



**TECHNICAL UNIVERSITY OF CRETE  
SCHOOL OF ELECTRICAL AND COMPUTER ENGINEERING**

**Diploma thesis**

**Design of an Electronic Control System  
for Maximizing the Energy Production of  
Photovoltaic Arrays based on Artificial  
Intelligence Techniques**

**Christos Kalogerakis**

**EXAMINATION COMMITTEE**

**Associate Professor Eftychios Koutroulis (Supervisor)**

**Associate Professor Michail Lagoudakis**

**Professor Kostas Kalaitzakis**

**Chania, September 2019**

# Abstract

The subject of this thesis is the design of an electronic energy management system for maximizing the power generated by a photovoltaic (PV) array. For that purpose, an innovative Maximum Power Point Tracking (MPPT) algorithm was developed, which is based on reinforcement learning, in order to operate the PV array at the Maximum Power Point (MPP) under uniform and non-uniform incident solar irradiation conditions. The PV system under study consists of an MPPT control unit, a DC/DC Boost-type power converter and a battery. For the implementation of the MPPT control system, four different Q-learning-based MPPT methods and a Particle Swarm Optimization-based (PSO) MPPT method were implemented. The Q-learning-based MPPT algorithms were simulated for multiple alternative shading patterns of the PV array and their performance was compared to that of the PSO-based MPPT method. The simulation results demonstrated that the Q-learning-based methods exhibit faster convergence to the global MPP (GMPP) than the PSO-based MPPT method when an appropriate learning process has been applied before their execution.

# CONTENTS

1. Introduction .....	7
1.1 The subject of this thesis .....	7
1.2 The structure of the thesis .....	10
2. Photovoltaic Systems .....	11
2.1 Technologies of photovoltaic systems .....	11
2.2 Estimation of equivalent circuit parameters of PV modules .....	13
2.3 Photovoltaic array curves under uniform irradiation .....	15
2.4 Photovoltaic array curves under non-uniform irradiation .....	16
2.5 DC/DC power converters .....	18
2.5.1 DC/DC Boost-type power converters .....	18
2.5.2 DC/DC Buck-type power converters .....	20
2.6 PV array topologies .....	23
2.7 Maximum Power Point Tracking techniques .....	25
2.7.1 Perturbation and observation MPPT algorithm .....	26
2.7.2 Incremental-Conductance MPPT algorithm .....	27
2.7.3 Particle Swarm Optimization MPPT algorithm .....	29
3. Maximum Power Point Tracking using Q-learning .....	32
3.1 Theoretical analysis .....	32
3.2 Q-learning for PV MPPT .....	35
3.2.1 Action policy selection .....	36
3.2.2 State-space .....	37
3.2.3 Action-space .....	38
3.2.4 Reward .....	40
3.2.5 Discount factor, learning rate .....	41
3.2.6 Q- learning-based algorithm process .....	42
4. The Simulink model of the PV system .....	46
4.1 Photovoltaic arrays .....	47
4.2 DC/DC Boost-type power converter .....	48
5. Simulation results .....	54
5.1 Analysis of the simulation results for shading pattern 1 .....	56

5.2 Analysis of the simulation results for shading pattern 2 .....	61
5.3 Analysis of the simulation results for shading pattern 3 .....	66
5.4 Analysis of the simulation results for shading pattern 4 .....	71
5.5 Analysis of the simulation results for shading pattern 5 .....	76
5.6 Analysis of the simulation results for shading pattern 6 .....	81
5.7 Analysis of the simulation results for shading pattern 7 .....	86
5.8 Analysis of the simulation results for shading pattern 8 .....	91
5.9 Analysis of the simulation results for shading pattern 9 .....	96
5.10 Analysis of the simulation results for shading pattern 10 .....	101
5.11 Analysis of the simulation results for shading pattern 11 .....	107
5.12 Analysis of the simulation results for shading pattern 12 .....	112
5.13 Analysis of the simulation results for shading pattern 13 .....	117
5.14 Analysis of the simulation results for shading pattern 14 .....	122
5.15 Analysis of the simulation results of the PSO-based MPPT method .....	127
5.16 Aggregated simulation results .....	131
5.17 Overall remarks .....	146
6. Conclusions .....	148
7. References .....	150



## Symbol Catalog

<i>V</i>	<i>Voltage (V)</i>
<i>I</i>	<i>Current (A)</i>
<i>P</i>	<i>Power (W)</i>
<i>G</i>	<i>Solar irradiation intensity (W/m<sup>2</sup>)</i>
<i>PV</i>	<i>Photovoltaic</i>
<i>STC</i>	<i>Standard Testing Conditions</i>
<i>SC</i>	<i>Short Circuit</i>
<i>OC</i>	<i>Open Circuit</i>
<i>I<sub>sc</sub></i>	<i>Short Circuit Current</i>
<i>V<sub>oc</sub></i>	<i>Open Circuit Voltage</i>
<i>MPP</i>	<i>Maximum Power Point</i>
<i>MPPT</i>	<i>Maximum Power Point Tracking</i>
<i>GMPP</i>	<i>Global Maximum Power Point</i>
<i>I<sub>MPP</sub></i>	<i>Current at maximum Power Point</i>
<i>V<sub>MPP</sub></i>	<i>Voltage at maximum Power Point</i>
<i>P<sub>MPP</sub></i>	<i>Power at maximum Power Point</i>
<i>DC</i>	<i>Duty cycle</i>
<i>PWM</i>	<i>Pulse Width Modulation</i>
<i>P&amp;O</i>	<i>Perturb &amp; Observe</i>
<i>INCOND</i>	<i>Incremental Conductance</i>
<i>lr</i>	<i>Learning Rate</i>
<i>SP</i>	<i>Series-Parallel</i>
<i>TCT</i>	<i>Total-Cross-Tied</i>
<i>BL</i>	<i>Bridge-Linked</i>
<i>PSO</i>	<i>Particle Swarm Optimization</i>

*R*      *Reward*

*$\gamma$*       *Discount factor*

*MDP*   *Markov Decision Process*

*MF*    *Model-free*

*D*      *Duty cycle in current time-step*

*DO*   *Duty cycle in a previous time-step*

*DP*   *Dynamic Programming*

# 1.

# Introduction

---

## 1.1 The subject of this thesis

During the last decade, a major priority of the scientific community is the global warming and the energy policies. The reduction of greenhouse gas emissions is a principal objective for most of the developed countries. Today, European Union (EU) has been able to reduce greenhouse gas emissions by 20% compared to the 1990s. In addition, EU countries target to capture at least 32% of their energy needs using renewable energy sources by 2030. The original target of at least 27% was revised upwards in 2018 [1]. In this context, the power generated by PV panels has an important role. Greenhouse gases are emitted only during the production of the PV array's installation components. When their installation is complete, electricity generation by solar irradiation is possible, without the emission of greenhouse gases. PV panels can produce more energy than that required for their manufacturing, during their lifetime. PV panels major advantage is that they can be installed in isolated areas, where there is not an electricity network available [2]. All of the above are designated as off-grid facilities and in areas such as remote islands or mountains, they can be an alternative, economical and environmentally sustainable solution. But in reality, in most cases, the power generated by renewable energy sources does not come from off-grid facilities, but from grid-connected facilities where the power is supplied to the grid. It is therefore understood that renewable energy production is a way of economic growth for EU countries.

RES power generation is also very useful for domestic use, because it is a way to increase financial revenues by supplying electricity to the power grid [3]. Despite its advantages, PV

power generation has some major disadvantages. It is more expensive than other resources, because of the high cost of its components (DC/AC inverter, wirings, batteries, solar panels and installation) [3]. Governments are promoting it with subsidies or feed-in tariffs, expecting the development of the technology, so that in the near future it will become a more competitive energy source compared to conventional energy sources. This is an important goal of the scientific community. This can be achieved by increasing RES efficiency, which will result in increased incomes and reduced costs of RES-generated energy. In this way, it will approach the cost of the power production by other sources [4]. During cloudy and rainy days, the PV systems energy production is low. That's the reason that batteries are being used, in order to store the PV-generated power during sunny days. Nowadays, modern battery technologies such as Lithium and Sodium-Sulphur batteries are most commonly being used for solar energy storage [5]. Furthermore, a PV power generation plant, needs a lot of space in order to be installed, and in some occasions, building roofs are not large enough to fit the number of PV panels necessary for covering the required energy demands. Lastly, a problem that occurs during the manufacturing process of the solar systems, is that many toxic materials are being released into the environment, which are associated with environmental pollution [6]. PV panels as a power source present many advantages. Their main advantage is that the sun is essentially an endless power source because it never ceases to emit light [6]. The technology that is used for this purpose is constantly evolving. For instance, there are materials now in use that are able to supply the required energy in small facilities with relatively small solar plants.

The existence of the photovoltaic phenomenon was first analyzed by Alexandre-Edmond Becquerel in the 19<sup>th</sup> century, however, today only 0.02% of the global energy demands are fulfilled by using solar energy [7]. Since 2009, there has been a steady decline in the cost of PV installations [8]. Nevertheless, the cost of PV installations can be further dramatically reduced. In order to achieve this goal, it is essential that solar energy technology to become cost-effective compared to fossil fuel or nuclear power systems. The aforementioned observations, combined with the slow but steady efficiency improvement in PV systems performance will make solar energy systems a dominant source of energy in the next 20 to 30 years [8].

The power generated by photovoltaic panels operating under real environmental conditions is affected by temperature and incident irradiance. In addition, PV modules are also affected by the partial shading caused by neighboring objects; i.e. buildings, trees etc. Thus, some of the PV modules may be subject to different shading levels [9]. This phenomenon can significantly influence the power output of the PV system, because a PV array that operates under partial

shading conditions presents multiple peaks at its PV array output power-voltage curve, due to the use of the bypass diodes [9]. For this reason, the PV system needs to be controlled, in order to operate at its maximum power point (MPP). To achieve this, the maximum power point tracking (MPPT) methods are necessary. These methods try to maximize the PV system power extraction under all environmental conditions in order to increase the PV plant efficiency. However, the efficiency of the power converter is also affecting the PV plant's efficiency. Improving the converter's efficiency is difficult, because more advanced technology is necessary, the components of which will certainly cost highly [10]. It is achievable though to increase the performance of MPPT control algorithms. There is no additional financial cost and it can be done in photovoltaic plants which are already operating by changing their control algorithm, which indirectly leads to an immediate reduction in total cost [10].

There is a wide variety of techniques for implementing the MPPT process, such as the Perturb and Observe (hill climbing method), Incremental conductance, Fractional Short-Circuit Current, Fractional Open-Circuit Voltage, Fuzzy Control, Neural Network Control, Particle Swarm Optimization-based algorithms etc. [11]. According to experimental results, the PSO-based MPPT method can detect the MPP in every shading pattern no matter where this point is located on the power-voltage curve of the PV source and that is the reason that it is considered to be the most efficient MPPT algorithm.

