0.0009851), which means that it has the smallest maximum error among compared algorithms, proving that it is the most effective among these six algorithms. In terms of stability, TLBODE obtained the smallest $RMSE_{std}$ (1.3404E-06), followed by HHARL (1.4055E-06) and ITLBO (1.4283E-06). In summary, HHARL is able to perform parameter identification closely to the experimental data.

In Photowatt-PWP-201, ITLBO, RLDE, and HHARL all obtained the best $RMSE_{min}$ (0.0024251), $RMSE_{mean}$ (0.0024251), and $RMSE_{max}$ (0.0024251), being able to extract more accurate parameters compared to the other algorithms. HHARL also obtained the best $RMSE_{std}$ (1.6653E-17), which indicates that it is not only the most powerful but also the most stable among the six algorithms in terms of finding the best performance.

In STM6-40/36, both HHARL and RLDE obtained the best $RMSE_{min}$ (0.0017298), $RMSE_{mean}$ (0.0017298), and $RMSE_{max}$ (0.0017298). In addition, TLBODE and ITLBO also obtained the same $RMSE_{min}$. In terms of the accuracy of parameter extraction, these are the best among the six algorithms. HHARL also obtained the best $RMSE_{std}$ (1.0551E-17), which demonstrates a substantial reduction when compared to the other five algorithms, followed by RLDE (5.2907E-17) and ITLBO (1.0563E-06). Thus, HHARL is the algorithm that can identify parameters most efficiently and most stably among compared algorithms, followed by RLDE and ITLBO.

In STP6-120/36, both RLDE and HHARL obtained the best $RMSE_{min}$ (0.0166006), $RMSE_{mean}$(0.0166006), and $RMSE_{max}$(0.016606). In addition, TLBODE also obtained the best $RMSE_{min}$ (0.016606), which indicates that RLDE and HHARL performed the best in parameter identification accuracy, followed by TLBODE. On this basis, RLDE obtained the best $RMSE_{std}$ (4.3629E-16), followed by HHARL (1.0865E-15) and TLBODE (5.4509E-05). In summary, RLDE and HHARL proved to be competing algorithms.

Table 7. Comparison between HHARL and other advanced correlation algorithms on eigenvalues of *RMSE*.

Model	Algorithm	*RMSE*			
		Min	Mean	Max	Std
SDM	TLBO	0.0009860	0.0010116	0.0011559	4.3616E-05
	TLBODE	0.0009860	0.0009867	0.0009980	2.6786E-06
	ITLBO	0.0009860	0.0009860	0.0009860	2.1553E-17
	RLDE	0.0009860	0.0009860	0.0009860	1.5967E-16
	IRLTLBO	0.0009928	0.0014096	0.0019318	2.8870E-04
	HHARL	0.0009860	0.0009860	0.0009860	2.7823E-17
DDM	TLBO	0.0009827	0.0012222	0.0023726	3.8201E-04
	TLBODE	0.0009826	0.0009857	0.0009910	1.3404E-06
	ITLBO	0.0009827	0.0009850	0.0009890	1.4283E-06
	RLDE	0.0009825	0.0009849	0.0009875	1.7580E-06
	IRLTLBO	0.0012586	0.0017865	0.0029082	3.9994E-04
	HHARL	0.0009825	0.0009847	0.0009851	1.4055E-06
Photowatt-PWP-201	TLBO	0.0024251	0.0024368	0.0026383	3.9996E-05
	TLBODE	0.0024251	0.0024254	0.0024348	1.7825E-06
	ITLBO	0.0024251	0.0024251	0.0024251	1.7671E-17
	RLDE	0.0024251	0.0024251	0.0024251	1.8989E-16
	IRLTLBO	0.0024255	0.0024724	0.0025545	3.6163E-05
	HHARL	0.0024251	0.0024251	0.0024251	1.6653E-17
STM6-40/36	TLBO	0.0017450	0.0075361	0.0683550	1.2846E-02
	TLBODE	0.0017298	0.0017929	0.0021545	9.9189E-05
	ITLBO	0.0017298	0.0017300	0.0017356	1.0563E-06
	RLDE	0.0017298	0.0017298	0.0017298	5.2907E-17
	IRLTLBO	0.0032877	0.0258860	0.0975421	2.8873E-02
	HHARL	0.0017298	0.0017298	0.0017298	1.0551E-17
STP6-120/36	TLBO	0.0166120	0.0412930	0.4987900	8.6945E-02
	TLBODE	0.0166006	0.0166301	0.0167810	5.4509E-05
	ITLBO	0.0166010	0.0166230	0.0169060	6.3914E-05
	RLDE	0.0166006	0.0166006	0.0166006	4.3629E-16
	IRLTLBO	0.0279416	0.0966340	1.2695257	2.2873E-01
	HHARL	0.0166006	0.0166006	0.0166006	1.0865E-15

4.3. Comparison with common optimization algorithms

Many common nature-inspired intelligent algorithms are also able to identify PV model parameters accurately. This subsection compares the parameters identified by particle swarm optimization (PSO) algorithm [15], grey wolf optimizer (GWO) [14], DE [7], gravitational search algorithm (GSA) [23], and artificial bee colony (ABC) algorithm [24] with HHARL. Tables 9–11 summarize the parameter results identified by each algorithm in the five PV models and the four characteristic values of *RMSE*. The parameters used by each algorithm in the experiment are summarized in Table 8.

Table 8. Parameter setting of each algorithm.

Algorithm	Parameter
PSO	NP=50, $w_{\min}$=0.5, $w_{\max}$=0.9
GWO	NP=50
DE	NP=50, F=0.5, CR=0.8
GSA	NP=50
ABC	NP=50, FoodNumber=25
HHARL	NP=50

Table 9 shows that, in SDM, both HHARL and DE obtained the best $RMSE_{min}$ (0.000986), while HHARL also obtained the best $RMSE_{mean}$ (0.000986) and $RMSE_{max}$ (0.000986); as such, HHARL extracted the most accurate parameters, followed by DE ($RMSE_{mean}$ = 0.001055, $RMSE_{max}$ = 0.001414) and PSO ($RMSE_{min}$ = 0.001006, $RMSE_{mean}$ = 0.001535, $RMSE_{max}$ = 0.002128). HHARL obtained the best $RMSE_{std}$ (2.78E-17), much smaller than those obtained by other algorithms, so it proved to be the most stable, followed by DE (9.64E-05) and PSO (0.000341).

