Combination of Artificial Neural Network-based Approaches to Control a Grid-connected Photovoltaic Source under Partial Shading Condition

Noureddine Akoubi ‡, Jamel Ben Salem, Lilia El Amraoui

Research Laboratory Smart Electricity & ICT, SE&ICT Lab., LR18ES44. National Engineering School of Carthage, University of Carthage, 4 Entrepreneurs Street, 2035 Charguia II, Tunisia

(yaakoubi.n@gmail.com, bsj_jamel@yahoo.fr, lilia.elamraoui@gmail.com)

‡ Noureddine Akoubi, Higher Institute of Technological Studies of Nabeul, 8000 Nabeul, Tunisia,

Tel: +216 25 284 387, Fax: +216 72 220 033, iset.contact@gmail.com

Received: 13.10.2022 Accepted:17.03.2023

Abstract- This paper proposes an approach based on artificial neural networks (ANN) to control a grid-connected photovoltaic system (PVS) under partial shading (PS) conditions. In PS conditions, the P-V curve exhibits multiple peaks, with only one representing the global maximum power point (GMPP), and the others representing local maximum power points (LMPP). Traditional Maximum Power Point Tracking (MPPT) methods are unable to identify the GMPP and get stuck around an LMPP, which results in reduced productivity of the PVS. The proposed approach combines supervised learning (SL) and deep reinforcement learning (DRL) techniques to design a controller with a hierarchical structure that can overcome the problem of identifying the GMPP in PVSs under PS conditions. The PVS under study consists of four identical solar panels. At the first control level, each solar panel has a sub-controller designed using ANN and the SL technique, which determines the appropriate duty cycle to extract the maximum power from the solar panel based on real-time weather conditions. At the second level, a DRL agent identifies the optimal duty cycle for the DC/DC converter from the duty cycles generated by the sub-controllers. The Deep Deterministic Policy Gradient (DDPG) and Twin Delayed DDPG (TD3) agents are implemented and evaluated for the second level of control. Simulation results using MATLAB/Simulink demonstrate the effectiveness of the proposed controller in tracking the GMPP.

Keywords photovoltaic source; Maximum Power Point Tracking; partial shading conditions; global maximum power point; artificial neural networks.

1. Introduction

Improving energy productivity and quality are the main challenges faced by photovoltaic systems in overcoming the high cost of installing this form of renewable energy [1, 2]. To meet this challenge, the research focuses on two main orientations. The first is related to photovoltaic cell technology [3, 4, 5] and the second is concerned with the control system of these photovoltaic panels [6, 7]. With the growth of PV installations, the impact of partial shading (PS), which is caused by several factors such as clouds, buildings, trees, and snow, represents a problem that significantly affects the productivity of photovoltaic systems and is, therefore, an important research topic. Indeed, traditional MPPT techniques, such as the perturbation and observation (PO) algorithm and the incremental conductance (IC) algorithm, are inefficient when the photovoltaic system (PVS) is composed of several solar panels and subjected to a partial shading condition. This weather situation unbalances the amount of solar irradiation captured by the different solar panels, and the MPPT controller is unable to track the true maximum power point. This limitation arises due to the presence of multiple local maximum power points (LMPPs) on the P-V curve, which makes it difficult for the traditional MPPT algorithm to determine the global maximum power point (GMPP). As a result, the algorithm gets stuck around an LMPP [8, 9]. The optimization of control strategies for PVSs to reduce energy losses resulting from PS conditions is a crucial area of

research in the solar energy field [10]. Various approaches have been proposed to mitigate the effects of PS, and among these, evolutionary algorithms have gained significant attention as an effective means of addressing power loss problems related to the PV effect and optimizing electrical energy production in PV systems [11,12]. As a result, researchers have been actively exploring and adapting this optimization approach to the control context of PVSs. This research has led to a surge in investigations aimed at advancing the development of novel control strategies that integrate evolutionary algorithms to improve the energy productivity and quality of PV systems in the presence of partial shading.

Indeed, in [13], the authors adopted the particle swarm optimization (PSO) algorithm to control a PVS under different weather conditions, including scenarios of PS. Nevertheless, the authors did not evaluate the accuracy of their PSO algorithm in tracking the GMPP, which is an important performance criterion for controlling PVS. This criterion highlights a weakness of most evolutionary algorithms. In the works [14, 15, 16], the authors are more interested in control performance, specifically the accuracy of tracking the global maximum power point (GMPP), by evaluating improved algorithms based on particle swarm optimization (PSO), grey wolf optimization (GWO), and Cat Swarm Optimization (CSO) metaheuristics, respectively. In [17], the authors combined the artificial bee colony (ABC) algorithm with the traditional perturb and observe (PO) method, aiming to improve the responsiveness and accuracy of tracking the GMPP. In [18], the use of the genetic algorithm (GA) combined with a backstepping controller yields better control performance.

These works [13-18], which propose improved solutions based on metaheuristics, have a critical drawback in terms of implementation in a real system. Indeed, the application of metaheuristics requires high computation time, and improving these algorithms increases their complexity and execution time. Additionally, the input variables are of an electrical nature characterized by rapid fluctuations that disturb and destabilize the operation of the controller based on these metaheuristics. In conclusion, the proposed solutions are hardly applicable to real PVSs.

Researchers have addressed the limitation of metaheuristic algorithms by using artificial neural networks (ANN), which can take in slow and smooth climatic variables as inputs for improved control stability of photovoltaic systems. Multiple studies, including [19] and [20], have validated the effectiveness of this approach in enhancing control performance. Nevertheless, some studies use an ANN with reduced structure to control PVSs by acquiring electrical variables [21]. Moreover, the authors [21] are obliged to combine the neural network with a conventional PO algorithm to guarantee an optimal performance. The studies [19, 20, 21] use supervised learning (SL) for ANN controller synthesis. This learning technique is insufficient in the case of partial shading condition, unlike deep reinforcement learning (DRL) which was adopted in [22, 23]. However, the DRL algorithm requires a large amount of training data with very fine tuning to obtain a valid control model under all weather conditions. In [22], the DRL agents are learned over 30000 episodes and require excessive time. Moreover, in [22], the solution is implemented on openAI Gym, which is not suitable for the complex dynamic system. The authors [22] proposed an analytically valid solution, but it does not take into account the internal dynamics of the PV system. It would be better to validate this solution on MATLAB/Simulink using the Simscape library to have a model closer to reality. In [23], the authors even combined DRL with fuzzy logic to increase the learning rate and reduce the training time.

Currently, consistent research work on the control of PVSs under SP conditions using DRL is too few, although this approach is very attractive since the synthesis of the control model does not require knowledge of the PVS model. Our proposal is to combine the benefits of SL with the performance of DRL to design a controller that can track the GMPP in any weather condition and can be implemented in a real PVS. Indeed, this paper proposes a novel approach that integrates supervised learning with deep reinforcement learning to effectively eliminate the issue of power loss in PVSs caused by PS conditions, which result in reduced productivity.

The proposed approach empowers the design of a hierarchical controller (ANN-RL controller) with the ability to accurately track the GMPP in real-time. The proposed hierarchical controller represents a significant advancement in PVS control by effectively addressing the issue of power loss due to PS. Through its ability to efficiently adapt to changing environmental conditions and optimize energy production. The ANN-RL controller is structured hierarchically and includes multiple ANN sub-controllers developed through SL, as well as a main controller developed through DRL. Each sub-controller receives weather data from an individual solar panel and determines the optimal duty cycle. After being generated by the ANN sub-controllers, the duty cycles are sent to the main controller to determine the optimal duty cycle for the entire PV system that aligns with the GMPP. The effectiveness of the main controller is evaluated using the Deep Deterministic Policy Gradient (DDPG) agent and the Twin Delayed DDPG (TD3) agent.

The paper is divided into five sections. The first section presents the PVS model. The second section details the implementation of the PO method and its limitations. The third section presents the design approach of the proposed ANN-RL controller, and simulation results are provided to demonstrate its effectiveness. The paper concludes with a summary of its contributions.