In this diploma thesis, an electronic energy management model was designed, in MATLAB<sup>TM</sup>/Simulink software. Specifically, a Q-learning-based MPPT method was developed for maximizing the power production of PV arrays which operate under partial shading conditions. For that purpose, a Simulink model was created, which consists of:

- a Photovoltaic array,
- a DC/DC Boost-type power converter for maximizing the energy produced at the converter's output and
- an MPPT control unit with the Q-learning-based algorithm, which determines the PWM signal of the DC/DC converter.

The control unit collects voltage and current signals, which change according to: the PWM signal's value, the solar radiation of the PV array and the temperature of the PV array. The Q-learning-based algorithm is used by the control unit of the PV system in order to help the learning agent to learn a map from states to optimal control actions online by updating the action values according to the received rewards. The experienced action values of every state that is visited are being stored in a Q-table. After the learning process, the learning agent is able to

know how to change the PWM signal value, in order to reach the MPP of the PV array, according to the state, under certain environmental conditions.

## **1.2 The structure of the thesis**

The structure of this diploma thesis is the following:

- Chapter 2: Information about photovoltaic systems and their functionality, different photovoltaic arrays structures analysis, study of current-voltage and power-voltage curves and description of already existing MPPT algorithms.
- Chapter 3: Description of the Q-learning-based algorithm for implementing the MPPT process.
- Chapter 4: Description of the Simulink model that was developed for simulating the operation of the PV system.
- Chapter 5: Presentation of the simulation results for different environmental conditions and scenarios. Comparison with conventional MPPT algorithms.
- Chapter 6: Conclusion analysis and proposals for future extensions.

---

# 2.

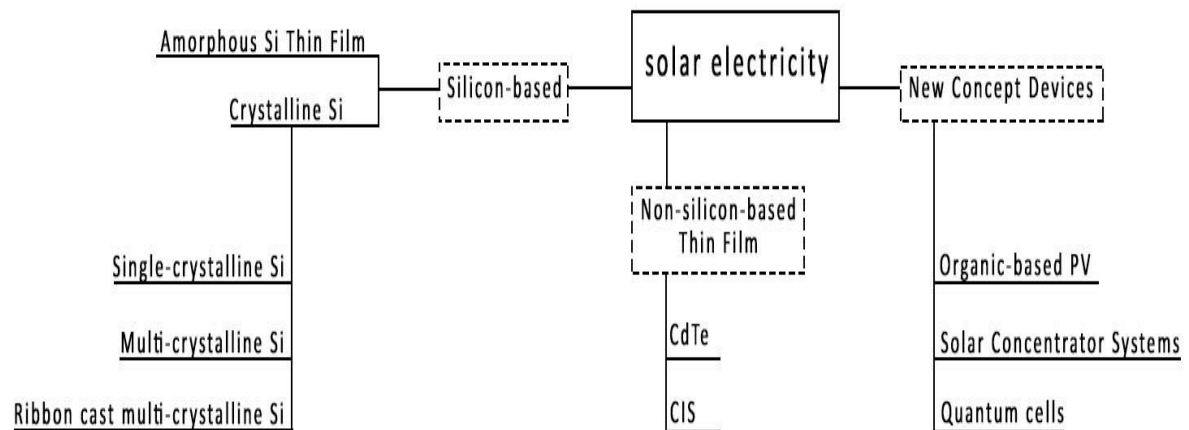
## Photovoltaic Systems

---

### 2.1 Technologies of photovoltaic systems

The most mature technology for PV systems that has been used for many decades is based on silicon as semiconductor. Alternative technologies that are based on non-silicon materials are also presented in Fig. 2-1 [12].

Multi-crystalline Si (mc-Si) wafers demand high amounts of energy in order to be produced, and were traditionally obtained through the re-melting of scraps of high-purity electronic grade silicon (EG-Si) into new ingots and subsequent mechanical slicing [12]. Single crystalline Si (Sc-Si) wafers, require an extra re-crystallization step, which further increases their energy demands. The industry has now introduced a new method for mc-Si ingots, which requires less energy for their production. In this method, primary metallurgical grade silicon (MG-Si) can directly produce a lower-grade silicon stock of sufficient purity for PV applications (solar grade silicon, SoG-Si). Recently, another type of Si-based modules was introduced in the market. They are called ribbon-Si. Their main advantage is that the time-consuming, energy-intensive and material-inefficient steps of Si development and shredding have been largely eliminated [12]. Because of the above it ends up being with higher energy yields and less financial cost, due to fact that approximately 100% of the Si feedstock ends up in the wafer material.

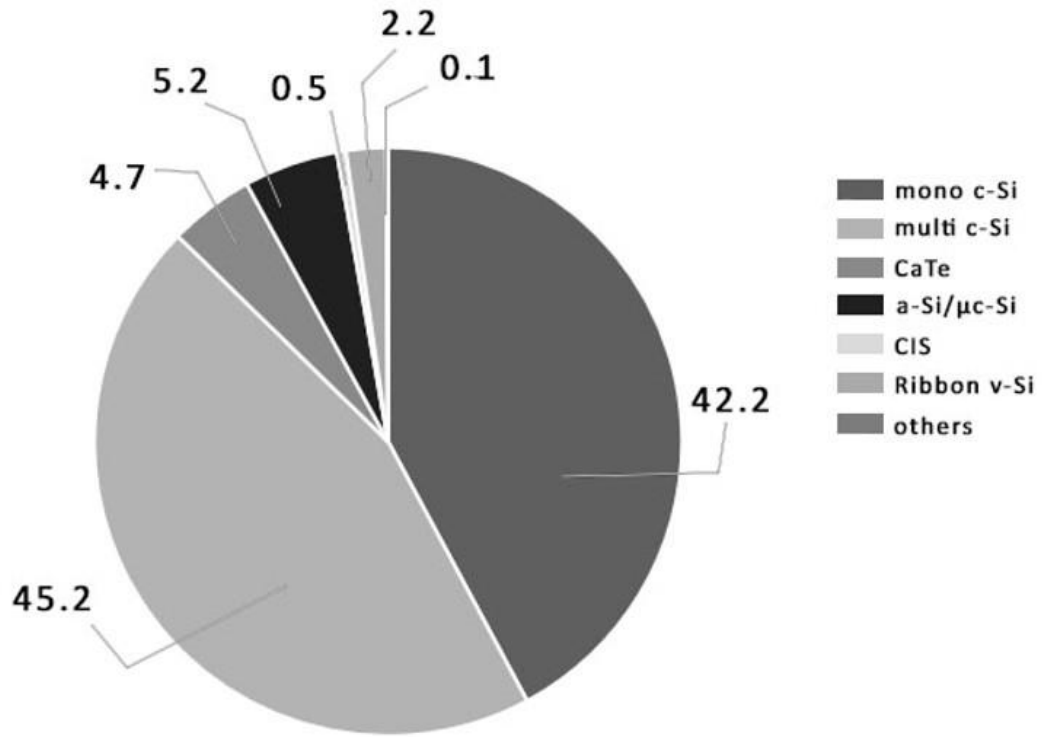


**Fig. 2-1.** PV technologies.

A thin layer of hydrogenated silicon deposited on glass is used in amorphous silicon (a-Si) modules. These materials have a low cost, but they have also a lower efficiency compared with others. They are usually used in consumer electronics or in building-integrated solutions [12].

Layers of thickness of a few nm are used in non-silicon-based PV modules (CIS and CdTe). Non-silicon-based PV modules are made by binary semiconductors, which are electrodeposited in glass panes. This kind of technologies are new in the industry. They were introduced in the early 2000s. Lastly, there is one more subcategory, the new concept devices. In this subcategory there are three families of technologies: (a) the first is based on organic polymer cells and it has low-medium efficiency and ultra-low cost, (b) the second is based on quantum cells and it has ultra-high efficiency and (c) the last family is called solar concentrator systems. With the exception of subcategory (c), these technologies are still in the experimental stage and they are called “third generation” PV technologies. The different PV technologies are presented in Fig. 2-2 [12].





**Fig. 2-2.** PV technologies market shares.

## 2.2 Estimation of equivalent circuit parameters of PV modules

An equivalent circuit shown in Fig. 2-3 is used together with the following set of circuit equations to express a typical current-voltage (I-V) characteristic of PV modules and arrays [13]:

$$I = I_{ph} - i_d - I_r \quad (2.1)$$

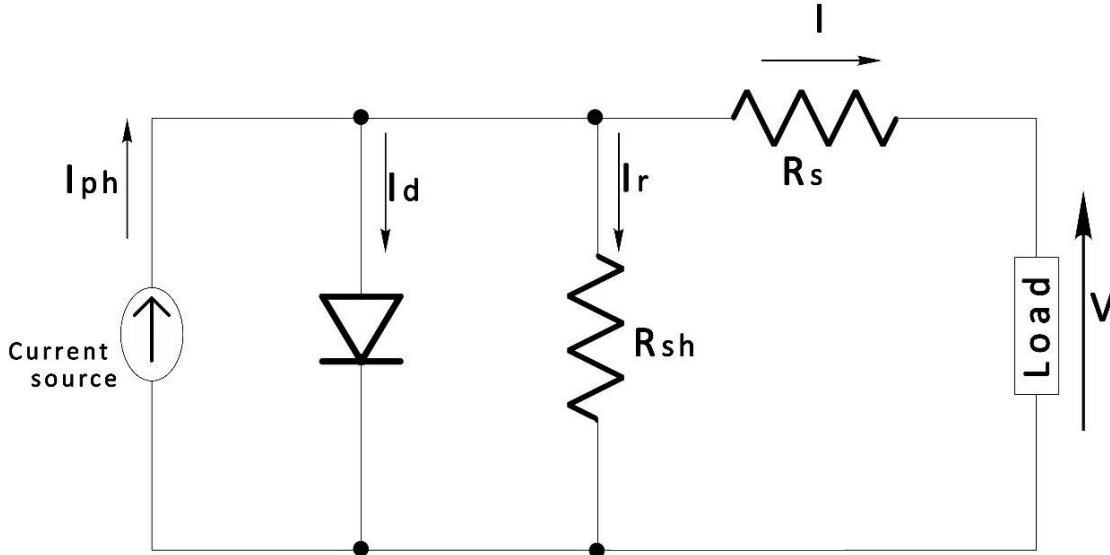
$$I_{ph} = I_{sho} \cdot \left( \frac{S}{1000} \right) + J_o \cdot (T - T_{ref}) \quad (2.2)$$

$$I_d = I_o \cdot \left[ \left( e^{\frac{q(V + R_s \cdot I)}{n \cdot k \cdot T}} \right) - 1 \right] \quad (2.3)$$

$$I_o = I_{d0} \cdot \left( \frac{T}{T_{ref}} \right)^3 \cdot e^{\frac{q \cdot E_g}{n \cdot k} \left( \frac{1}{T_{ref}} - \frac{1}{T} \right)} \quad (2.4)$$

$$I_r = \frac{V + R_s}{R_{sh}} \quad (2.5)$$

$$E_g = m \cdot \left( 1.16 - 7.02 \cdot 10^{-4} \cdot \frac{T^2}{T - 1108} \right) \quad (2.6)$$