Table 9. Comparison between HHARL and common optimization algorithms on SDM.

Algorithm	I_{ph}(A)	I_{sd}(μA)	R_s(Ω)	R_{sh}(Ω)	n	$RMSE$			
						Min	Mean	Max	Std
PSO	0.7604	0.7320	0.0329	93.8257	1.5683	0.001006	0.001535	0.002128	0.000341
GWO	0.7717	0.9593	0.0245	12.3928	1.6053	0.001345	0.005915	0.038261	0.008888
DE	0.7606	0.4406	0.0351	63.9155	1.5131	0.000986	0.001055	0.001414	9.64E-05
GSA	0.7310	0.6899	0.0146	43.6258	1.5662	0.006716	0.029845	0.050575	0.011234
ABC	0.7798	0.7144	0.0316	18.2214	1.5646	0.007919	0.035132	0.090773	0.019560
HHARL	0.7608	0.3230	0.0364	53.7185	1.4812	0.000986	0.000986	0.000986	2.78E-17

Table 10 shows that, in DDM, HHARL obtained the best $RMSE_{min}$ (0.000982), followed by DE (0.000983) and PSO (0.000984), although the difference between the top three was relatively small. HHARL obtained both the best $RMSE_{mean}$ (0.000985) and $RMSE_{max}$ (0.000986), followed by DE ($RMSE_{mean}$ = 0.001002, $RMSE_{max}$ = 0.001225) and PSO ($RMSE_{mean}$ = 0.001743, $RMSE_{max}$ =0.002506). Therefore, in parameter extraction accuracy ranking, HHARL comes first, followed by DE and PSO. In addition, HHARL also obtained a far better $RMSE_{std}$ (1.14E-06), followed by DE (4.66E-05) and PSO (0.000466), showing that HHARL is the most stable.

Table 10. Comparison between HHARL and common optimization algorithms on DDM.

Algorithm	I_{ph}(A)	I_{sd1}(μA)	R_s(Ω)	R_{sh}(Ω)	n_1	I_{sd2}(μA)	n_2	$RMSE$			
								Min	Mean	Max	Std
PSO	0.7604	0.5299	0.0340	78.4145	1.5344	0.1309	1.9054	0.000984	0.001743	0.002506	0.000466
GWO	0.7709	0.0737	0.0248	13.7185	1.5677	0.9449	1.6159	0.001236	0.009023	0.040950	0.011658
DE	0.7606	0.2378	0.0347	64.7110	1.5484	0.2464	1.5039	0.000983	0.001002	0.001225	4.66E-05
GSA	0.7416	6.7849	0.0213	49.1814	1.6050	0.6613	1.1782	0.011807	0.027881	0.052164	0.010803
ABC	0.7628	0.0125	0.0250	58.4838	1.5677	0.9842	1.6052	0.006646	0.031299	0.071819	0.014580
HHARL	0.7608	0.7494	0.0367	55.4854	2.0000	0.2260	1.4510	0.000982	0.000985	0.000986	1.41E-06

Table 11 shows that, in the three PV panel module models, HHARL obtained the best $RMSE_{min}$ (0.002425, 0.001730, 0.016601) for the Photowatt-PWP-201, STM6-40/36, and STP6-120/36 models, respectively. In Photowatt-PWP-201, DE and HHARL obtained the same $RMSE_{min}$ (0.002425), and PSO (0.002427) and GWO (0.002612) obtained the second and third, respectively. In STM6-40/36, DE (0.002350) and PSO (0.006322) obtained the second and third. In STP6-120/36, GWO (0.018770) and DE (0.019871) obtained the second and third. HHARL obtained the best $RMSE_{mean}$ (0.002425, 0.001730, 0.016601). In Photowatt-PWP-201, DE obtained the same $RMSE_{mean}$ as HHARL (0.002425), while PSO (0.002995) and GWO (0.003625) obtained the second and third, respectively. In STM6-40/36, DE (0.002350) and PSO (0.006322) obtained the second and third. In STP6-120/36, GWO (0.018770) and DE (0.019871) were the second and third. In terms of $RMSE_{max}$, HHARL continued to obtain the best $RMSE_{max}$ (0.002425, 0.001730, 0.016601). In Photowatt-PWP-201, DE (0.002435) and PSO (0.006790) were the second and third, respectively. In STM6-40/36 and STP6-120/36, DE (0.003356, 0.138950) and GWO (0.310800, 0.282620) were second and third. Therefore, HHARL is able to identify parameters more accurately compared to other algorithms. In terms of $RMSE_{std}$, HHARL also obtained the best $RMSE_{std}$ (1.67E-17, 1.06E-17, 1.09E-15). In Photowatt-PWP-201, DE (1.76E-06) and PSO (0.001060) obtained the second and third. In STM6-40/36, DE (0.000267) and GSA (0.001528) obtained the second and third. In STP6-120/36, DE (0.020359) and GWO (0.76263) obtained the second and third. Therefore, HHARL exhibits a higher level of performance stability when compared to the other algorithms.

Table 11. Comparison between HHARL and common optimization algorithms on PV panel module models.

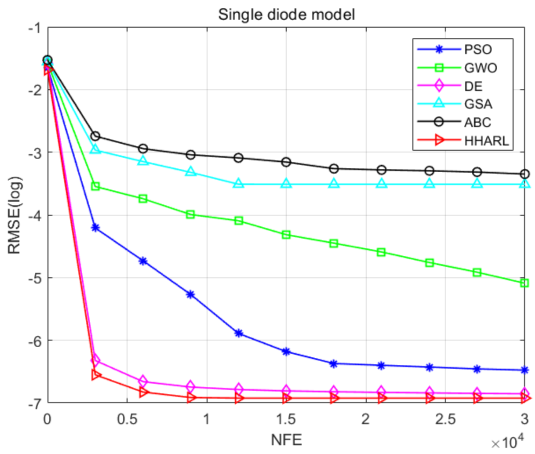
Model	Algorithm	I_{ph}(A)	I_{sd}(μA)	R_s(Ω)	R_{sh}(Ω)	n	$RMSE$			
							Min	Mean	Max	Std
Photowatt-PWP-201	PSO	1.0286	4.0270	1.1884	1.37E+03	49.2008	0.002427	0.002995	0.006790	0.001060
	GWO	1.0371	4.8662	1.1464	614.2467	49.9950	0.002612	0.003625	0.008283	0.001548
	DE	1.0305	3.4823	1.2013	981.9822	48.6428	0.002425	0.002425	0.002435	1.76E-06
	GSA	1.1100	1.7115	0.7568	966.2868	47.1528	0.014806	1.216800	7.962900	2.062600