2. Presentation of the studied photovoltaic system

The studied system, which is a grid-connected PVS, consists of a photovoltaic array with four solar panels, a DC-DC boost converter, and a DC-AC inverter [24, 25]. This power conversion chain is interfaced with the electrical grid using a filtering system and a transformer, as shown in Fig. 1. The PVS is driven by a MPPT controller and a Voltage Source Converter (VSC) controller that control the DC-DC converter and the DC-AC converter, respectively. The essential modules of the grid connected PVS, which are implemented using MATLAB/Simulink, will be detailed in the following paragraph.

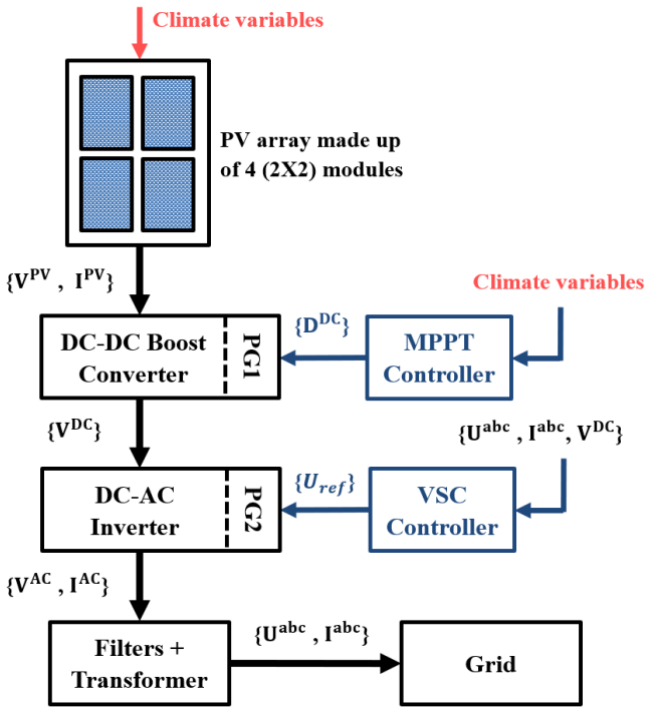

Fig. 1. Grid-Connected PVS.

Where PG1 and PG2 are both pulse width modulation (PWM) signal generators which control respectively the DC-DC Boost converter and the AC-DC inverter.

2.1. Structure of the photovoltaic cell array

The photovoltaic array is composed of two series of SunPower SPR-E18-295-COM reference panels connected in parallel, where each series is composed of two panels, as shown in Fig. 2. The maximum power output of the four modules is 120 kW. The capacitors C11, C12, C21, and C22 have identical values of 400 µF.

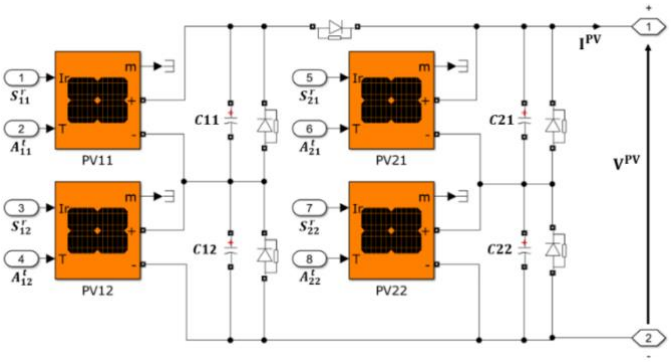

Fig. 2. The connection of the four solar panels.

Each PV panel consist of 34 parallel strings made up of three modules connected in series. The maximum power of the PV panel is 30 kW.

2.2. The DC-DC boost converter

The DC-DC boost converter, as illustrated in Fig. 3, controls the power generated by the PVS and provides an output voltage whose magnitude is greater than the input voltage [25, 26]. To activate the DC-DC boost converter, the duty cycle D^{DC} is converted into a PWM signal with a frequency of 5 kHz using a PG1 generator.

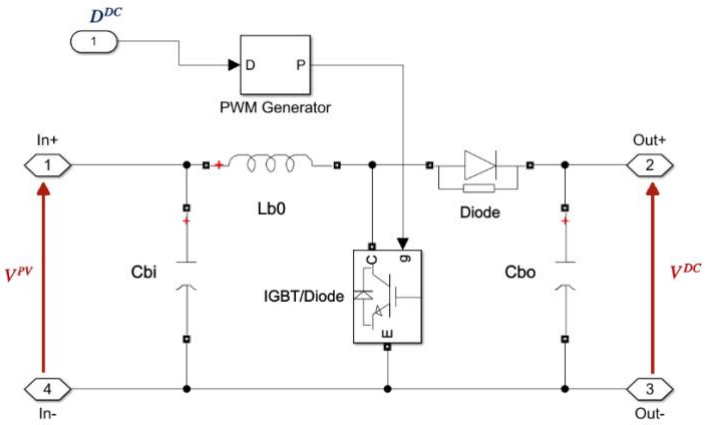

Fig. 3. DC-DC Boost Converter

Where the Input capacitor Cbi=200 µF, the inductor Lb0=10 mH and the output capacitor Cbo=1200 µF.

2.3. VSC controller setting for DC-AC inverter

The DC-AC inverter operates with a VSC controller to convert 500 V DC to 260 V AC while maintaining a constant power factor of 1. The VSC controller is composed of four modules [26 ,27]. The first module determines the value of the reference current I^{d}_{ref} using the following equation:

$$I^{d}_{ref} = k^{p}_{v}\left(V^{DC}_{mes} - V^{DC}_{ref}\right) + k^{i}_{v}\int\left(V^{DC}_{mes} - V^{DC}_{ref}\right)dt \qquad (1)$$

Where $k^{p}_{v} = 7$ and $k^{i}_{v} = 800$.

The second module is a phase-locked loop (PLL) that detects the phase angle of the grid and synchronizes it with the grid voltage. The three phase voltages measured on the grid side are the inputs to the PLL model, and the output is the phase angle. The phase angle is used in the Park transformation. The third module is a Feed Forward Current Controller which generates the voltages V^{d}_{conv} and V^{q}_{conv} according to the following equations:

$$V^{d}_{conv} = u^{d} - L\omega I^{q}_{ref} + RI^{d}_{ref} + V^{d}_{mes} \qquad (2)$$

$$V^{q}_{conv} = u^{q} - L\omega I^{d}_{ref} + RI^{q}_{ref} + V^{q}_{mes} \qquad (3)$$

$$u^{d} = k^{p}_{i}\left(I^{d}_{ref} - I^{d}_{mes}\right) + k^{i}_{i}\int\left(I^{d}_{ref} - I^{d}_{mes}\right)dt \qquad (4)$$

$$u^{q} = k^{p}_{i}\left(I^{q}_{ref} - I^{q}_{mes}\right) + k^{i}_{i}\int\left(I^{q}_{ref} - I^{q}_{mes}\right)dt \qquad (5)$$

Where $k^{p}_{i} = 0.3$, $k^{i}_{i} = 20$, $L = 0.25$ mH and $R = 2$ mΩ

The fourth module performs the inverse Park transformation and generates the reference voltage U_{ref}.

3. Control of the PVS by the traditional MPPT technique

The MPPT controller operates a DC-DC converter that is connected to 4 solar panels operating under identical ambient temperature A^{t} and uniform solar irradiation S^{r} with a single point of maximum power. The multiple maximum power points in the PV characteristic are caused by the partial

shading of one or more panels. In partial shading conditions, the shaded panels absorb less sunlight, generating less current, and consequently reducing the power generated by the panel. This causes a mismatch between panels, leading to multiple maximum power points in the P-V characteristic. In this case, the traditional MPPT controller will be unable to identify the GMPP and will be trapped in an LMPP. If photovoltaic panels of different age and technology are connected in series, the PS phenomenon will be present constantly.

The first section of this paragraph describes the design of the traditional MPPT controller, which is based on the perturb and observe (PO) method [28]. This control technique's performance in tracking the maximum power point is evaluated in this section. In the second section, the limitations of the traditional PO method are presented under partial shading (PS) conditions [29]

3.1. Performance of traditional PO method

The diagram illustrating the PO method used to control the power generated by the PVS is shown in Fig. 4.