**Fig. 2-3.** Equivalent circuit for PV modules and PV arrays.

where:

$S$  is the solar insolation ( $W/m^2$ )

$T_{ref}$  is the reference temperature (298 K)

$T$  is the cell temperature (K)

$E_g$  is the band gap energy of the cell semiconductor (eV)

$k$  is the Boltzmann constant

$R_s$  is the series resistance (Ohm)

$R_{sh}$  is the shunt resistance (Ohm)

$n$  is the diode emission factor

$I_{d0}$  is the diode reverse current (A)

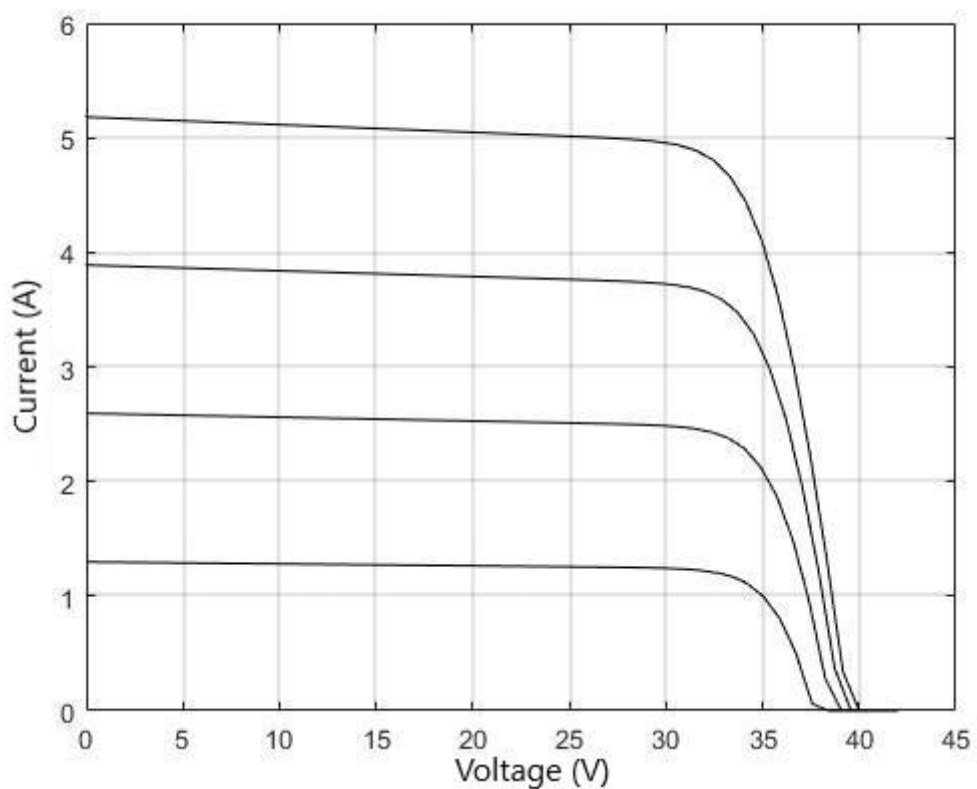
$I_{sho}$  is the short-circuit current at reference state (A)

$J_0$  is the temperature coefficient ( $mA/K$ ) and

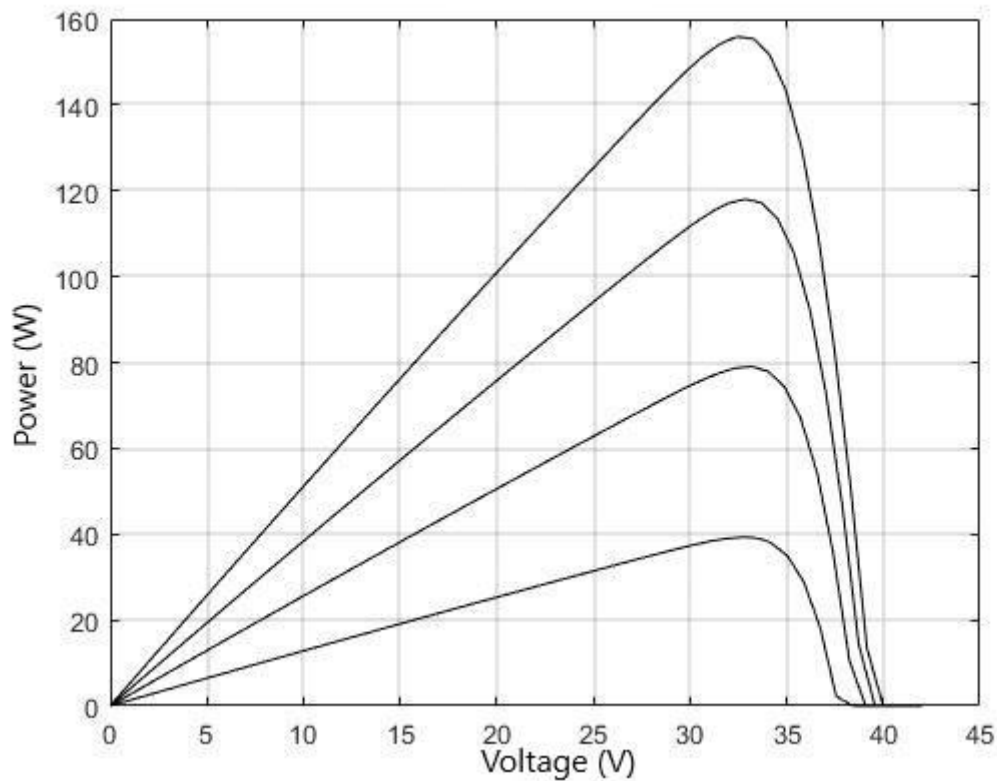
$m$  is the number of cells connected in series.

## 2.3 Photovoltaic array curves under uniform irradiation

The energy that is produced by a photovoltaic array is proportional to the intensity of the solar irradiation that strikes the panel. If the irradiation is uniform over the PV array (same intensity of solar irradiation in every PV panel of the array), then the resulting power-voltage curve will have an operation point, which will be the Maximum Power Point of the PV array under certain environmental conditions (temperature, solar Irradiation etc.). The solar irradiance is calculated in  $W/m^2$ . In Fig. 2-4 and 2-5 the I-V and P-V curves, respectively, are displayed.



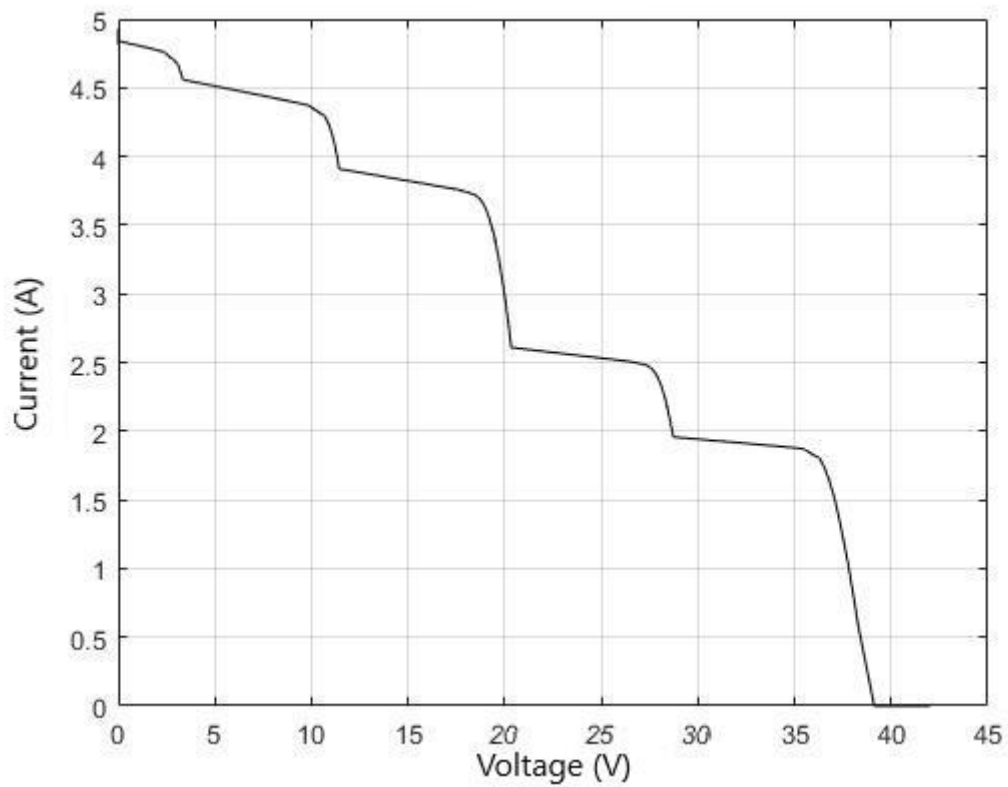
**Fig. 2-4.** Current-voltage curve for uniform incident solar irradiation.



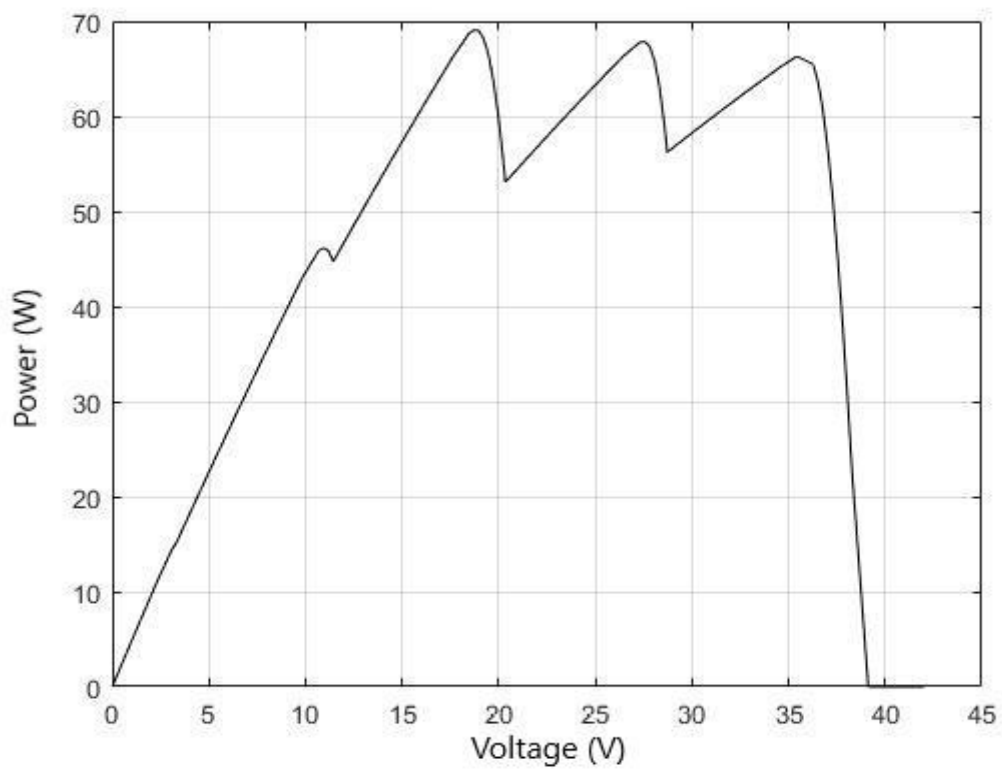
**Fig. 2-5.** Power-voltage curve for uniform incident solar irradiation.