Continued on next page

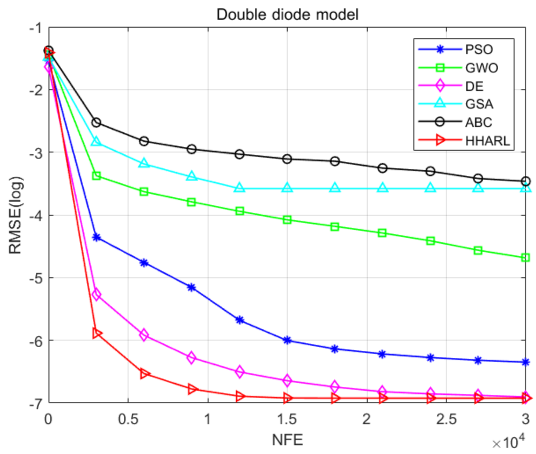
Model	Algorithm	I_{ph}(A)	I_{sd}(μA)	R_s(Ω)	R_{sh}(Ω)	n	$RMSE$			
							Min	Mean	Max	Std
Photowatt-PWP-201	ABC	1.0184	2.8958	0.9896	1.84E+03	48.0643	0.011660	0.046286	0.090023	0.025312
	HHARL	1.0305	3.4823	1.2013	981.9823	48.6428	0.002425	0.002425	0.002425	1.67E-17
STM6-40/36	PSO	1.6608	6.6694	0.0015	460.0947	1.6817	0.006322	0.049561	0.314900	0.072418
	GWO	1.6771	1.4503	0.0020	7.8853	1.5017	0.006604	0.053510	0.310800	0.102760
	DE	1.6628	3.2091	0.0021	19.0510	1.5908	0.002350	0.002698	0.003356	0.000267
	GSA	1.4882	9.0723	0.1740	132.5741	15.1763	0.354600	0.362600	0.362990	0.001528
	ABC	1.6260	3.8587	0.0143	3.5787	20.0067	0.114110	0.294060	0.360110	0.074962
	HHARL	1.6639	1.7387	0.0043	15.9283	1.5203	0.001730	0.001730	0.001730	1.06E-17
STP6-120/36	PSO	7.5755	6.8583	0.0047	134.5487	1.3673	0.099349	0.507660	1.662200	0.523960
	GWO	7.4522	1.6834	0.0048	734.4850	1.2330	0.018770	0.073987	0.282620	0.076263
	DE	7.4869	7.8872	0.0039	984.3619	1.3708	0.019871	0.034766	0.138950	0.020359
	GSA	6.1341	47.5621	0.0005	776.0472	25.4692	1.607600	1.711800	1.715400	0.019676
	ABC	6.1341	29.6721	0.0123	313.8091	23.6143	0.250480	1.585900	1.715600	0.291870
	HHARL	7.4725	2.3350	0.0046	22.2199	1.2601	0.016601	0.016601	0.016601	1.09E-15

To clearly show the changing trends of these algorithms with the iteration progress, the convergence curves of these algorithms applied to five models are summarized in Figure 12. Figure 12(a) shows that DE and HHARL converge earlier than the other four algorithms in SDM. Figure 12(b) shows that HHARL is the first to reach convergence in DDM. As can be seen from Figure 12(c), DE and HHARL converge significantly faster than the other four algorithms in Photowatt-PWP-201. In STM6-40/36 and STP6-120/36, ABC and GSA converge prematurely, and HHARL has the fastest rate of convergence among other comparison algorithms, as seen from Figure 12(d)–(e). Overall, HHARL is the fastest at parameter identification.

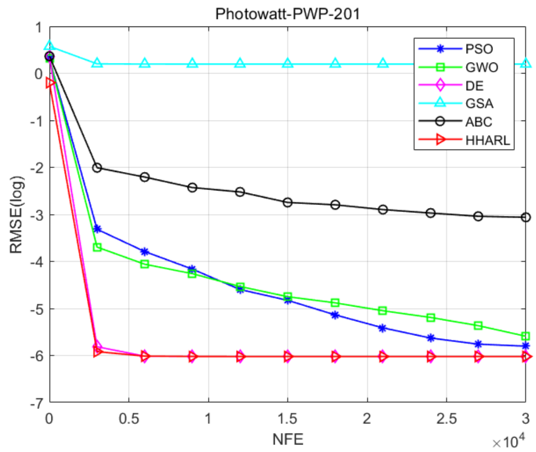
Overall, HHARL consistently outperforms the competing algorithms in terms of parameter identification accuracy, stability, and convergence speed across both the SDM and DDM. DE and PSO rank second and third, respectively, indicating their relatively competitive optimization performance. For the Photowatt-PWP-201 dataset, HHARL achieves the best performance across all evaluated criteria, followed by DE and PSO. For the STM6-40/36 and STP6-120/36 datasets, HHARL also demonstrates superior parameter identification accuracy and stability compared with the other algorithms, with DE and GWO ranking second and third, respectively. Furthermore, the results of the Friedman statistical tests, presented in Figure 13, consistently rank HHARL first among the evaluated algorithms, providing statistical evidence of its overall superiority. These results demonstrate that HHARL possesses strong optimization capability and robustness, enabling accurate, stable, and efficient parameter identification for photovoltaic models.