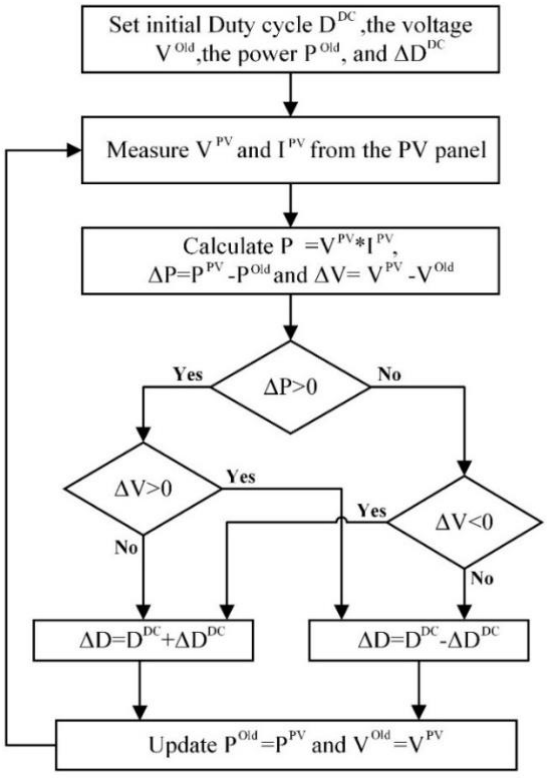

Fig. 4. Diagram of the PO algorithm

When all four solar panels are subjected to the same climatic conditions, the values of S_1^r, S_2^r, S_3^r, S_4^r and A_1^t, A_2^t, A_3^t, A_4^t are equal to S^r and A^t respectively. The results of the simulation for different levels of solar radiation and temperature are illustrated in Figures 5 and 6. The unit of time used in all figures depicting the simulation results is hours (h).

This evolution scenario of climatic variables exhibits several levels of variation. For instance, at t=1h, there is a small increase in solar irradiation of 100 W/m². At t=2h, the increase in solar irradiation is of an average value of 200 W/m², and at t=3h, the increase is significant with a value of 500 W/m². Similarly, in the case of a decrease in solar irradiation, two levels of variation of 200 W/m² at t=6h and 500 W/m² at t=8.5h are observed. The A^t temperature evolution also displays multiple levels of temperature with various increases and decreases. Figure 7 illustrates the power supplied to the grid for different levels of climatic variables.

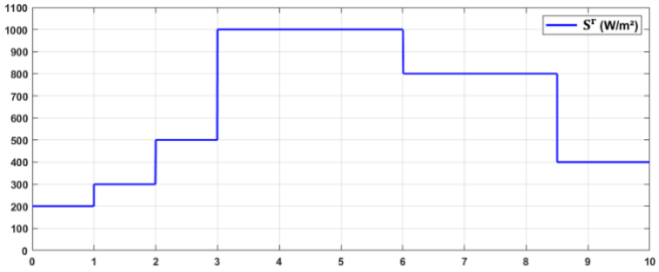

Fig. 5. Evolution of the solar irradiation S^r.

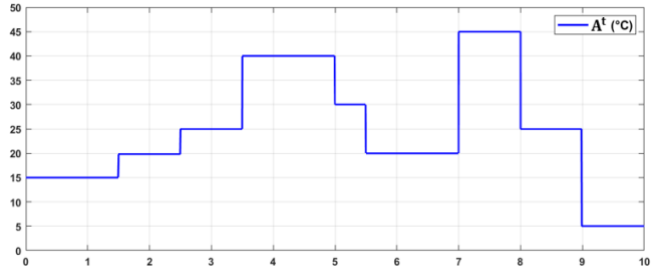

Fig. 6. Evolution of the ambient temperature A^t.

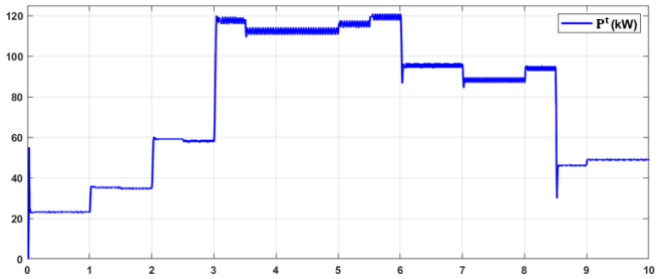

Fig. 7. Power supplied to the grid.

To assess the performance of the system, the maximum power point (MPP) is determined from the P-V curves of the photovoltaic system (PVS) for every level of solar irradiation S^r and ambient temperature A^t, as depicted in Figure 8. The gap between the transmitted power and the MPP, which reflects the accuracy of the MPPT controller using the PO method, is given by the following formula:

$$GAP(\%) = \frac{MPP - P^t}{MPP} 100 \quad (6)$$

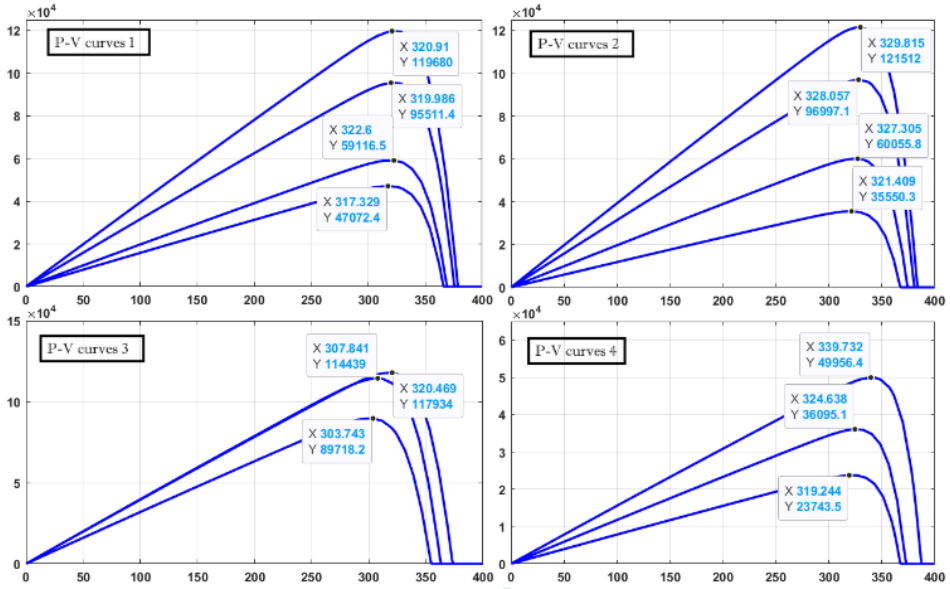

Fig. 8. P-V curves for different levels of S^r and A^t

Figure 9 illustrates the GAP (%) value over each time interval.

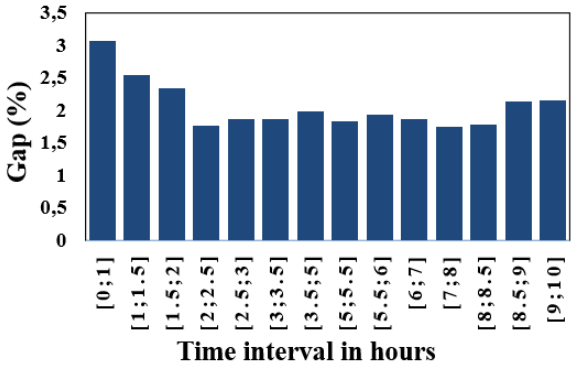

Fig. 9. Error rate of the PO algorithm.

The average gap between the MPP and the generated power is 1.95%. In the analysed case, the PO algorithm can identify and tracking the MPP with an efficiency of 98.05%. The P-V curve consistently exhibits a single GMPP. It has been observed that solar irradiation has a greater impact on the generated power than temperature. The subsequent analysis primarily focuses on evaluating the effect of changes in irradiation on the performance of the PO method.