## 2.4 Photovoltaic array curves under non-uniform irradiation

It was observed that during uniform solar radiation on the solar panels, there is only one peak in the power-voltage curve. The MPP can be easily located under these environmental conditions. The MPPT process is more complicated when the entire PV array is operating under partial shading conditions. Typically, partial shading is caused by the clouds that strike on certain spots of the PV array, while other parts of the PV array are left differently irradiated. The problem is even more complicated in PV arrays with long strings of PV panels. Partial shading is also created by the PV module irregularities; for instance, the presence of cracks on one or more PV modules of the array. During partial shading, the bypass diodes connected in parallel with each PV module create multiple peaks in the I-V and P-V curves characteristics, i.e. several local and one global peak. This effect can be observed in Fig. 2-6 and 2-7, respectively.



**Fig. 2-6.** Current-voltage curve of a PV array under partial shading conditions.



**Fig. 2-7.** Power-voltage curve of a PV array under partial shading conditions.

## 2.5 DC/DC power converters

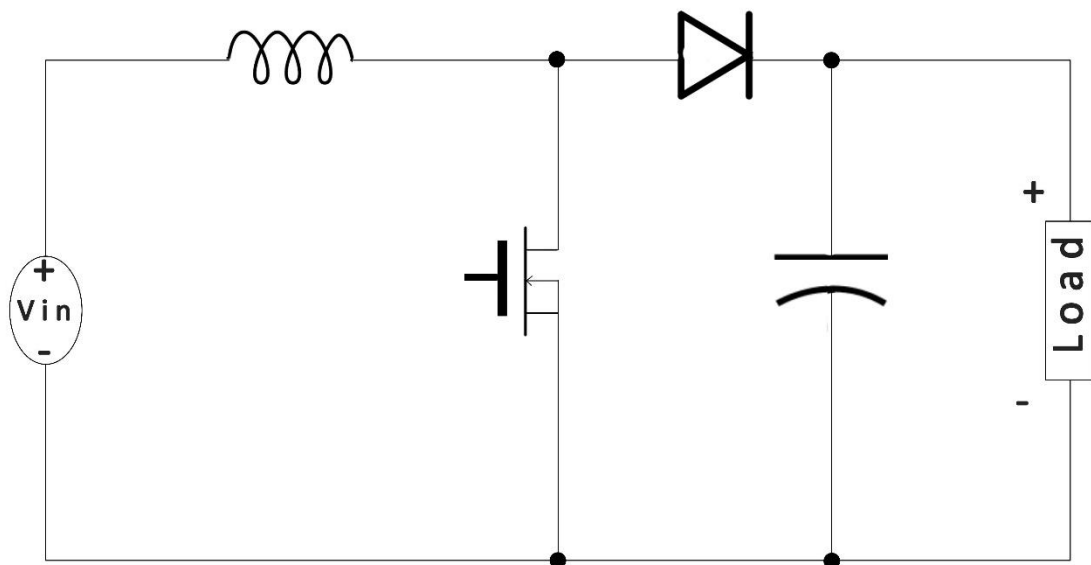
DC/DC power converters are electronic circuits which convert a source of DC from one voltage level into another voltage level. DC/DC power converters have a wide variety of applications:

- Power supply systems for telecommunications
- Power transmission
- Photovoltaic systems etc.

For photovoltaic systems, two kinds of DC/DC power converters are primarily being used. Boost-type power converters and buck-type power converters.

### 2.5.1 DC/DC Boost-type power converters

The boost-type power converter electronic circuit is displayed in Fig. 2-8. The DC output voltage, can be up to four times higher than the input voltage. This depends on the value of the input inductor and the semiconductor switch. The minimum output voltage value is equal to the input voltage value.



**Fig. 2-8.** Boost-type power converter circuit.

A boost-type power converter may operate in two distinct states: the on-state and the off-state, respectively. When the power converter is operating in the on-state, the Mosfet is closed, resulting in the inductor's current increment. On the other hand, in the off-state, the switch is

open and the only available path for the inductor current is through the flyback diode, the capacitor and the load [14]. Then, the energy that was accumulated during the on-state, is directly transferred into the output capacitor.

A DC/DC boost-type power converter may operate in the continuous conduction mode or in the discontinuous conduction mode. In continuous conduction mode, the current flowing through the inductor is always higher than zero. The waveforms of current and voltage of the power converter during continuous conduction mode are presented in Fig.2-9. After each switching period the inductor returns to the same state and therefore the inductor average DC voltage is equal to 0 at steady state [14]. The input-output voltages of the Boost-type power converter are related as follows:

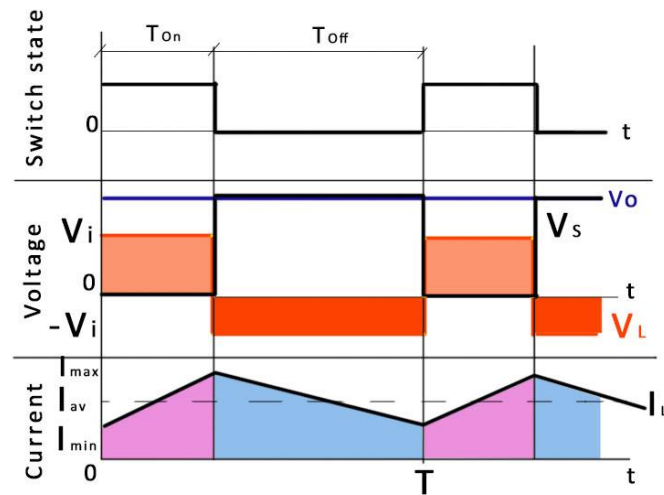
$$V_i = (1 - D) \cdot V_o \quad (2.7)$$

where:

$D$  is the duty cycle value

$V_o$  is the output voltage (V) and

$V_i$  is the input voltage (V).



**Fig. 2-9.** Current and voltage waveforms in continuous conduction mode

[14].

Under light loads, the inductor can be fully discharged before the commutation cycle is fully completed. This can happen provided that the ripple amplitude of the current is very high. As it can be observed in Fig. 2-9, during this operating mode the current that flows through the

inductor is becoming equal to 0 before the end of a switching period. This change greatly affects the voltage gain that becomes equal to:

$$\frac{V_o}{V_i} = \frac{1 + \sqrt{1 + \frac{4 \cdot D^2}{K}}}{2} \quad (2.8)$$

where:

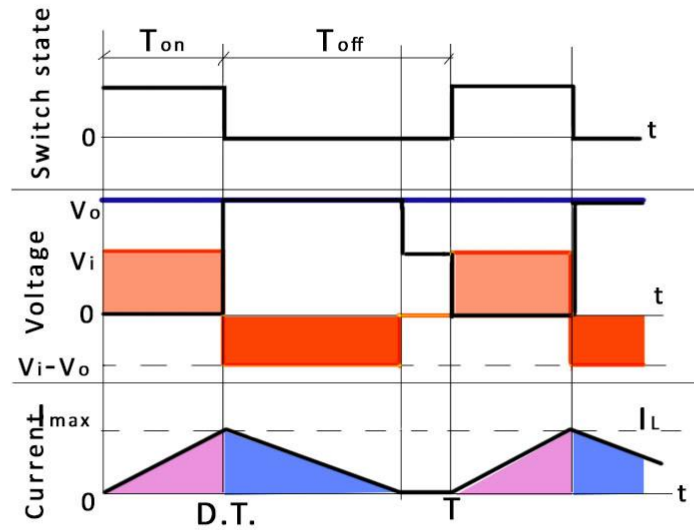
$D$  is the duty cycle

$K$  is equal to  $\frac{2L}{RT}$  [14]

$T$  is the switching period (s)

$L$  is the inductance of the coil (H) and

$R$  is the load resistance (Ohm).



**Fig. 2-10.** Current and voltage waveforms in discontinuous conduction mode

[14].

## 2.5.2 DC/DC Buck-type power converters

In Buck-type DC/DC power converters, the output voltage  $V_o$  is less than the input voltage  $V_i$ . The quotient of the time called  $T_{on}$  (the time during which the switch is ON) to the total time  $T$  of



the operating period, is called Duty Cycle ( $D$ ) of the switch and its value is given by the following equation [15]:

$$D = \frac{T_{on}}{T_{on} + T_{off}} \quad (2.9)$$

where:

$V_o$  is the output voltage (V) and

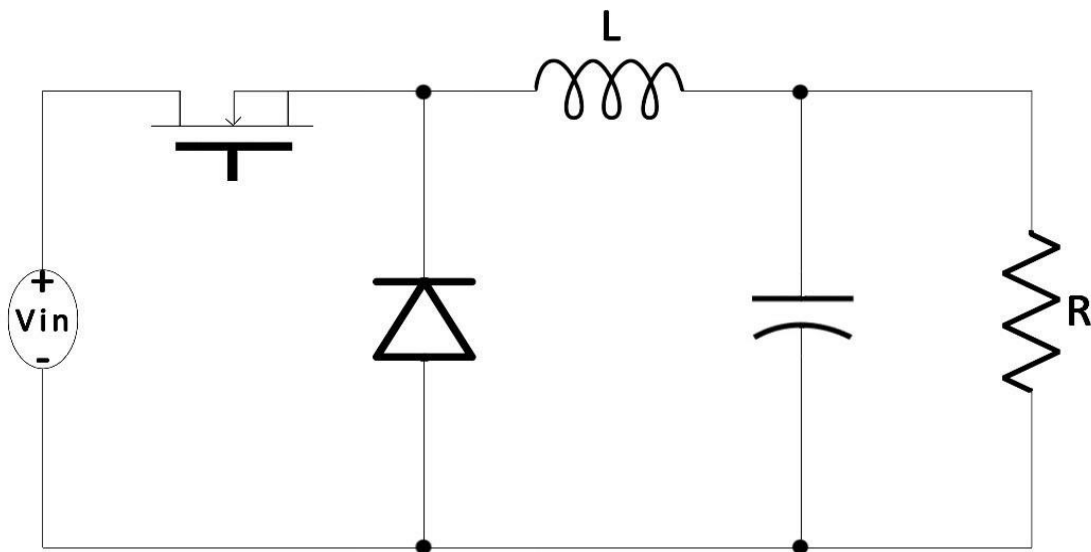
$V_i$  is the input voltage (V).