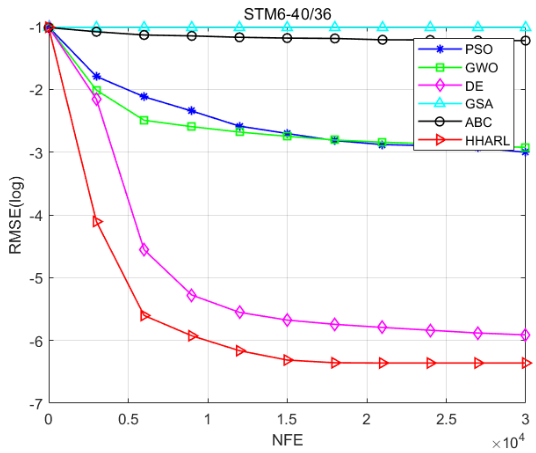
(a) SDM



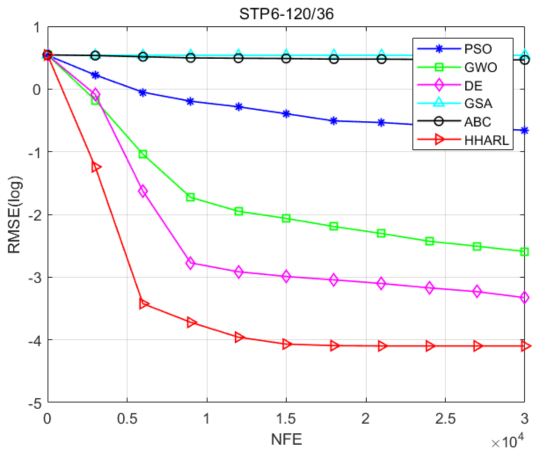
(b) DDM



(c) Photowatt-PWP-201



(d) STM6-40/36



(e) STP6-120/36

Figure 12. Convergence curves of common optimization algorithms for PV models: (a) SDM; (b) DDM; (c) Photowatt-PWP-201; (d) STM6-40/36; (e) STP6-120/36.

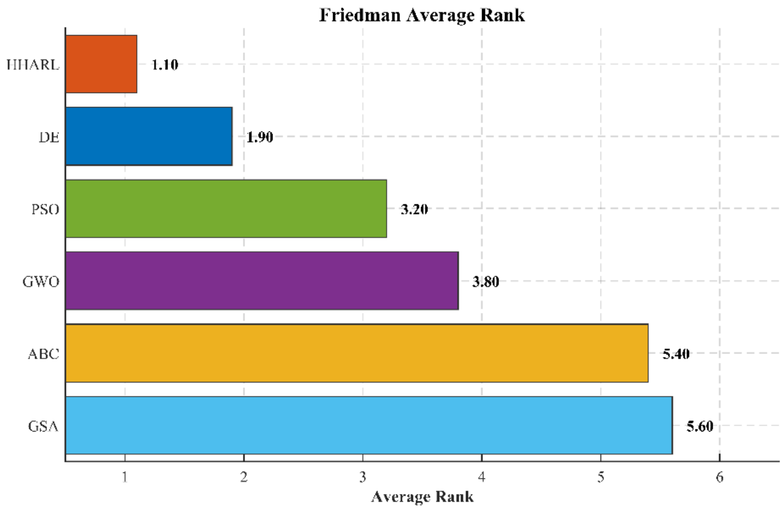


Figure 13. Friedman average rank results.

4.4. Comparison with other results published in the literature

To further demonstrate the efficacy of HHARL in identifying solar system parameters, HHARL is compared with conclusions drawn by other intelligent algorithms published in recent years. Tables 12–15 summarize the parameters of five models extracted by each algorithm, as well as $RMSE_{min}$ and the number of fitness evaluations (NFE).

Table 12 shows that almost all algorithms obtained $RMSE_{min}$=9.8602E-04, and the I_{ph} identified by these algorithms are all the same; the other four parameter values are also very close. However, compared with other algorithms, the NFE of HHARL and OLBGWO is 30,000, so they consumed the least resources when converging to the optimal solution. Hence, HHARL shows reliability and accuracy in SDM parameter identification.

Table 12. Comparison of HHARL with other results published in the literature on SDM.

Algorithm	I_{ph}(A)	I_{sd}(μA)	R_s(Ω)	R_{sh}(Ω)	n	$RMSE_{min}$	NFE
SATLBO [25]	0.7608	0.3232	0.0364	53.7526	1.4812	9.8602E-04	50000
IJAYA [8]	0.7608	0.3228	0.0364	53.7595	1.4811	9.8603E-04	50000
DE/WOA [26]	0.7608	0.3230	0.0364	53.7185	1.4812	9.8602E-04	50000
OBWOA [27]	0.7608	0.3232	0.0363	53.6836	1.5208	9.8602E-04	1500000
TLABC [28]	0.7608	0.3230	0.0364	53.7167	1.4812	9.8602E-04	50000
PGJAYA [18]	0.7608	0.3230	0.0364	53.7185	1.4812	9.8602E-04	50000
EGBO [29]	0.7608	0.3230	0.0364	53.7185	1.4811	9.8602E-04	200000
EABOA [30]	0.7608	0.3230	0.0364	53.7660	1.4811	9.8602E-04	50000
OLBGWO [31]	0.7608	0.3230	0.0364	53.7188	1.4819	9.8602E-04	30000
DOLADE [19]	0.7608	0.3230	0.0364	53.7185	1.4819	9.8602E-04	50000
PSMDE[32]	0.7608	0.3230	0.0364	53.7185	1.4812	9.8602E-04	50000
HHARL	0.7608	0.3230	0.0364	53.7185	1.4812	9.8602E-04	30000

Table 13 shows that EABOA obtained the worst $RMSE_{min}$= 9.8607E-04. DE/WOA, CBSA, EGBO, DOLADE, and HHARL obtained the best $RMSE_{min}$=9.8248E-04. Although OLBGWO (9.8256E-04), MRFO (9.8443E-04), and HHARL (9.8248E-04) use the same *NFE*, the $RMSE_{min}$ of HHARL is the smallest, so HHARL is more efficient at identifying parameters. Therefore, HHARL is also effective in DDM parameter identification.

Table 13. Comparison of HHARL with other results published in the literature on DDM.