3.2. Limitation of traditional methods in the case of partial shading conditions

In scenarios where the level of solar irradiation varies among panels, the P-V curve may display two points of maximum power. The higher one is called the GMPP, and the other is referred to as the LMPP. Let us consider a specific climatic situation (SC1), defined as follows: $S_1^r = S_3^r$=1000 W/m², S_2^r=300 W/m², S_4^r=500 W/m² and $A_1^t = A_2^t = A_3^t = A_4^t = A^t$ =25°C. Figure 11 illustrates that there are two maximum power points in this case. The first, which is the GMPP, corresponds to V^{PV} = 169.944 V for P^t=59220.5W, and the second, which is the LMPP, corresponds to V^{PV} = 355.997 V for P^t=51093.7 W. The power difference between the two points is ΔP=8126.8 W.

To assess the effectiveness of the traditional PO control method, a scenario involving variations in solar irradiance, as illustrated in Fig. 10, is considered. The PVS is initially placed under SC1 conditions, which correspond to the P-V characteristics depicted in Fig. 11, for a duration of 2 hours. Following this, the PVS is exposed to changes in irradiance levels for 6 hours before returning to the original SC1 climatic conditions, which are maintained for an additional 2 hours. Figure 12 illustrates that the PVS produces a derived power of 57.93 kW under partial shading conditions corresponding to SC1. This power output is near the GMPP point, which equals 59.22 kW. The system reaches this point due to a specific initialization of the duty cycle value in the PO algorithm, which is set at 0.75. However, following an irradiance change, the PO controller is unable to identify the GMPP again and becomes trapped around the LMPP, resulting in a transmitted power of 50.03 kW.

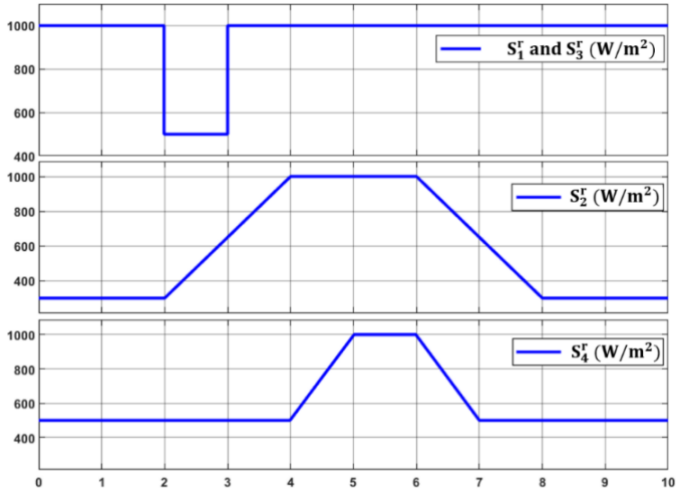

Fig. 10. Variations of solar irradiation for each photovoltaic panel

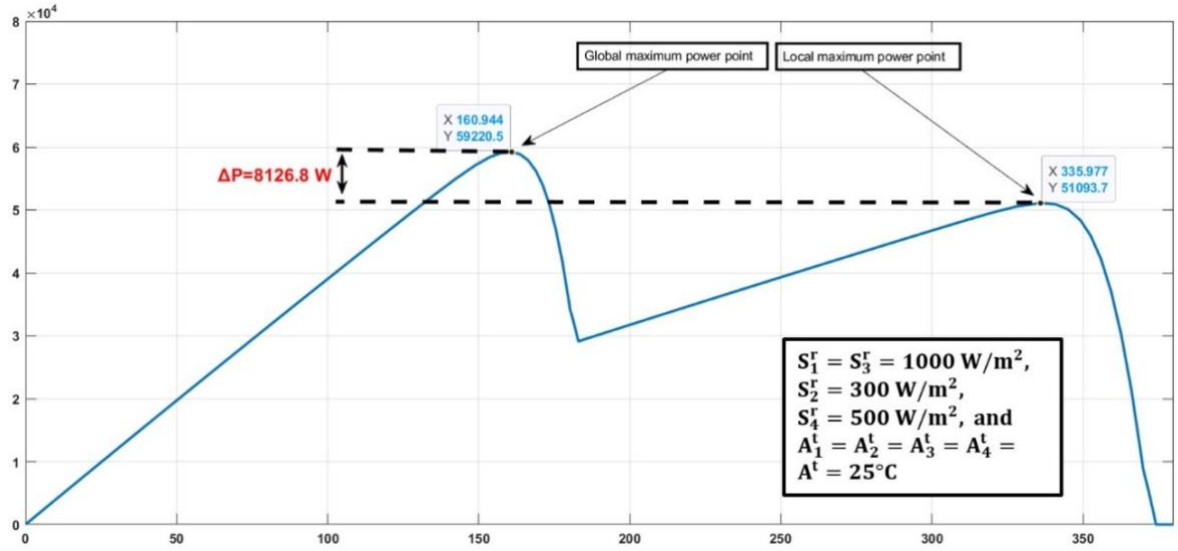

Fig. 11. P-V curve under partial shading condition

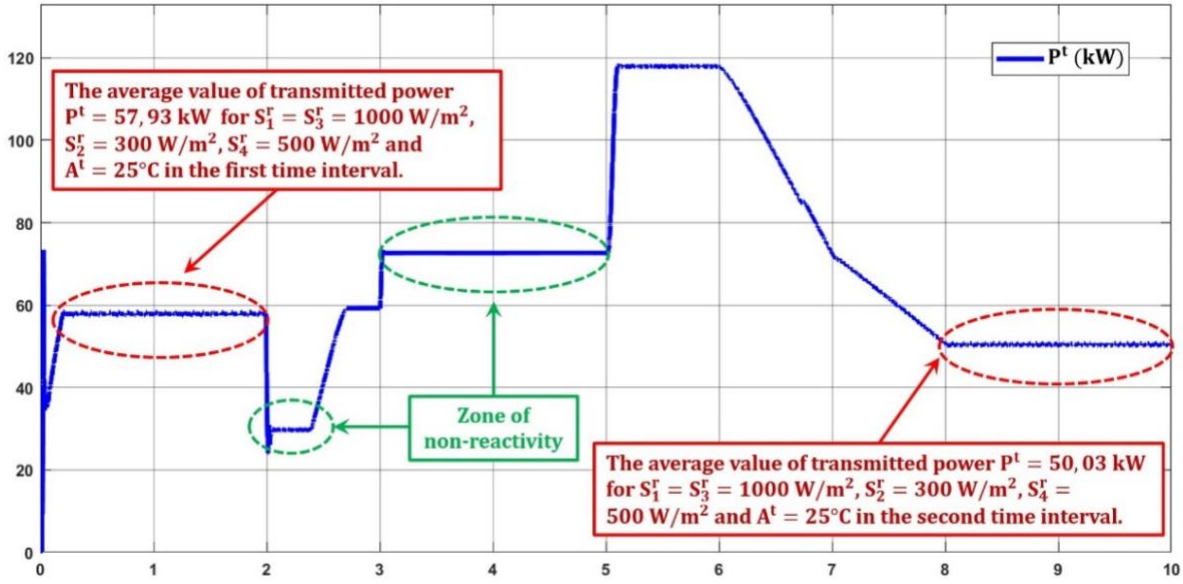

Fig. 12. Evolution of the transmitted power to the grid using PO method.

4. Proposed approach for the control of a PVS using ANNs

The proposed approach involves a two-level controller structure, as shown in Fig.13. At the first level, a local ANN sub-controller is directly connected to a single solar panel. The ANN sub-controller, which is implemented using supervised learning techniques, generates the duty cycle based on the climatic conditions of the individual solar panel. At the second level, the main controller, developed using Deep Reinforcement Learning (DRL), receives duty cycles from the various ANN sub-controllers and determines the optimal duty cycle, D_{opt}^{DC}.

It should be noted that the duty cycle values, cycles D_1^{DC}, D_2^{DC}, D_3^{DC} and D_4^{DC}, generated by the ANNs must be highly accurate to enable the RL agent to identify the optimal duty cycle. To address this challenge, a careful approach is employed in the design of the ANN sub-controller, which is described in the following paragraph.