In Fig. 2-11 the buck converter electronic circuit is presented. There are two operating modes in a Buck-type power converter. When the switch is open (OFF-state), the flyback diode creates a closed current loop, in order to discharge the stored energy in the load [15]. When the switch is closed (ON-state), the diode is reversed polarized, and the current flows through the inductor in order to charge it. The output voltage of the converter is given by the following equation:

$$V_o = D \cdot V_{in} \quad (2.10)$$

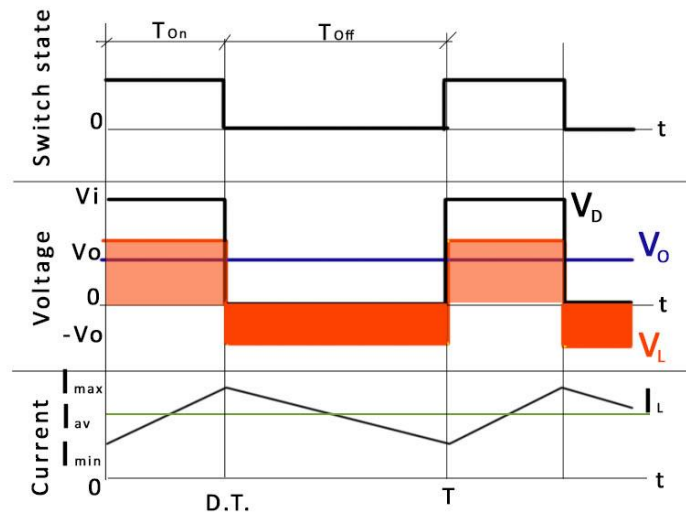
where:

$D$  is the duty cycle ( $0 \leq D \leq 1$ ).

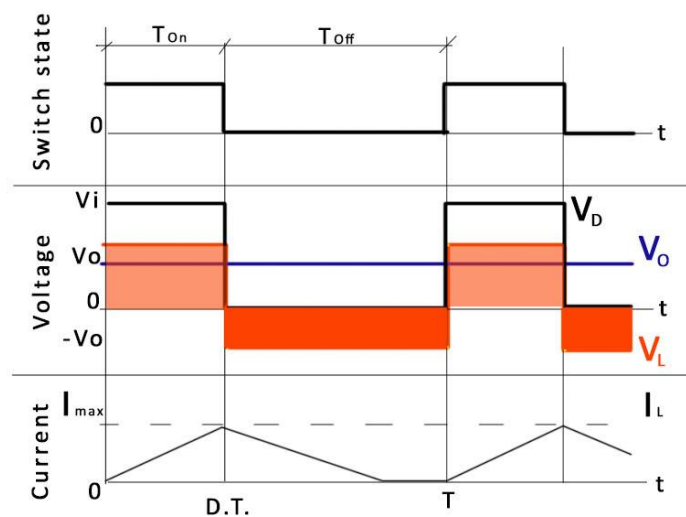


**Fig. 2-11.** Buck-type power converter circuit.

In Figs. 2-12 and 2-13 the waveforms of current and voltage for both operation states, during continuous conduction mode and discontinuous conduction mode, respectively, are presented.



**Fig. 2-12.** Current and voltage waveforms in continuous conduction mode of a Buck-type DC/DC power converter.



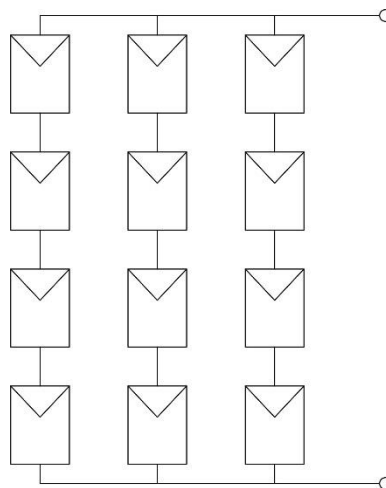
**Fig. 2-13.** Current and voltage waveforms in discontinuous conduction mode of a Buck-type DC/DC power converter.

## 2.6 PV array topologies

Four types of PV array topologies have been developed [16]:

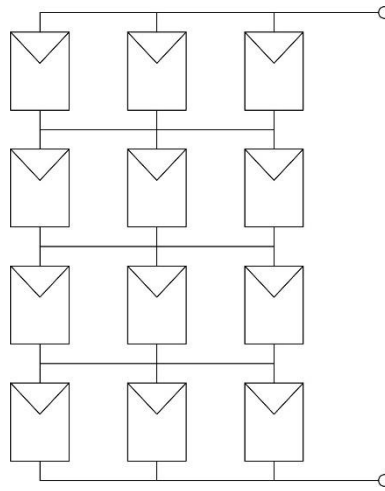
- Series-parallel topology (SP)
- Total-cross-tied topology (TCT)
- Bridge-linked topology (BL)
- Honey-comb topology (HC)

The SP PV topology [16], [17] is used very often, because of its low financial cost and its simplicity. It is obtained, by connecting solar modules in series and creating a string. In order to increase the total current of the PV array, many strings are connected in parallel. The SP PV topology is presented in Fig. 2-14.



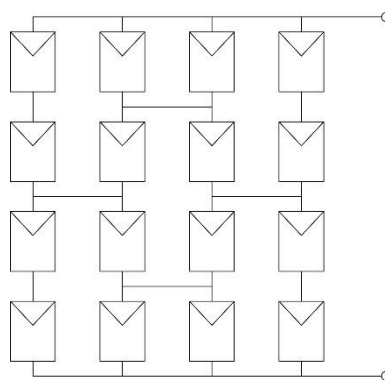
**Fig. 2-14.** The series-parallel PV array topology.

In the TCT PV topology, strings are created by parallel-connected PV modules and then, all of those strings are connected in series. In TCT PV topology, the currents are summed up and the voltages are equal. This type of connection has two advantages and one disadvantage. Advantages: (a) the overall effect of PV modules mismatch is reduced and (b) the losses of the bypass diodes are decreased. Disadvantages: (a) a lot of wires are required, which increase the financial cost [17], [18]. The TCT PV topology is presented in Fig. 2-15.



**Fig. 2-15.** The TCT PV array topology.

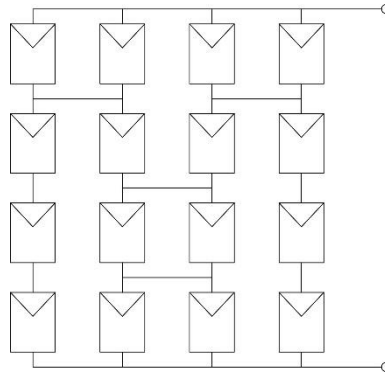
The BL PV topology is very similar with the TCT connection. Its main difference is that half of the TCT connections are removed. It is being called Bridge-Linked, because the modules are connected in bridge rectifier architecture. The mismatching power loss of BL PV topology is less than in the SP PV topology, because of its smaller number of modules that are connected in series. On the contrary, the BL PV topology has a higher number of modules connected in series with respect to the TCT PV topology, thus leading to an increment of the mismatching power losses [17], [18], [19]. This type of connection requires less wiring compared with the other two PV topologies, which results in a lower financial cost. The BL topology is presented in Fig. 2-16.



**Fig. 2-16.** The BL PV array topology.

The HC PV topology is a modified structure of the BL PV topology. The PV modules are connected similar to the hexagon shape of the honey-comb architecture [19]. It combines the advantages of the BL and TCT PV topologies (low financial cost, small effect of the PV modules mismatch) [17], [20]. The HC PV topology has a higher number of series connections than the BL and TCT PV topologies, which leads to an increment of the mismatching power losses [19],

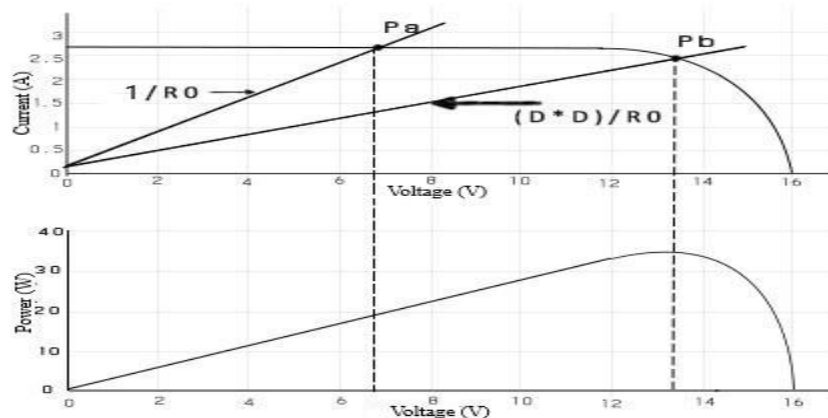
[20]. The mismatching power losses of the SP PV topology are more than the power losses of the HC PV topology, because of the higher number of modules connected in series. It requires the same number of wires with the BL PV topology, but in a different way. The HC PV topology is presented in Fig. 2-17.



**Fig. 2-17.** The HC PV array topology.

## 2.7 Maximum Power Point Tracking techniques

In Fig. 2-18 a constant load  $R_O$  is presented by the line having slope  $1/R_O$ . If  $R_O$  is connected to the PV array output then power  $P_a$  will be generated, which is different than the power  $P_b$ , which is the maximum power that can be obtained by the PV array[21]. For that reason, the PV arrays and the load are connected with a power converter. This converter adapts the load to the array, so that the maximum power is produced by the PV array. The role of the power converter duty cycle is to help the PV system to reach the peak power point [21]. This can be achieved by continuously modifying the duty cycle value, until the MPP is reached.