Algorithm	I_{ph}(A)	I_{sd1}(μA)	R_s(Ω)	R_{sh}(Ω)	n_1	I_{sd2}(μA)	n_2	$RMSE_{min}$	*NFE*
SATLBO [25]	0.7608	0.2509	0.0366	55.1170	1.4598	0.5454	1.9994	9.8280E-04	50000
IJAYA [8]	0.7601	0.0050	0.0376	77.8519	1.2186	0.7590	1.6247	9.8293E-04	50000
DE/WOA [26]	0.7608	0.2260	0.0367	55.4854	1.4510	0.7493	2.0000	9.8248E-04	50000
OBWOA [27]	0.7607	0.2299	0.0367	55.3990	1.4915	0.7493	2.0000	9.8251E-04	1500000
TLABC [28]	0.7608	0.4239	0.0367	54.6680	1.9075	0.2401	1.4567	9.8415E-04	50000
PGJAYA [18]	0.7608	0.2103	0.0368	55.8135	1.4450	0.8853	2.0000	9.8263E-04	50000
EGBO [29]	0.7608	0.2500	0.0367	55.4855	1.4510	0.7490	2.0000	9.8248E-04	200000
EABOA [30]	0.7608	0.2507	0.0366	55.3660	1.4599	0.7207	2.0000	9.8607E-04	50000
OLBGWO [31]	0.7608	0.2260	0.0367	55.3078	1.4513	0.6432	1.9618	9.8256E-04	30000
DOLADE [19]	0.7608	0.2260	0.0367	55.4854	1.4510	0.7493	2.0000	9.8248E-04	50000
MRFO[4]	0.7608	0.2135	0.0365	54.1795	1.9913	0.2924	1.4727	9.8443E-04	30000
HHARL	0.7608	0.7494	0.0367	55.4854	2.0000	0.2260	1.4510	9.8248E-04	30000

Table 14 shows that, in Photowatt-PWP-201, most algorithms obtained the same $RMSE_{min}$=2.425075E-03, and the five parameters identified had little differences in value, so the parameters identified by HHARL were accurate and effective. In STM6-40/36, BHCS, CPMPSO, EGBO, and HHARL got the best $RMSE_{min}$=1.7298E-03, and the parameters identified by these algorithms also presented little differences, so the parameters identified by HHARL can be regarded as valid. In STP6-120/36, CWOA obtained the worst $RMSE_{min}$=1.7601E-02, while the other algorithms obtained $RMSE_{min}$=1.6601E-02, and the values of the parameters identified were close, proving that the parameters identified by HHARL were accurate. Moreover, while obtaining the same optimal solution, the *NFE* of HHARL was smaller, which shows that it can use computing resources more efficiently and find high-quality solutions faster.

Table 14. Comparison of HHARL with other results published in the literature on PV panel module models.

Model	Algorithm	$I_{ph}(A)$	$I_{sd}(\mu A)$	R_s(Ω)	R_{sh}(Ω)	n	$RMSE_{min}$	*NFE*
Photowatt-PWP-201	PGJAYA[18]	1.0305	3.4818	1.2013	981.8545	48.6424	2.425075E-03	50000
	LCJAYA [33]	1.0305	3.4823	1.2024	981.9828	48.6684	2.425075E-03	50000
	SGDE [34]	1.0305	3.4823	1.2013	981.9822	48.6428	2.425075E-03	50000
	EGBO[29]	1.0305	3.4823	1.2013	981.9822	48.6428	2.4251E-03	200000

Continued on next page

Model	Algorithm	$I_{ph}(A)$	$I_{sd}(\mu A)$	$R_s(\Omega)$	$R_{sh}(\Omega)$	n	$RMSE_{min}$	NFE
Photowatt-PWP-201	EABOA[30]	1.0304	3.5084	1.2006	991.9831	48.6713	2.4252E-03	50000
	DOLADE[19]	1.0305	3.4823	1.2013	981.9823	48.6428	2.425075E-03	50000
	MRFO [4]	1.0302	3.9750	0.0330	30.3291	1.3654	2.455927E-03	30000
	HHARL	1.0305	3.4823	1.2013	981.9823	48.6428	2.425075E-03	30000
STM6-40/36	CWOA [35]	1.7000	1.6338	0.0050	15.4000	1.5000	1.8E-03	150000
	OBWOA [27]	1.6642	1.6503	0.0044	15.5299	1.5142	1.753E-03	1500000
	BHCS [36]	1.6639	1.7387	0.0043	15.9283	1.5203	1.7298E-03	50000
	GWOCS [37]	1.6641	1.7449	0.0042	15.7326	1.5207	1.7337E-03	50000
	CPMPSO [38]	1.6639	1.7387	0.0043	15.9286	1.5203	1.7298E-03	50000
	EGBO [29]	1.6639	1.7387	0.0043	15.9283	1.5203	1.7298E-03	200000
	DOLADE [19]	1.6640	2.0000	0.1049	570.2588	55.1860	1.783E-03	50000
	MRFO [4]	1.6638	1.6818	0.0044	16.0572	1.5166	1.741E-03	30000
	HHARL	1.6639	1.7387	0.0043	15.9283	1.5203	1.7298E-03	30000
STP6-120/36	CWOA [35]	7.4760	1.2000	4.9E-06	9.745	1.2072	1.7601E-02	150000
	BHCS [36]	7.4725	2.3350	0.0046	22.2199	1.2601	1.6601E-02	50000
	CPMPSO [38]	7.4725	2.3350	0.0046	22.2199	1.2601	1.66006E-02	50000
	EGBO[29]	7.4725	2.3350	0.0046	22.2199	1.2601	1.6601E-02	200000
	MRFO[4]	7.4625	2.6252	0.0045	122.6983	1.2699	1.6368E-02	30000
	HHARL	7.4725	2.3350	0.0046	22.2199	1.2601	1.66006E-02	30000

4.5. Discussion

4.5.1. The ablation study

HHARL mainly consists of the LS strategy, GS strategy, ranking mechanism, and RL-based parameters. To individually test the contribution of these components to the algorithm, we conducted an ablation study to assess their impact on the experimental results. To ensure fairness, the control variable principle was adopted. In the HHA-LS test, HHA was optimized using only one LS strategy while keeping the other conditions unchanged. The same approach was applied in the HHA-GS testing, HHARL-NR with no ranking, and HHA-FP with fixed parameters. Meanwhile, we designed a simple deterministic adaptation rule to control F and C. When the offspring is better, we increase F and C; otherwise, we decrease F and C. The algorithm is HHA-DA.

The experimental parameters employed are consistent with those utilized by HHARL for identifying DDM, as it has seven parameters, being a little harder than the remaining four models. The population size is 50, the maximum number of evaluations is 30,000, and the final results are obtained after 30 runs. Table 15 shows the optimization results of the ablation study. In terms of the min value, HHARL and HHARL-NR obtained the best results. In terms of mean and max values, HHARL achieved the best results. It can be seen from the test results that HHARL is superior to these variants. The inferior performance of HHA-DA can be mainly attributed to the limited adaptability of its deterministic parameter adjustment rule. Specifically, the fixed adaptation logic cannot adequately respond to the changing evolutionary dynamics, whereas the Q-learning-based mechanism can dynamically select parameter adjustment actions according to the current evolutionary state. The superiority of HHARL can be attributed to the complementary roles of the different components. The LS strategy emphasizes localized exploitation around individual solutions. Conversely, the GS strategy enhances global exploration and population diversity. The ranking mechanism strengthens the GS strategy by directing additional search efforts toward high-quality solutions, thereby improving the exploitation capability. Meanwhile, the adaptive parameter adjustment dynamically regulates key algorithmic parameters according to the evolving search state, allowing HHARL to adapt its search behavior throughout the evolutionary process. The coordinated interaction of these mechanisms enables a more effective exploration–exploitation trade-off, which is essential for achieving improved convergence performance and search robustness.