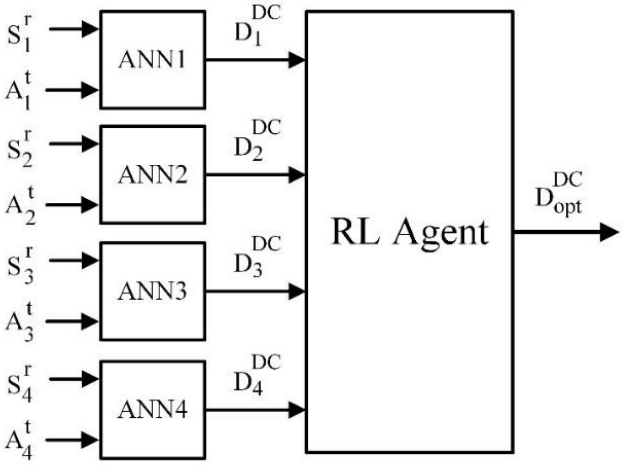

Fig. 13. Structure of the proposed ANN-RL controller

4.1. Design of the ANN sub-controller using supervised learning

The PO controller relies on acquiring and interpreting internal electrical variables V^{PV} and I^{PV}, which are influenced by the dynamics of various elements within the PVS. The design of the ANN sub-controller aims to replicate the dynamic behavior of the PO controller using the two climate variables, S^r and A^t, as inputs. The position of the maximum power is suitably tracked by the weather variables SS^r and A^t, without any disturbance. To achieve this, a well-thought-out approach is employed, covering the entire operating space (OS) of the solar panel and generating the duty cycle $D^{DC}_{j,k}$ for each value of S^r and A^t using the PO algorithm [19].

The synthesis of the ANN sub-controller involves defining a set of equilibrium points (EP) within the OS of the PVS. These EPs serve as a basis for training the ANN sub-controller. To ensure the accuracy and robustness of this first level of control [19], a uniform distribution of EPs throughout the OS is proposed, as illustrated in Fig. 14.

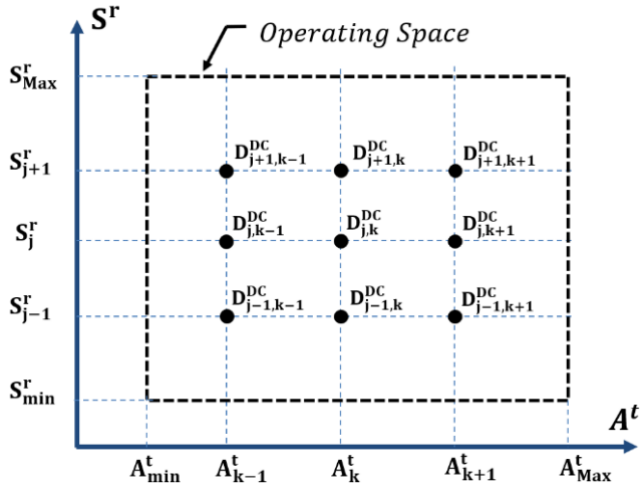

Fig. 14. Uniform distribution of the equilibrium points.

Solar irradiance S^r and temperature $A^t$are used as inputs into the ANN sub-controller to determine the output duty cycle D^{DC}. The equilibrium point is defined by the triplet (S^r, A^t, D^{DC}). The PVS operating space must be covered by all these EPs. The OS for S^r is $\Omega S = [S^r_{min}; S^r_{Max}]$ and for A^t is $\Omega A = [A^t_{min}; A^t_{Max}]$. Then, for each $S^r_j \in \Omega S$ and $A^t_k \in \Omega A$ the duty cycle $D^{DC}_{j,k}$.With $j \in [1; N]$ and $k \in [1; M]$. The OS for D^{DC} is $\Omega S.\Omega A$. The total number of EP is P=MxN. In the synthesis of a supervised learning controller, it is necessary to determine beforehand the optimal value of D^{DC} for each pair (S^r and A^t). This optimal value of D^{DC} allows to generate the maximum power of the attached PVS. The optimal D^{DC} value is identified using a conventional MPPT algorithm. The model of the PVS controlled by the PO method will be used to generate the training data for the ANN sub-controller. The training data generation algorithm is illustrated in Fig. 15.

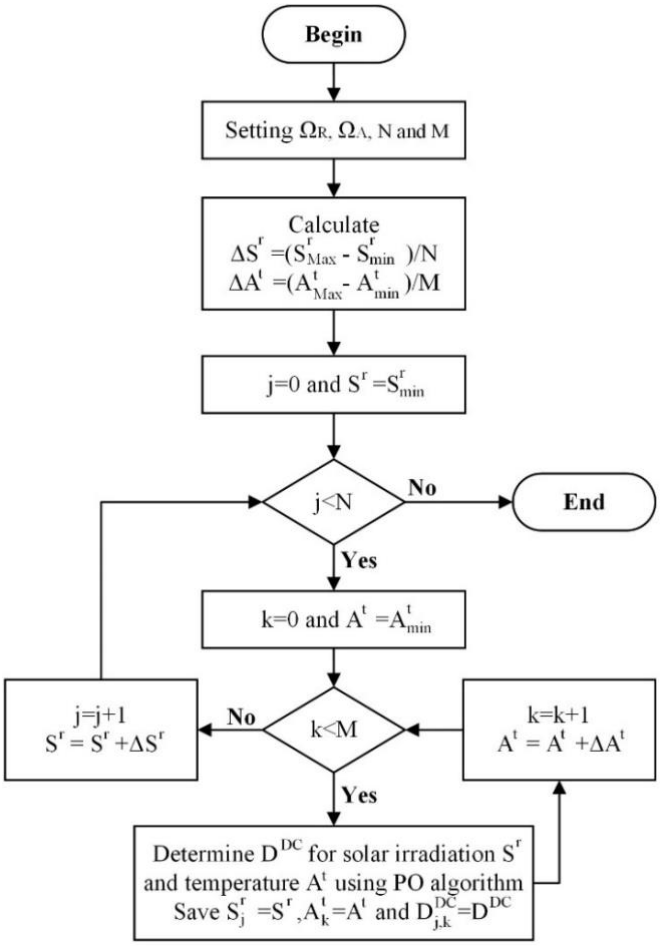

Fig. 15. Diagram of the training data generation algorithm

Where $A^t_{min} = 0$, $A^t_{Max} = 50$, $S^r_{min} = 200$, $S^r_{Max} = 1000$, N=10 and M=40.

The proposed ANN sub-controller is structured on three layers: an input layer with two acquisition neurons, a hidden layer of ten neurons that mix the acquired climate variables S^r and A^t, and an output layer with a single artificial neuron that delivers the value of the duty cycle. The ANN sub-controller is trained using the MATLAB software application "Neural Net Fitting". Of the 400 equilibrium points, 72% will be used for training, and the remaining 28% will be divided equally for testing and validation.

4.2. Design of the RL main controller using deep reinforcement learning

Deep reinforcement learning is a technique to design a control policy without prior knowledge of the dynamic behavior of the system. This approach is based on the interaction of an agent with its environment [31, 32]. The agent chooses an action based on its observations and rewards to update its policy settings during learning. A DRL algorithm will then try to maximize the total reward received by the agent. In the literature, there are several DRL algorithms, such as SARSA, DQN, PPO, TD3 and DDPG [33, 34]. In this context, we proceed to the exploitation of DRL for the main control of a PVS. The problem is to identify the optimal duty

cycle D_{opt}^{DC} that follows the GMPP. Figure 16 illustrates the integration of the agent in the environment.

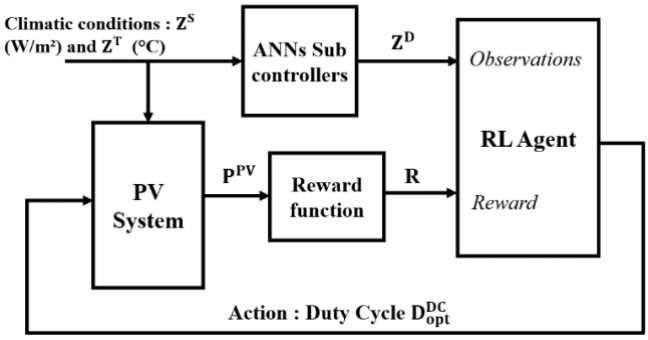

Fig. 16. Training model of the RL Agent

The training of the RL agent is carried out through an iterative process with the aim of optimizing the power output of the PVS. The climatic condition is defined by two vectors $Z^S = [S_1^r, S_2^r, S_3^r, S_4^r]$ and $Z^T = [A_1^t, A_2^t, A_3^t, A_4^t]$. Each pair of climate data (S_i^r, A_i^t) are processed by the ANNi sub-controller, as detailed in the previous Fig. 14, to generate the duty cycle D_i^{DC} with i=1,2,3 and 4. The observation data of the agent RL is defined by the vector $Z^D = [D_1^{DC}, D_2^{DC}, D_3^{DC}, D_4^{DC}]$.