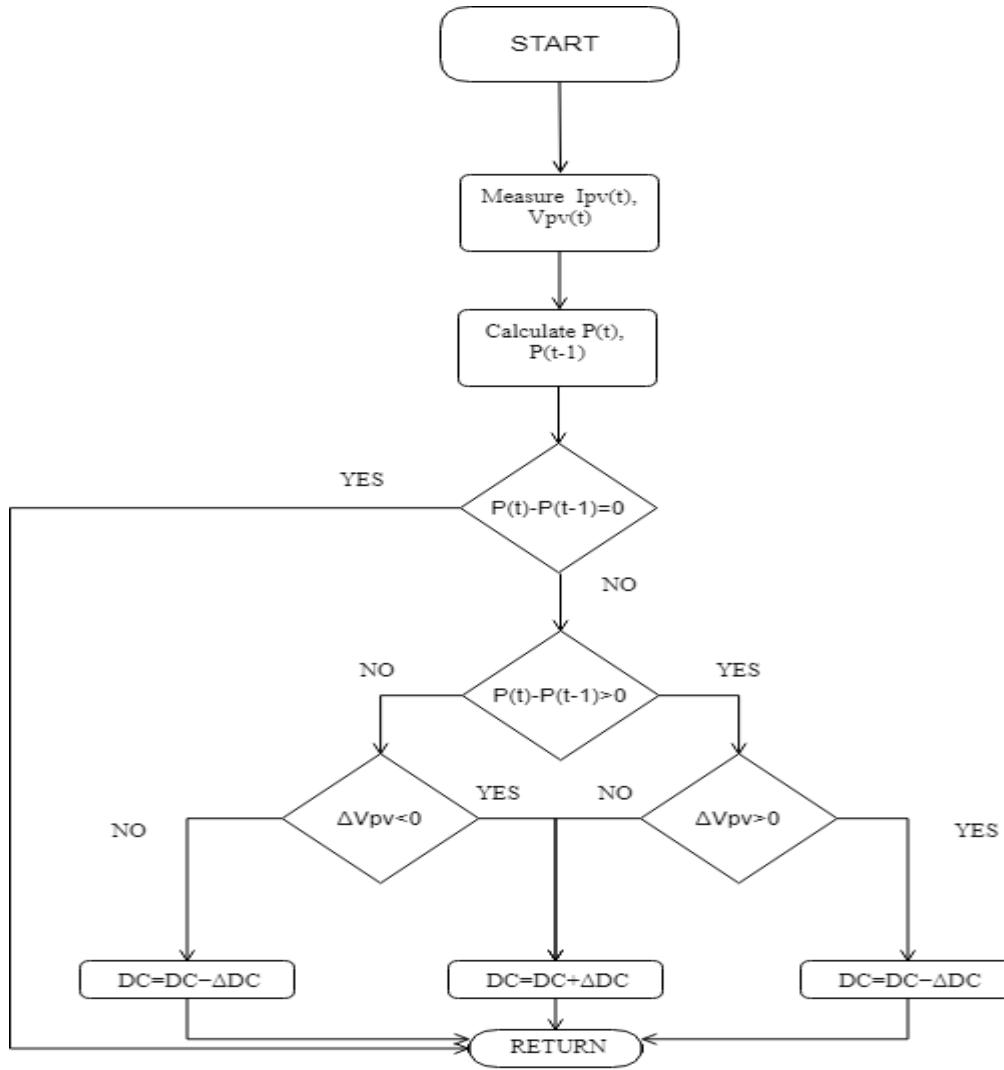


**Fig. 2-18.** Intersection between the load line and the power-voltage and current-voltage curves [21].

Many techniques have been developed for implementing the MPPT process. Among them, the Perturbation and Observation (P&O), Incremental-Conductance and Particle Swarm Optimization algorithms will be introduced in the following.

### **2.7.1 Perturbation and observation MPPT algorithm**

In this method, the PV array voltage and current have to be sensed. The PV array voltage is varied in order to control the output power of the PV system. If the PV array output voltage and power are increasing then the duty cycle keeps increasing. Otherwise, the MPPT algorithm starts decreasing the duty cycle. Similarly, when the PV array voltage is decreased and the PV array output power is increased, then the duty cycle is decreased. This process is continued, until the peak power point of the PV array is reached. The flowchart of this method is presented in Fig. 2-19 [22].



**Fig. 2-19.** P&O MPPT technique flowchart.

### 2.7.2 Incremental-Conductance MPPT algorithm

The method of Incremental Conductance is based on the fact that the slope of the power-voltage curve of the photovoltaic array at the MPP is zero. It is positive on the left-hand side of the MPP and negative to the right-hand side of the MPP, and it is increased or decreased respectively. The mathematical equations of this method can be described as follows [22], [23]:

- $\frac{dP}{dV} = 0$  at the MPP (2.11)

- $\frac{dP}{dV} > 0$  at the left side of the MPP (2.12)

- $\frac{dP}{dV} < 0$  at the right side of the MPP (2.13)

where:

$dV$  is the difference between previous and current PV voltage values

$dI$  is the difference between previous and current values of the PV current

$dP$  is the difference between previous and current PV power values.

$V$  is the voltage value (V) and

$I$  is the current value (A).

However:

$$\frac{dP}{dV} = \frac{d(IV)}{dV} = I + V \cdot \frac{dI}{dV} \quad (2.14)$$

Using (2.14), (2.11) - (2.13) are transformed to:

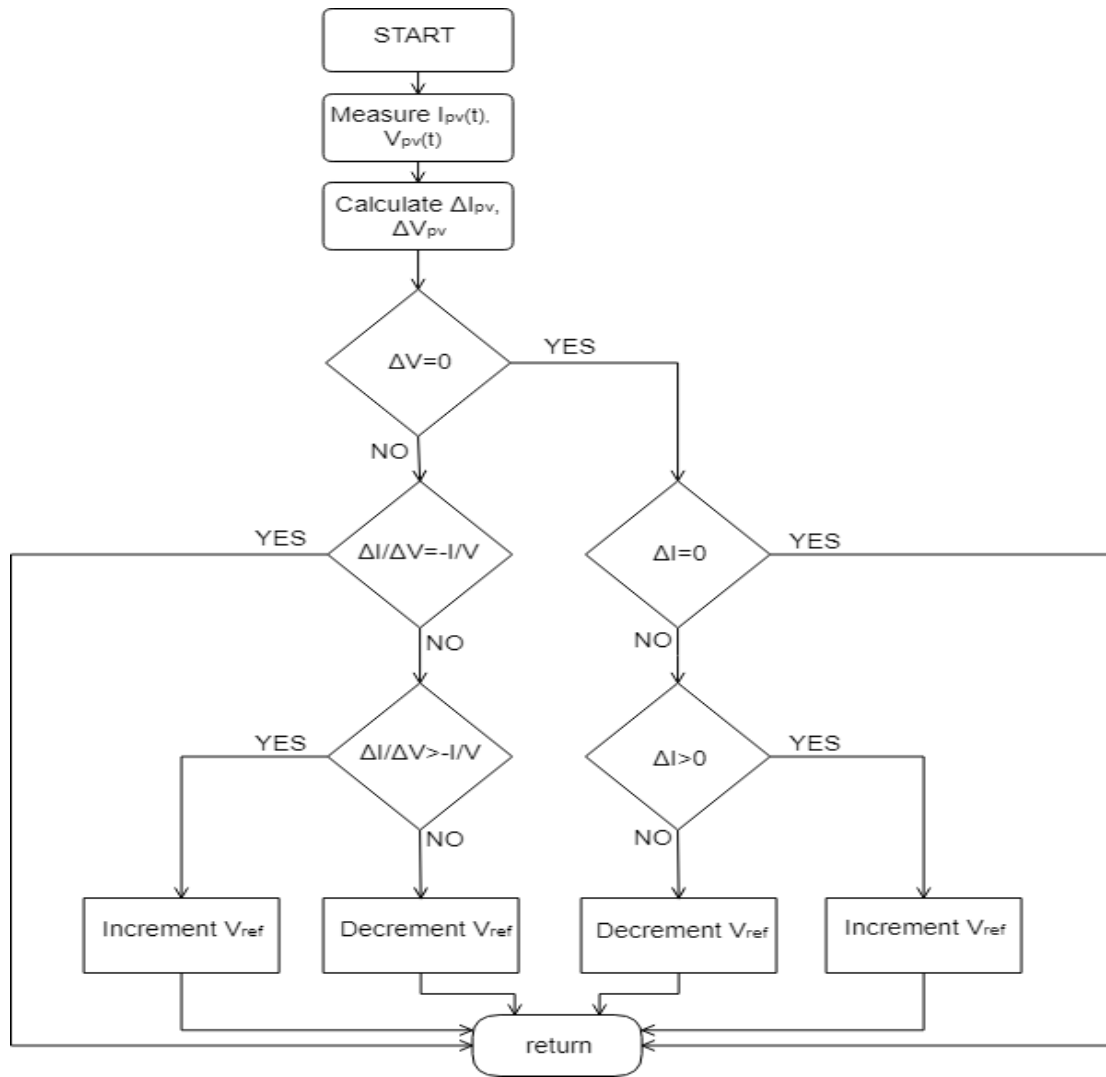
- $\frac{dI}{dV} = -\frac{I}{V}$  at the MPP (2.15)

- $\frac{dI}{dV} > -\frac{I}{V}$  at the left side of the MPP (2.16)

- $\frac{dI}{dV} < -\frac{I}{V}$  at the right side of the MPP (2.17)

The flowchart of this method is presented in Fig. 2-20.





**Fig. 2-20.** Flowchart of the Incremental-Conductance MPPT algorithm.

### 2.7.3 Particle Swarm Optimization MPPT algorithm

The Particle Swarm Optimization MPPT algorithm executes the following steps [24]:

- The duty cycle of the power converter is defined as the particles position. The power generated by the PV array is the fitness value evaluation function. A large number of particles leads to a more accurate detection of the GMPP however, this requires more time.
- During the initialization, the particles are placed on a specific position. If there is any prior knowledge of where the GMPP is located, then all the particles should be placed around it. In addition, it should be mentioned that the fixed particles should be placed between the range  $[D_{\min}, D_{\max}]$ , where  $D_{\max}$  and  $D_{\min}$  are the maximum and minimum values of the duty cycle, respectively.