Table 15. Optimization results of the ablation study on DDM

	RMSE			
	Min	Mean	Max	Std
HHA-LS	0.005155	0.015769	0.045271	0.009998
HHA-GS	0.000984	0.000989	0.001059	1.38E-05
HHARL-NR	0.000983	0.001013	0.001302	7.02E-05
HHA-FP	0.000991	0.001238	0.001699	2.27E-04
HHA-DA	0.000984	0.000997	0.001070	2.17E-05
HHARL	0.000983	0.000985	0.000985	1.41E-06

4.5.2. Computation time evaluation

HHARL performs well in identifying parameters on DDM, so we further discuss its computational time cost. The computing time of the algorithms and the optimal *RMSE* value in identifying DDM parameters are summarized in Table 16. To make the comparison clearer, Figure 14 shows a line graph of the time taken by these algorithms. In terms of time alone, HHARL ranks fourth, while TLBODE takes the shortest time. However, the main purpose of this study is to optimize *RMSE*, so a comprehensive evaluation is needed. Overall, these results demonstrate that HHARL provides a competitive and robust solution for photovoltaic parameter identification, achieving a favorable balance among identification accuracy, convergence performance, stability, and computational efficiency.

Table 16. The computing time (s) and optimal *RMSE* of each algorithm in DDM.

Algorithm	Computing time (s)	$RMSE_{min}$
TLBO	10.0123	0.0009827
TLBODE	9.9806	0.0009826
ITLBO	10.6294	0.0009827
RLDE	20.6520	0.0009825
IRLTLBO	14.2764	0.0012586
HHARL	13.9003	0.0009825

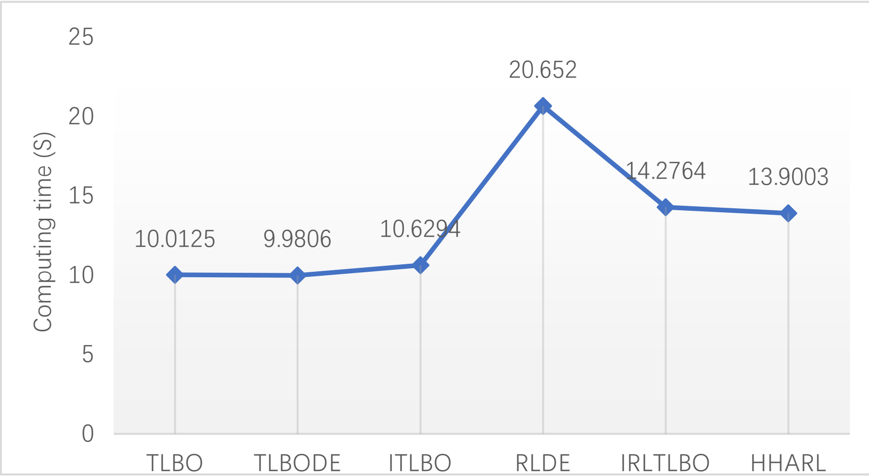


Figure 14. Computing time (s) in DDM.

4.5.3. Discussion

The experimental results show that HHARL performs well in terms of accuracy and efficiency of identifying solar system parameters. First, HHARL leverages the idea of LS and GS to guide individuals to balance exploration and exploitation in the search space. LS is mainly concerned with the search space near the current solution, which allows HHARL to find high-quality solutions in a relatively short time. LS is able to search the solution space more precisely because it focuses on optimizing local regions near the current solution. This allows HHARL to fine-tune and optimize the solution around the current optimal solution to get better results. GS is capable of searching the solution space extensively, trying different solutions, and discovering possible global optimal solutions. This gives HHARL greater exploration power and helps to find potentially better solutions. By exploring multiple regions in the solution space, GS can find more optimal solutions and avoid being trapped by local optimal solutions.

Second, HHARL uses a ranking strategy to select search strategies. This capability enables HHARL to adapt to varying local optimization and global exploration requirements by selecting different search strategies based on the specific regions and characteristics of the problem space.

Third, F and C adjusted by RL are introduced to further balance the exploration and exploitation capabilities of search strategies. Besides, HHARL uses fitness feedback and reward mechanisms to evaluate the behavior of individuals and update the Q table. This fitness feedback mechanism enables individuals to guide their behaviors through reward and punishment according to the optimization goal of problems, which makes individuals more likely to search in the direction of a better solution and expedites the convergence speed of the algorithm.

In conclusion, HHARL balances the exploration and exploitation capabilities of the algorithm with LS and GS strategies selected by the ranking strategy and parameters adjusted by RL technique. HHARL makes use of these advantages to enable individuals to utilize existing knowledge and experience to search more effectively. Therefore, it improves the searchability of the algorithm and the quality of the optimal solution.

5. Conclusion

Accurate identification of PV parameters to optimize the design, operation, and maintenance of solar systems is one of the key measures to increase efficiency when converting solar energy into electricity. Thus, this paper proposes a hybrid heuristic intelligent algorithm based on reinforcement learning (HHARL) to accurately identify PV parameters. In HHARL, local search (LS) and global search (GS) methods are proposed, which use the ranking parameter R to select the appropriate search method for individuals. LS mainly searches for optimal solutions in a local region of the current optimal solution, while GS focuses more on exploring potential global optimal solutions in the entire solution space. To better balance the exploration and exploitation of the algorithm, RL-adjusted parameters are introduced. During parameter adjustment, using the Q-learning algorithm, the agent iteratively optimizes values in the Q table by constantly updating the Q value, and finally determines the optimal action by selecting the highest Q value to obtain the maximum cumulative reward, gradually finding the optimal search strategy.