The design of a controller based on an ANN requires an essential learning phase, which is validated by satisfying a set of performance criteria. In the case of DRL, the learning phase involves conducting a set of experiments. Due to the strong influence of irradiation on the power output of PVS with respect to ambient temperature, we propose a uniform distribution of irradiation, as shown in Algorithm 1 below. The temperatures are defined randomly.

Algorithm 1 :Algorithm for generating the learning scenario of the DRL agent

Intialize $\Delta S^r = (S_{Max}^r - S_{min}^r)/L$
for i=1, L do
  for j=1, L do
    for k=1, L do
      for h=1, L do
        $Z^s(i,j,k,h) = [S_1^r, S_2^r, S_3^r, S_4^r]$
        A_1^T = random(5,50)
        A_2^T = random(5,50)
        A_3^T = random(5,50)
        A_4^T = random(5,50)
        $A^T(i,j,k,h) = [A_1^T, A_2^T, A_3^T, A_4^T]$
        $S_4^r = S_4^r + \Delta S^r$
        End for
      $S_3^r = S_3^r + \Delta S^r$
      End for
    $S_2^r = S_2^r + \Delta S^r$
    End for
  $S_1^r = S_1^r + \Delta S^r$
End for

The function random(5,50) returns a random value between 5 and 50. L represents the number of solar irradiance levels, which is equal to 6. The total number of solar irradiance combinations is 6^4 =1296. The reward function, proportional to the P^{PV} power value, is expressed as $R = \alpha P^{PV} - \beta$, where $\alpha = 10^{-4}$ and $\beta = 10$. The learning process is carried out for both the TD3 and DDGP agents. It should be noted that the reward value is proportional to the power generated by the PVS.

Figures 17 and 18 depict the results of the training process for the TD3 and DDGP agents, respectively, displaying the reward value obtained for each episode. An increase in this value indicates that the RL agent is approaching the optimal duty cycle and generating more power from the PVS.

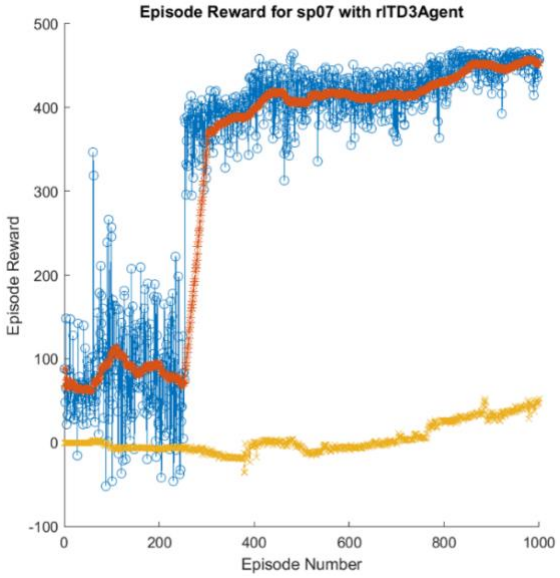

Fig. 17. Evolution of the reward value during the training of the TD3 agent.

The TD3 agent quickly identified the optimal duty cycle for power generation within the first 200 episodes and continued to improve it. Its learning process was completed in less than 1000 episodes. These results demonstrate the success of the proposed approach for designing the two-level controller. The learning process of the DDGP agent is conducted for 1000 episodes, the same number as for the TD3 agent.

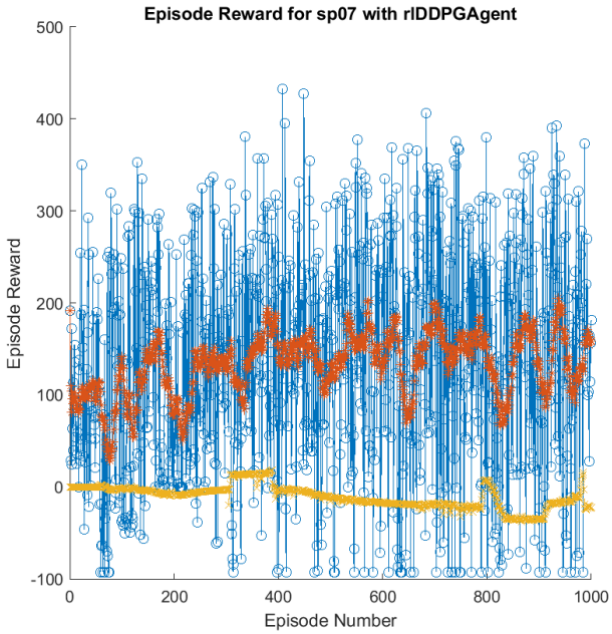

Fig. 18. Evolution of the reward value during the training of the DDGP agent.

The behaviour of the TD3 and DDGP agents during the learning process was assessed by visualizing the reward value

obtained in each experiment, as shown in Fig. 17 and Fig.18. The TD3 agent exhibited a rapid improvement in its strategy after exploring the environment, whereas the DDGP agent failed to converge to an optimal strategy and remained in a state of environment exploitation. This demonstrates the advantage of TD3 in efficiently adapting its policy to the environment and achieving better performance in learning tasks compared to DDGP.

Figure 19 and Figure 20, shows respectively the simulation results of 6 experiments performed by the DDGP agent and the TD3 agent.

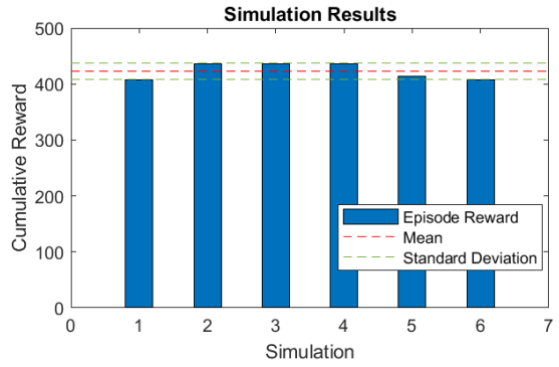

Fig. 19. Reward value for the DDGP trained agent.

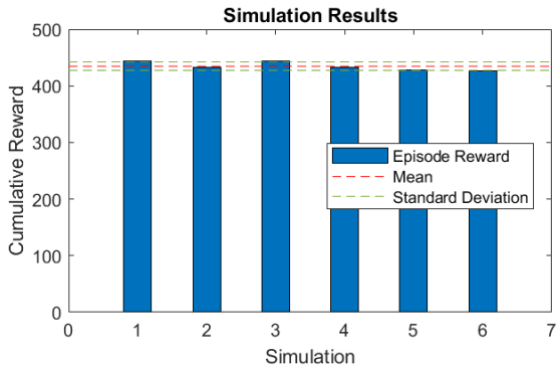

Fig. 20. Reward value for the TD3 trained agent.

Table 1 summarizes the maximum, minimum, and average reward values obtained by the TD3 and DDGP agents across six experiments.

Table 1. Synthesis of the rewards obtained for the TD3 and DDPG agents.

	DDGP Agent	**TD3 Agent**
Average reward value	422.766	434.77
Maximum reward value	435.845	443.845
Minimum reward value	407.615	427.335

The TD3 agent achieved a higher reward value and better performance in tracking the GMPP compared to the DDGP agent. Therefore, in the following evaluation of the simulation results of the PVS under PS condition, the TD3 agent was used as the main controller.