- Every duty cycle value corresponds to a pair of current and voltage values. This pair is used in order to calculate the power generated by the PV array, which is the fitness value of every particle. It should be noted that in order to acquire correct samples, the time interval between successive particle evaluations has to be larger than the power converter's settling time.
- A criterion for the evaluation of how close a given solution is to the optimum solution of a problem should be defined. In the MPPT problem this criterion is the generated power of the PV system and it is defined as the fitness function.
- For every particle, the fitness value with the highest value is called  $P_{\text{BEST},i}$ . If a particular particle finds out in the future a higher fitness value, then this value is set as the new  $P_{\text{BEST},i}$ . After that, the particle with the highest fitness value than every other particle is chosen as  $G_{\text{BEST}}$ .
- Then, after the evaluation of every particle, the position and velocity of each particle are updated. The velocity and position of each particle are defined by:

$$U_i(k+1) = w \cdot U_i(k) + c_1 \cdot r_1 \cdot (P_{\text{best},i} - x_i(k)) + c_2 \cdot r_2 \cdot (g_{\text{best}} - x_i(k)) \quad (2.18)$$

$$x_i(k+1) = x_i(k) + u_i(k+1) \quad (2.19)$$

where:

$x_i$  is the position of particle  $i$

$u_i$  is the velocity of particle  $i$

$k$  is the iteration number and

$w$  is the inertia weight

$r_1, r_2$  are random variables uniformly distributed within  $[0,1]$  and

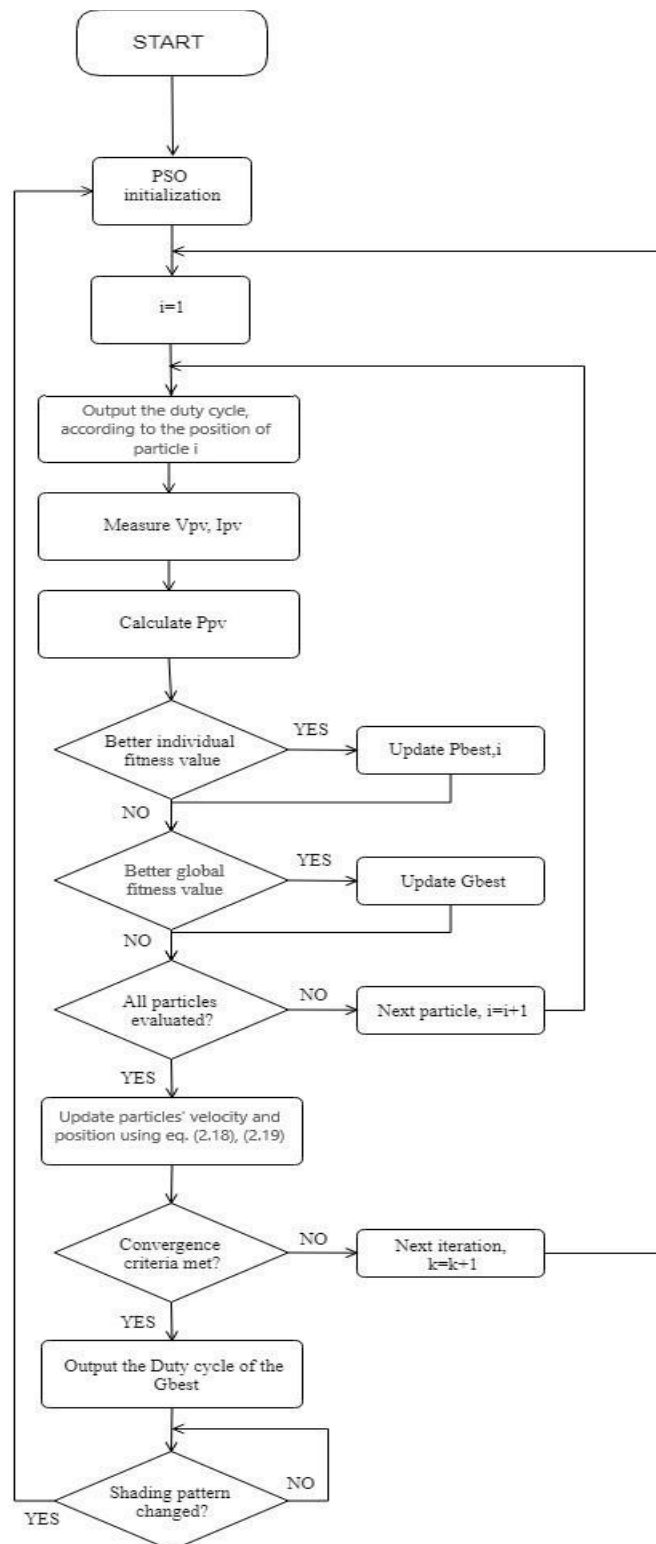
$c_1, c_2$  are the cognitive and social coefficient.

The parameter settings of the PSO-based MPPT method are presented in Chapter 5. If the velocity of every particle becomes smaller than a threshold, or if the maximum number of iterations is reached, the MPPT algorithm will stop and output the obtained  $G_{\text{BEST}}$  solution [24].

The algorithm is reinitialized when the following equation is satisfied:

$$\frac{|P_{pv,new} - P_{pv,last}|}{P_{pv,last}} \geq \Delta P(\%) \quad (2.20)$$

The flowchart of the PSO MPPT method is presented in Fig. 2-21.



**Fig. 2-21.** PSO MPPT technique flowchart.

# 3.

## Maximum Power Point Tracking using Q-learning

---

### 3.1 Theoretical analysis

Q-learning is a model-free (MF) reinforcement learning algorithm. MF algorithms have the characteristic that they are able to solve a problem without using the transition probability distribution associated with the Markov decision process. Thus, a MF algorithm is considered as an “explicit” trial and error algorithm [25].

It is considered that a computational agent is able to explore a discrete, finite world. During its exploration, in every time-step, it can select one action from a finite collection of actions. The agent is considered as a controller of a controlled Markov process. Depending on the state and the action, the agent receives a reward  $r_n$ . Every state changes probabilistically to  $y_n$ , according to the following equation [26]:

$$\Pr ob[y_n \equiv y | x_n, a_n] = P_{x_n y[a_n]} \quad (3.1)$$

The target of the algorithm is to help the agent to maximize the total discounted expected reward, by determining the most optimal policy. Discounted reward is the reward that the agent will receive in  $s$  time-steps. It is used in order to balance immediate and future rewards. Therefore, the future rewards are worth less than immediate rewards. Its less worth depends on one factor called discount factor ( $\gamma$ ). In  $s$  steps this future reward is multiplied by  $\gamma^s$  ( $0 < \gamma < 1$ ). Under a specific policy  $\pi$ , the value of state  $x$  is calculated as follows [26]:

$$V^{\pi}(x) \equiv R_x(\pi(x)) + \gamma \sum_y P_{xy}[\pi(x)] V^{\pi}(y) \quad (3.2)$$

The agent chooses an action that a policy  $\pi$  recommends. By choosing this action it receives an immediate reward and moves to a specific state, whose worth is equal to  $V^{\pi}(y)$  and its probability is equal to  $P_{xy}[\pi(x)]$ . The theory of DP ensures that there is at least one optimal stationary policy  $\pi^*$  which is given by [26]:

$$V^*(x) \equiv V^{\pi^*}(x) = \max_a \left\{ R_x(a) + \gamma \sum_y P_{xy}[a] V^{\pi^*}(y) \right\} \quad (3.3)$$

where:

$R_x$  is the immediate reward

$P_{xy}[a]$  is the probability of moving from state  $x$  to state  $y$  under action  $a$

$a$  is the action

$\pi^*$  is the action policy and

$V$  is the state's value.

The Q-learning update function is defined by the following formula [25]:

$$Q(S_t, a_t) \leftarrow Q(S_t, a_t) + \alpha \cdot [R_t + \gamma \cdot \max_a Q(S_{t+1}, a_{t+1}) - Q(S_t, a_t)] \quad (3.4)$$

where:

$\alpha$  is the learning rate

$R_t$  is the immediate reward

$\gamma$  is the discount factor

$Q(S_{t+1}, a_{t+1})$  is the Q-value of the next state and

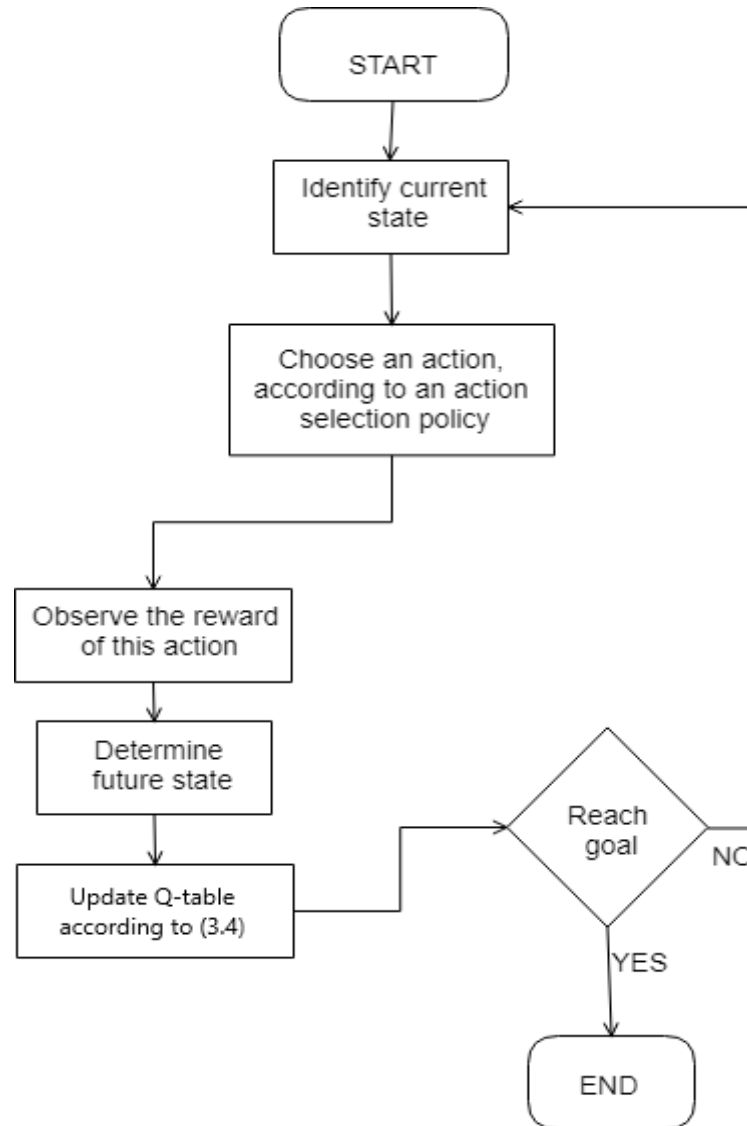
$Q(S_t, a_t)$  is the Q-value of the current state.

The **learning rate or  $\alpha$**  is defined as how much the agent accepts the estimate derived from the new value versus the old value. Thus, the difference between the new estimate and the old Q-value is calculated and then this value is multiplied by the learning rate. This result is added to  $Q(S_t, a_t)$ , which is the previous Q-value.

**Discount factor or  $\gamma$**  is used to balance immediate and future reward. Using (3.4), it is observed that a discount to the future reward is applied. The range of this factor is usually between 0.8 and

0.99. **Reward** is the value that the agent will receive after choosing a specific action at a given state [27].

A flowchart of the Q-learning method is presented in Fig. 3-1.



**Fig. 3-1.** A flowchart of the Q-learning algorithm.

All of the above can be described in the following simple steps:

**Step 1:** The agent identifies the state that it has visited. A discretized state-space is necessary in order to help the agent not to confuse different states as identical.

**Step 2:** The agent chooses a proper action, according to the action policy selection. The actions are the execution acts that enable the agent to do exploration within the state-space.