The great performance of HHARL in parameter identification was verified by comparisons with advanced correlation algorithms, common optimization algorithms, and other optimization algorithms in five PV models. Primarily, HHARL obtained $RMSE_{min}$ values of 9.86E-04, 9.82E-04, 0.0024, 0.0017, and 0.0166 for SDM, DDM, Photowatt-PWP-201, STM6-40/36, and STP6-120/36, respectively. Compared with advanced related algorithms (TLBO, TLBODE, ITLBO, RLDE, and IRLTLBO), HHARL performed the best in accurately identifying PV parameters. At the same time, HHARL also ranked highly in stability among these algorithms. Compared with common optimization algorithms (PSO, GWO, DE, GSA, and ABC), HHARL ranked first in both accuracy and stability. Compared with other results published in the literature, HHARL got the best $RMSE_{min}$ among five PV models with less NFE, and identified parameter values very closely to the other algorithms, which proves its superiority in parameter extraction. Therefore, HHARL can be regarded as a highly efficient approach for accurately identifying parameters of solar systems.

Future research will focus on incorporating more informative performance metrics, such as the relative root mean square error (RRMSE), into the objective function to enhance the comprehensiveness of the optimization criterion. Furthermore, HHARL will be benchmarked against a broader range of state-of-the-art optimization algorithms to provide a more rigorous assessment of its performance, robustness, and generalization capability.

Data availability

The data used in this study are available from the corresponding author upon reasonable request.

Author contributions

Xuming Wang: Methodology, Software, Validation, Writing - review & editing; Wen Zhang: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Visualization, Writing - original draft; Xiaobing Yu: Conceptualization, Funding acquisition, Resources, Supervision, Validation, Writing - review & editing.

Use of Generative-AI tools declaration

During the preparation of this manuscript, the authors used ChatGPT for language polishing and expression improvement. All AI-assisted content was critically reviewed and revised by the authors, who take full responsibility for the final published version.

Acknowledgments

The authors are very grateful to those who provided valuable comments on this article.

Conflict of interest

On behalf of all authors, the corresponding author states that there is no conflict of interest.

References

1. Y. B. Assoa, D. Valencia-Caballero, E. Rico, T. D. Caño, J. V. Furtado, Performance of a large size photovoltaic module for façade integration, *Renew. Energy,* **211** (2023), 903–917. https://doi.org/10.1016/j.renene.2023.04.087
2. F. Nicoletti, M. A. Cucumo, N. Arcuri, Building-integrated photovoltaics (BIPV): A mathematical approach to evaluate the electrical production of solar PV blinds, *Energy,* **263** (2023), 126030. https://doi.org/10.1016/j.energy.2022.126030
3. S. Li, W. Gong, L. Wang, X. Yan, C. Hu, A hybrid adaptive teaching–learning-based optimization and differential evolution for parameter identification of photovoltaic models, *Energ Convers Manage,* **225** (2020), 113474. https://doi.org/10.1016/j.enconman.2020.113474
4. H. Bakır, Elevating PV model performance: Accurate and reliable parameter extraction of solar cell models with state-of-art metaheuristic algorithms, *Next Energy,* **11** (2026), 100513. https://doi.org/10.1016/j.nxener.2026.100513
5. X. Mi, Z. Liao, S. Li, Q. Gu, Adaptive teaching–learning-based optimization with experience learning to identify photovoltaic cell parameters, *Energy Rep,* **7** (2021), 4114–4125. https://doi.org/10.1016/j.egyr.2021.06.097
6. Z. Liao, Z. Chen, S. Li, Parameters extraction of photovoltaic models using triple-phase teaching-learning-based optimization, *IEEE Access,* **8** (2020), 69937–69952. https://doi.org/10.1109/access.2020.2984728
7. R. Storn, K. Price, Differential evolution – A simple and efficient heuristic for global optimization over continuous spaces, *J. Glob. Optim.,* **11** (1997), 341–359. https://doi.org/10.1023/A:1008202821328
8. K. Yu, J. J. Liang, B. Y. Qu, X. Chen, H. Wang, Parameters identification of photovoltaic models using an improved JAYA optimization algorithm, *Energy Convers. Manage.,* **150** (2017), 742–753. https://doi.org/10.1016/j.enconman.2017.08.063
9. F. Hui, S. Li, K. Peng, S. Jing, M. Liu, UMRL: A unified multi-modal reinforcement learning for autonomous driving decision-making at unsignalized intersections, *Inf. Fusion,* (2026), 104721. https://doi.org/10.1016/j.inffus.2026.104721
10. Z. Hu, W. Gong, S. Li, Reinforcement learning-based differential evolution for parameters extraction of photovoltaic models, *Energy Rep.,* **7** (2021), 916–928. https://doi.org/10.1016/j.egyr.2021.01.096
11. A. M. Humada, M. Hojabri, S. Mekhilef, H. M. Hamada, Solar cell parameters extraction based on single and double-diode models: A review, *Renew. Sustain. Energy Rev.,* **56** (2016), 494–509. https://doi.org/10.1016/j.rser.2015.11.051
12. F. J. Toledo, V. Galiano, J. M. Blanes, V. Herranz, E. Batzelis, Photovoltaic single-diode model parametrization. An application to the calculus of the Euclidean distance to an I–V curve, *Math. Comput. Simul.,* **225** (2024), 794–819. https://doi.org/10.1016/j.matcom.2023.01.005
13. J. Dang, G. Wang, C. Xia, R. Jia, P. Li, Research on the parameter identification of PV module based on fuzzy adaptive differential evolution algorithm, *Energy Rep.,* **8** (2022), 12081–12091. https://doi.org/10.1016/j.egyr.2022.09.057
14. S. Mirjalili, S. M. Mirjalili, A. Lewis, Grey wolf optimizer, *Adv. Eng. Softw.,* **69** (2014), 46–61. https://doi.org/10.1016/j.advengsoft.2013.12.007