4.3. Simulation result and discussion performance of the proposed controller

The following section presents the simulation results of the power transmitted to the power grid using the ANN-RL controller. Thus, a comparative study is conducted with the traditional PO controller. The result of simulation, for the two studied controllers, is given by Fig. 21.

Based on the results presented in Fig. 21, the power generated by the PVS for both controllers are averaged over a 10-hour simulation period. Table 2 shows the average power produced by the PVS when using the conventional PO controller and the proposed new ANN-RL controller, and the corresponding power gain achieved.

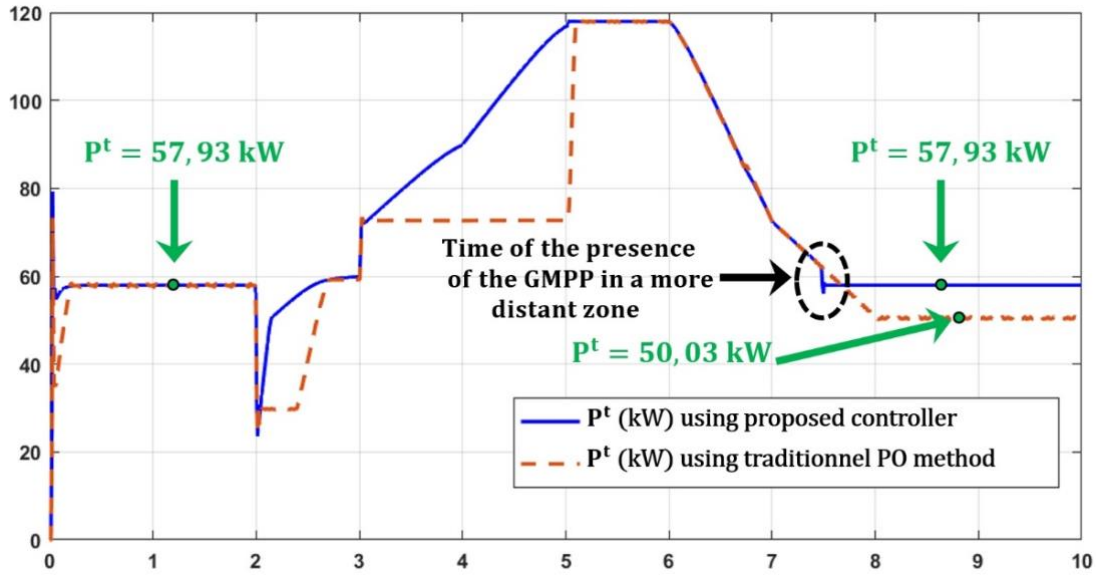

Fig. 21. Transmitted power using the ANN-RL controller and using the PO method.

Table 2: Average power output for each controller and power gain obtained.

Average power produced using the conventional PO controller	74.68 kW
Average power produced using the proposed new ANN-RL controller	67.51 kW
Power gain achieved	10.62 %

The average value of the transmitted power using the proposed ANN-RL controller is equal to 74.68 kW against 67.51 kW using the PO method. Therefore, the ANN-RL controller increases the power generated by the PVS by 10.62% $\left\{\frac{74.68-67.51}{67.51} * 100 = 10.62\%\right\}$ compared to the conventional PO controller. This is the minimum gain that can be achieved since the PO controller is well initialized at the start of the algorithm. Analytically, the maximum power under the studied PS condition is equal to 59.22 kW and the power generated by the ANN-RL controller is equal to 57.93 kW. The deviation from the maximum power value is 2.17 % which is a small and acceptable deviation even in a uniform shading condition.

In addition, the learning process of the main controller agent is significantly reduced to only 1000 episodes, providing evidence that the ANN sub-controllers effectively simplified the main controller's task. The integration of SL with DRL eliminates the complexities that arise from implementing DRL for PV system control under partial shading conditions. This promising combination opens new opportunities to apply this approach in various domains of electrical system control.

5. Conclusion

In this paper, a robust controller that combines SL with DRL is proposed to address the PS problem in a grid-connected PVS composed of four solar panels. The proposed ANN-RL controller is composed of a set of sub-controllers developed using SL and a main controller implemented using DRL. The comparative study reveals that the TD3 agent outperforms the DDGP agent in tracking the maximum power point during the development of the main controller. Unlike the traditional MPPT controller, the proposed ANN-RL controller can detect the GMPP of a PVS under PS conditions, resulting in a significant increase in energy production. Specifically, for the case study presented in this paper, the gain is at least 10%, which validates the integration of ANN approaches for PVS control. Ultimately, the combination of supervised learning and reinforcement learning simplifies controller design and implementation, while also presenting potential opportunities for enhancing the efficacy of control systems across diverse industrial settings.

References

[1] Ratnakar Babu Bollipo, Suresh Mikkili and Praveen Kumar Bonthagorla, "Critical Review on PV MPPT Techniques: Classical, Intelligent and Optimisation", IET Renewable Power Generation Vol. 14 Iss. 9, pp. 1433-1452, June 12, 2020.

[2] Indresh Yadav, Sanjay Kumar Maurya and Gaurav Kumar Gupta, "A literature review on industrially accepted MPPT techniques for solar PVS", International Journal of Electrical and Computer Engineering (IJECE) Vol. 10, No. 2, pp. 2117-2127, April 20, 2020.

[3] Divya Sharma, Rajesh Mehra, Balwinder Raj, "Comparative analysis of photovoltaic technologies for high efficiency solar cell design", Superlattices and Microstructures, Elsevier, Mars 2021.

[4] Neelam Rathore, Narayan Lal Panwar, Fatiha Yettou, Amor Gama, "A Comprehensive review on different types of solar photovoltaic cells and their applications", nternational Journal of Ambient Energy, Taylor & Francis & Informa UK Limited, trading as Taylor & Francis Group, Mars 11, 2019.

[5] Anshul Awasthi, Akash Kumar Shukla, Murali Manohar S.R., Chandrakant Dondariya, K.N. Shukla, Deepak Porwal, Geetam Richhariya, "Review on sun tracking technology in solar PVS",Energy Reports, February 17, 2020.

[6] Mingxuan Mao, Lichuang Cui, Qianjin Zhang, Ke Guo, Lin Zhou, Han Huang, "Classification and summarization of solar photovoltaic MPPT techniques: A review based on traditional and intelligent control strategies", Energy Reports Vol. 6, Pages 1312-1327, November 2020.

[7] Mohammad Sarvi and Ahmad Azadian, "A comprehensive review and classifed comparison of MPPT algorithms in PVS", Energy Systems Vol. 13, pp 281–320 ,March 18, 2021.

[8] Manoja Kumar Behera, Lalit Chandra Saikia, "A new combined extreme learning machine variable steepest gradient ascent MPPT for PVS based on optimized PI-FOI cascade controller under uniform and partial shading conditions", Sustainable Energy Technologies and Assessments, Elsevier, 26 October 2020.

[9] Yihao Wan, Mingxuan Mao, Lin Zhou, Qianjin Zhang, Xinze Xi, Chen Zheng, "A Novel Nature-Inspired Maximum Power Point Tracking (MPPT) Controller Based on SSA-GWO Algorithm for Partially Shaded Photovoltaic Systems", Electronics 2019, June 15, 2019.

[10] G. Sai Krishna, Tukaram Moger, "Reconfiguration strategies for reducing partial shading effects in photovoltaic arrays: State of the art", Solar Energy, Elsevier, February 23, 2019.

[11] Mohamed A. Mohamed, Ahmed A. Zaki Diab and Hegazy Rezk, "Partial Shading Mitigation of PVSs via Different Meta-Heuristic Techniques", Renewable Energy, August 22, 2018.

[12] Cliford Choe Wei Chang, T. J. Ding M. A. S. Bhuiyan, K. C. Chao, M. Ariannejad, Haw Choon Yian, "Nature-Inspired Optimization Algorithms in Solving Partial Shading Problems: A Systematic Review", Archives of Computational Methods in Engineering (2023), Vol 30, pp. 223–249, Springer, https://doi.org/10.1007/s11831-022-09803-x, 19 August 2022.