15. J. Kennedy, R. Eberhart, Particle swarm optimization, *Proc. ICNN'95,* **4** (1995), 1942–1948. https://doi.org/10.1109/ICNN.1995.488968
16. A. Taheri, K. RahimiZadeh, R. V. Rao, An efficient Balanced Teaching-Learning-Based optimization algorithm with Individual restarting strategy for solving global optimization problems, *Inf. Sci.,* **576** (2021), 68–104. https://doi.org/10.1016/j.ins.2021.06.064
17. T. Easwarakhanthan, J. Bottin, I. Bouhouch, C. Boutrit, Nonlinear minimization algorithm for determining the solar cell parameters with microcomputers, *Int. J. Sol. Energy,* **4** (1986), 1–12. https://doi.org/10.1080/01425918608909835
18. K. Yu, B. Qu, C. Yue, S. Ge, X. Chen, J. Liang, A performance-guided JAYA algorithm for parameters identification of photovoltaic cell and module, *Appl. Energy,* **237** (2019), 241–257. https://doi.org/10.1016/j.apenergy.2019.01.008
19. J. Zhou, Y. Zhang, Y. Zhang, W. L. Shang, Z. Yang, W. Feng, Parameters identification of photovoltaic models using a differential evolution algorithm based on elite and obsolete dynamic learning, *Appl. Energy,* **314** (2022), 118877. https://doi.org/10.1016/j.apenergy.2022.118877
20. S. Li, W. Gong, X. Yan, C. Hu, D. Bai, L. Wang, L. Gao, Parameter extraction of photovoltaic models using an improved teaching-learning-based optimization, *Energy Convers. Manage.,* **186** (2019), 293–305. https://doi.org/10.1016/j.enconman.2019.02.048
21. R. V. Rao, V. J. Savsani, D. P. Vakharia, Teaching–learning-based optimization: A novel method for constrained mechanical design optimization problems, *Comput. Aided Des.,* **43** (2011), 303–315. https://doi.org/10.1016/j.cad.2010.12.015
22. D. Wu, S. Wang, Q. Liu, L. Abualigah, H. Jia, N. Razmjooy, An improved teaching-learning-based optimization algorithm with reinforcement learning strategy for solving optimization problems, *Comput. Intell. Neurosci.,* **2022** (2022), 1–24. https://doi.org/10.1155/2022/1535957
23. E. Rashedi, H. Nezamabadi-pour, S. Saryazdi, GSA: A gravitational search algorithm, *Inf. Sci.,* **179** (2009), 2232–2248. https://doi.org/10.1016/j.ins.2009.03.004
24. D. Karaboga, B. Basturk, A powerful and efficient algorithm for numerical function optimization: artificial bee colony (ABC) algorithm, *J. Glob. Optim.,* **39** (2007), 459–471. https://doi.org/10.1007/s10898-007-9149-x
25. K. Yu, X. Chen, X. Wang, Z. Wang, Parameters identification of photovoltaic models using self-adaptive teaching-learning-based optimization, *Energy Convers. Manage.,* **145** (2017), 233–246. https://doi.org/10.1016/j.enconman.2017.04.054
26. G. Xiong, J. Zhang, X. Yuan, D. Shi, Y. He, G. Yao, Parameter extraction of solar photovoltaic models by means of a hybrid differential evolution with whale optimization algorithm, *Sol. Energy,* **176** (2018), 742–761. https://doi.org/10.1016/j.solener.2018.10.050
27. M. A. Elaziz, D. Oliva, Parameter estimation of solar cells diode models by an improved opposition-based whale optimization algorithm, *Energy Convers. Manage.,* **171** (2018), 1843–1859. https://doi.org/10.1016/j.enconman.2018.05.062
28. X. Chen, B. Xu, C. Mei, Y. Ding, K. Li, Teaching–learning–based artificial bee colony for solar photovoltaic parameter estimation, *Appl. Energy,* **212** (2018), 1578–1588. https://doi.org/10.1016/j.apenergy.2017.12.115
29. I. Ahmadianfar, W. Gong, A. A. Heidari, N. A. Golilarz, A. Samadi-Koucheksaraee, H. Chen, Gradient-based optimization with ranking mechanisms for parameter identification of photovoltaic systems, *Energy Rep.,* **7** (2021), 3979–3997. https://doi.org/10.1016/j.egyr.2021.06.064

30. W. Long, T. Wu, M. Xu, M. Tang, S. Cai, Parameters identification of photovoltaic models by using an enhanced adaptive butterfly optimization algorithm, *Energy,* **229** (2021), 120750. https://doi.org/10.1016/j.energy.2021.120750
31. F. J. Xavier, A. Pradeep, M. Premkumar, C. Kumar, Orthogonal learning-based Gray Wolf Optimizer for identifying the uncertain parameters of various photovoltaic models, *Optik,* **247** (2021), 167973. https://doi.org/10.1016/j.ijleo.2021.167973
32. Y. Xiong, Y. Luo, J. Peng, Q. Zhang, S. Huang, An improved population segmentation-based multi-mutation differential evolution algorithm for parameter extraction of photovoltaic models, *Energy Convers. Manage.,* **327** (2025), 119553. https://doi.org/10.1016/j.enconman.2025.119553
33. X. Jian, Z. Weng, A logistic chaotic JAYA algorithm for parameters identification of photovoltaic cell and module models, *Optik,* **203** (2020), 164041. https://doi.org/10.1016/j.ijleo.2019.164041
34. J. Liang, K. Qiao, M. Yuan, K. Yu, B. Qu, S. Ge, et al., Evolutionary multi-task optimization for parameters extraction of photovoltaic models, *Energy Convers. Manage.,* **207** (2020), 112509. https://doi.org/10.1016/j.enconman.2020.112509
35. D. Oliva, M. A. E. Aziz, A. E. Hassanien, Parameter estimation of photovoltaic cells using an improved chaotic whale optimization algorithm, *Appl. Energy,* **200** (2017), 141–154. https://doi.org/10.1016/j.apenergy.2017.05.029
36. X. Chen, K. Yu, Hybridizing cuckoo search algorithm with biogeography-based optimization for estimating photovoltaic model parameters, *Sol. Energy,* **180** (2019), 192–206. https://doi.org/10.1016/j.solener.2019.01.025
37. W. Long, S. Cai, J. Jiao, M. Xu, T. Wu, A new hybrid algorithm based on grey wolf optimizer and cuckoo search for parameter extraction of solar photovoltaic models, *Energy Convers. Manage.,* **203** (2020), 112243. https://doi.org/10.1016/j.enconman.2019.112243
38. J. Liang, S. Ge, B. Qu, K. Yu, F. Liu, H. Yang, et al., Classified perturbation mutation based particle swarm optimization algorithm for parameters extraction of photovoltaic models, *Energy Convers. Manage.,* **203** (2020), 112138. https://doi.org/10.1016/j.enconman.2019.112138