[13] Ali M Eltamaly, Mamdooh S Al-Saud, Ahmed G Abokhalil, Hassan MH Farh, "Photovoltaic maximum power point tracking under dynamic partial shading changes by novel adaptive particle swarm optimization strategy", Transactions of the Institute of Measurement and Control 2020, Vol. 42, Issue 1, August 28, 2019.

[14] Mingxuan Mao, Qichang Duan, Pan Duan and Bei Hu1, "Comprehensive improvement of artificial fish swarm algorithm for global MPPT in PVS under partial shading conditions", Transactions of the Institute of Measurement and Control, 2017.

[15] Dipanwita Debnath, Nirmala Soren, Arun Dev Pandey, and Noman Hanif Barbhuiya, "Improved Grey Wolf assists MPPT Approach for Solar Photovoltaic System under Partially Shaded and Gradually Atmospheric Changing Condition", International Energy Journal, Vol 20, pp. 87 – 100, 2020.

[16] T. Nagadurga, P. V. R. L. Narasimham and V. S. Vakula, "Global Maximum Power Point Tracking of Solar Photovoltaic Strings under Partial Shading Conditions Using Cat Swarm Optimization Technique", Sustainability, October 2021.

[17] Deepthi Pilakkat, S. Kanthalakshmi, "Single phase PVS operating under Partially Shaded Conditions with ABC-PO as MPPT algorithm for grid connected applications", Energy Reports Vol. 6, Pages 1910-1921, November 2020.

[18] Cheikhne Cheikh Ahmed, Mohamed Cherkaoui, Mohcine Mokhlis, and Mouad Bahij, "Genetic Algorithm and Backstepping Controller for Photovoltaic System under Partial Shading Effect", INTERNATIONAL JOURNAL of RENEWABLE ENERGY RESEARCH, Vol.11, No.1, March, 2021.

[19] Noureddine Akoubi, Jamel Ben Salem and Lilia El Amraoui, "Contribution on the Combination of Artificial Neural Network and Incremental Conductance Method to MPPT Control Approach", 2022 5th International Conference on Advanced Systems and Emergent Technologies (IC_ASET), IEEE, May 25, 2022.

[20] Adi Kurniawana, and Eiji Shintaku, "A Neural Network-Based Rapid Maximum Power Point Tracking Method for Photovoltaic Systems in Partial Shading Conditions", Applied Solar Energy, Vol. 56, No. 3, pp. 157–167, January 7, 2020.

[21] Senthamizh Selvan Sakthivel, Venkadesan Arunachalam, "Artificial Neural Network Assisted P&O-Based MPPT Controller for a Partially Shaded Grid-Connected Solar PV Panel", Arabian Journal for Science and Engineering, Springer, December 2022.

[22] L. Avila, M. De Paula, M. Trimboli et al., "Deep reinforcement learning approach for MPPT control of partially shaded PV systems in Smart Grids", Applied Soft Computing Journal (2020), https://doi.org/10.1016/j.asoc.2020.106711, September 2020.

[23] Yaduvir Singh, Nitai Pal, "Reinforcement learning with fuzzified reward approach for MPPT control of PV systems", Sustainable Energy Technologies and Assessments, Volume 48, ScienceDirect, Elsevier, December 2021.

[24] Muhammad Abu Bakar Siddique, Adeel Asad, Rao M. Asif, Ateeq Ur Rehman, Muhammad Tariq Sadiq and Inam Ullah, "Implementation of Incremental Conductance MPPT Algorithm with Integral Regulator by Using Boost Converter in Grid-Connected PV Array", IETE Journal of Research, Taylor & Francis Group, May 10, 2021.

[25] Shaowu Li, "A variable-weather-parameter MPPT control strategy based on MPPT constraint conditions of PVS with inverter", Energy Conversion and Management Vol. 197, October 1, 2019.

[26] Unal Yilmaz, Omer Turksoy, Ahmet 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 pp. 634-651, December 2019.

[27] Aurobinda Bag, Bidyadhar Subudhi and Pravat Kumar Ray, "A Combined Reinforcement Learning and Sliding Mode Control Scheme for Grid Integration of a PVS", CSEE Journal of power and energy systems Vol. 5, No. 4, December 2019.

[28] Mohd Faisal Jalil, Shahida Khatoon, Ibraheem Nasiruddin and R. C. Bansal, "Review of PV array modelling, configuration and MPPT techniques", International Journal of Modelling and Simulation, Taylor & Francis Group, Jun 21, 2021.

[29] Mazen Abdel-Salam, Mohamed-Tharwat EL-Mohandes and Mohamed Goda, "History of Maximum Power Point Tracking", , Modern Maximum Power Point Tracking Techniques for Photovoltaic Energy Systems pp 1-19, Green Energy and Technology, July 31, 2019.

[30] Min Ding, Dong Lv, Chen Yang, Shi Li, Qi Fang, Bo Yang, and Xiaoshun Zhang, "Global Maximum Power Point Tracking of PVSs under Partial Shading Condition: A Transfer Reinforcement Learning Approach", Applied Sciences, July 9, 2019.

[31] Bin Huang and Jianhui Wang, "Deep-Reinforcement-Learning-Based Capacity Scheduling for PV-Battery Storage System", IEEE Transation on smart grid Vol. 12, No. 3, May 2021.

[32] Paulo Lissaa, Conor Deane, Michael Schukat, Federico Seri Marcus Keane and Enda Barrett, "Deep

reinforcement learning for home energy management system control", Energy and AI Vol. 3, March 2021.

[33] Daniel J.B. Harrold, Jun Cao, Zhong Fan, "Renewable energy integration and microgrid energy trading using multi-agent deep reinforcement learning", Applied Energy, April 14, 2022.

[34] Waldemar Kolodziejczyk, Izabela Zoltowska and Pawel Cichosz, "Real-time energy purchase optimization for a storage-integrated photovoltaic system by deep reinforcement learning", Control Engineering Practice Vol. 106, January 2021.

[35] Saad Motahhir, Aboubakr El Hammoumi and Abdelaziz El Ghziza, "The Most Used MPPT Algorithms: Review and the Suitable Low-cost Embedded Board for Each Algorithm", Journal of Cleaner Production, Elsevier, October 20, 2019.

[36] CH Hussaian Basha and C Rani, "Different Conventional and Soft Computing MPPT Techniques for Solar PVSs with High Step-Up Boost Converters: A Comprehensive Analysis", Energies Vol. 13, January 12, 2020.

[37] S Satheesh Kumar and A Immanuel Selvakumar, "Maximum power point tracking and power flow management of hybrid renewable energy system with partial shading capability: A hybrid technique", Transactions of the Institute of Measurement and Control, 2020.

[38] César G. Villegas-Mier, Juvenal Rodriguez-Resendiz, José M. Álvarez-Alvarado, Hugo Rodriguez-Resendiz, Ana Marcela Herrera-Navarro, and Omar Rodríguez-Abreo, "Artificial Neural Networks in MPPT Algorithms for Optimization of Photovoltaic Power Systems: A Review", Micromachines 2021, October 17, 2021.

[39] Ferdaws Ben Naceur, Chokri Ben Salah, Achraf Jabeur Telmoudi and Mohamed Ali Mahjoub, "Intelligent approach for optimal sizing in photovoltaic panel-battery system and optimizing smart grid energy", Transactions of the Institute of Measurement and Control, 2021.

[40] Milad Fathi and Jafar Amiri Parian, "Intelligent MPPT for photovoltaic panels using a novel fuzzy logic and artificial neural networks based on evolutionary algorithms", Energy Reports Vol. 7 pp. 1338–1348 , March 2, 2021.

[41] Bao Chau Phan, Ying-Chih Lai, and Chin E. Lin, "A Deep Reinforcement Learning-Based MPPT Control for PVSs under Partial Shading Condition", Sensors, 27 May 2020.

[42] A.T.D. Perera, P.U. Wickramasinghe, Vahid M. Nik, Jean-Louis Scartezzini, "Introducing reinforcement learning to the energy system design process", Applied Energy Vol. 262, March 15, 2020.

[43] P. Kofinas, S. Doltsinis, A.I. Dounis, G.A. Vouros, "A reinforcement learning approach for MPPT control method of photovoltaic sources", Renewable Energy Vol. 108, pp. 461-473, August 2017.