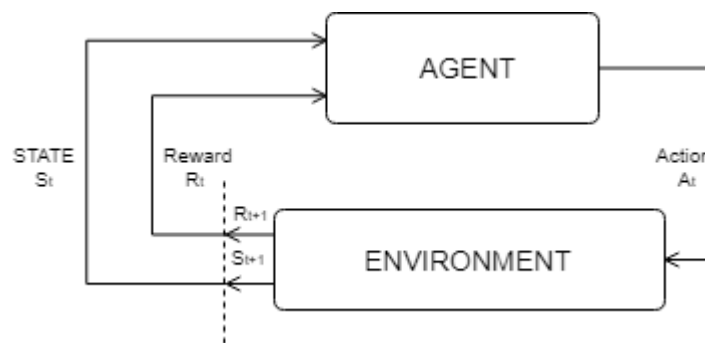
**Step 3:** The agent receives a reward, depending on the result of the action that it has chosen.

**Step 4:** The agent determines its future action and the Q-table is updated, according to (3.4).

**Step 5:** The agent checks if the goal has been achieved.

### 3.2 Q-learning for PV MPPT

The PV MPPT problem is a discrete time stochastic control process. It provides a mathematical framework for modelling decision making in situations where outcomes are partly random and partly under the control of a decision maker. For this reason, it can be considered as a Markov Decision Process (MDP) [21]. A MDP is used to model the interaction between the agent and the controlled environment. It consists of: (a) the state-space, (b) the set of all possible actions and (c) the reinforcement reward function, which represents the reward when applying the action  $a$  in the state  $s$  which leads to the state  $s'$  [21]. MDPs are useful for studying optimization problems, such as the MPPT problem. One way to solve this kind of problems is the use of a temporal difference Q-learning algorithm. Q-learning's goal is to find an action selection policy, which will maximize the total discounted rewards it receives over the future. The simplest form of a Q-learning algorithm is presented in Fig. 3-2.



**Fig. 3-2.** One simple form of Q-learning algorithm.

It was mentioned before, that a Q-learning algorithm learns from a set of observations: (a) states, (b) actions and (c) rewards. In Section 3.1, it has been mentioned that Q-learning is defined by (3.4),  $Q(S_t, a_t)$  is the current value of the action value function that will be updated. In each time-step, the agent observes its current state. After that, it chooses its action, according to its action selection policy  $\pi^*$ . After the action selection, it observes its future state ( $S_{t+1}$ ). Then, it receives an immediate reward  $R_t$  and selects the future Q-value with the maximum value

over actions,  $\max_a Q(S_{t+1}, a)$ . Learning rate  $\alpha$  determines, the effect of the estimate based on the new Q-value over the last Q-value in the last Q-update.

### 3.2.1 Action policy selection

In order to ensure that every Q-value that has been calculated is the optimal one, every pair of state-action must be estimated. For each state that the agent has visited, all possible actions should be selected. This can be achieved by selecting the most suitable action policy. In this study, Boltzmann exploration has been used. Boltzmann exploration is a classic strategy for sequential decision-maker under uncertainty and is one of the most standard tools in Reinforcement Learning [28]. Instead of always taking the optimal action or taking a random action, this approach involves an action with weighted probabilities. Each probability is calculated by the following formula:

$$p(S, a_i) = \frac{e^{\frac{Q(S, a_i)}{\tau}}}{\sum_{a_i} e^{\frac{Q(S, a_i)}{\tau}}} \quad (3.5)$$

where:

$\tau$  is the Boltzmann exploration's parameter

$a_i$  is the current action and

$Q(S, a_i)$  is the Q-table's cell current value.

The factor  $\tau$  is called temperature. It is a positive value which controls the randomness of the action selection. If temperature has a high value, then all the probabilities values are similar. This concludes to a random action selection, which depends on a factor, called  $N_r$ .  $N_r$  is a random number between the range (0, 1). As long as visits  $N$  are increased in a specific state, then temperature  $\tau$  is decreased. After a certain number of visits  $N$ ,  $\tau$  is becoming equal to its minimum value. This means that exploration is over and the agent chooses greedily the action with the highest Q-value. The temperature function is presented in the following formula:

$$\tau = \begin{cases} \tau_{min} + \left(1 - \frac{N}{N_{max}}\right) \cdot (\tau_{max} - \tau_{min}), & \text{if } (N \leq N_{max}) \\ \tau_{min}, & \text{if } (N > N_{max}) \end{cases} \quad (3.6)$$

where:

$\tau_{min}$  is the minimum temperature value



$\tau_{max}$  is the maximum temperature value

$N$  is the number of visits and

$N_{max}$  is the maximum number of visits.

### 3.2.2 State-space

The state-space in our problem is formed with two different ways. In both ways, there is a discretized state-space [29]. In the first way, the state depends on the present duty cycle of the DC/DC boost-type power converter and the power generated by the PV plants:

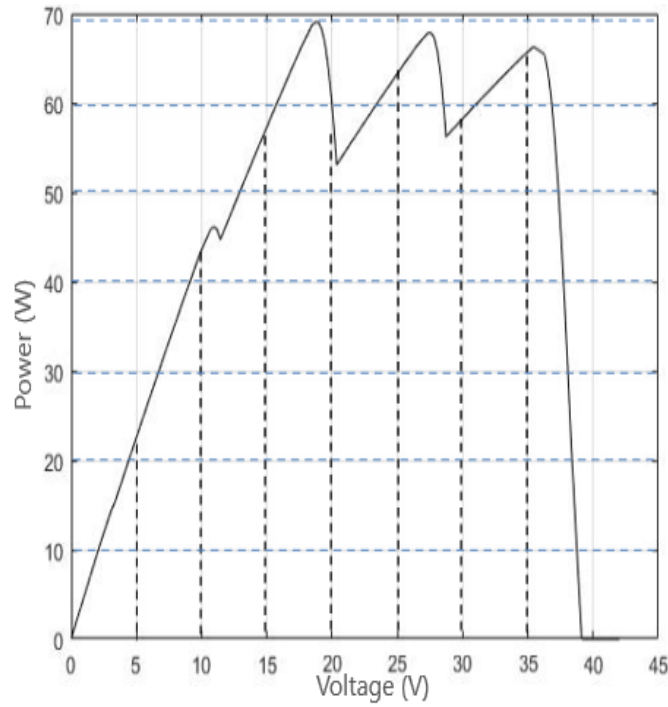
$$state = \{S \mid S_{i,j} = (D_i, P_j), i \in [1, 2, 3 \dots n], j \in [1, 2, 3 \dots m] \mid \} \quad (3.7)$$

where:

$n$  is the number of equal quantized steps of the duty cycle (D) and

$m$  is the number of equal quantized steps of the PV power (P).

Two points should be mentioned. The first one is related to the duty cycle. The duty cycle range is between (0, 1), but practically it can't take values near 1 or near 0. That's the reason its range was formed as (0.19, 0.88). The second point is related to the numbers of the equal quantized steps of  $n$ ,  $m$ . Large numbers of  $n$ ,  $m$  will conclude to a better approach of the MPP, but it will also conclude in a longer learning time. On the other hand, small numbers of  $n$ ,  $m$ , will conclude to a worse approach of the MPP in a shorter learning period. A trade-off is created, which is associated with the convergence time and accuracy. In Fig. 3-3 the discrete state-space is described, using two parameters (duty cycle steps are the vertical black dotted line and PV power steps are the horizontal blue dotted lines). In this example, the duty cycle is defined by 8 equal quantized steps and the PV power is defined by 7 equal quantized steps. In every duty cycle value, the state is determined jointly by the step that the duty cycle belongs to and by the step that the PV array output power belongs to. It should be noted that each value of the duty cycle corresponds to a specific value of the PV array output voltage, according to (2.7). Assuming that a duty cycle value corresponds to a PV array output power value equals to 53 W and a voltage value equals to 18 V, it is assumed that the agent's state is equal to (6,4), because 53 W belong to the 6<sup>th</sup> step of the PV power and 18 V belong to the 4<sup>th</sup> step of the voltage.



**Fig. 3-3.** Discrete state-space using duty cycle's steps and PV array output power's steps.

In the second way of the state-space definition, there is one more parameter which is taken into account. The duty cycle of a previous time-step. This extra information is being used in order to check if a better accuracy can be achieved. This concludes in a longer learning time. It is described by the following formula:

$$state = \{S \mid S_{i,j} = (D_i, P_j, DO_k), i \in [1, 2, 3 \dots n], j \in [1, 2, 3 \dots m], k \in [1, 2, 3 \dots p] \} \quad (3.8)$$

where:

$n$  is the number of equal quantized steps of the duty cycle (D)

$m$  is the number of equal quantized steps of the PV power (P) and

$k$  is the number of equal quantized steps of the of duty cycle of a previous time-step (DO).

### 3.2.3 Action-space

The PV system's controller action-space is expressed by the following set:

$$A = \{a \mid +0.04, -0.04, +0.12, -0.12, +0.28, -0.28, 0\} \quad (3.9)$$

where:

+0.04, -0.04, +0.12, -0.12, +0.28, -0.28, 0 indicate changes over the current duty cycle.

It must be mentioned that +0.04 is action number 1, -0.04 is action number 2, +0.12 is action number 3, -0.12 is action number 4, +0.28 is action number 5, -0.28 is action number 6 and 0 is action number 7. A wide variety of different actions is chosen in order to ensure that most of the duty cycle values are selected. This is necessary in order to minimize the probability that the agent will converge at a wrong peak. At the beginning of the implementation of the method, only actions +0.04, -0.04 and 0 were adopted. However, it was observed that under non-uniform solar irradiation conditions the agent was unable to detect the global MPP (GMPP), because it was concluding that the first peak detected, was the maximum one. Thus, actions +0.12 and -0.12 were also added to the list of the possible actions. However, again it was observed that under conditions the agent was unable to detect GMPPs, which were not near the initial state. For these reasons, the actions +0.28 and -0.28 were added to the actions list. Thus, it has been observed that the agent is capable of detecting the GMPP in most of the times.

If duty cycle is equal to 0.64 and the agent selects action number 3, then the next duty cycle value will be equal to  $0.64 + 0.12 = 0.76$ . A maximum (0.83) and a minimum (0.19) value for duty cycle was defined (i.e. if duty cycle's value is equal to 0.83 and actions 1, 3 or 5 are selected, then duty cycle will remain to 0.83 in the next time-step). It is worth mentioning that if the agent selects action number 7 instead, then it will remain in the same state during the next time-step. Each action is selected according to Boltzmann exploration, which was analyzed in Section 3.2.1. When exploration is over, the action with the highest Q-value is selected for that particular state. It is considered that the GMPP of the PV array is detected, when action number 7 ("no-change" in duty cycle's value) has the highest Q-value. In this case, Q-values for actions number 1, 2, 3, 4, 5 and 6 are smaller, due to the reduction of the PV array output power, by the selection of the corresponding actions compared to the Q-value for action number 7. All of the above lead to the conclusion that the GMPP is the point where the agent is located at. The duty cycle is updated according to the following equation:

$$D_{t+1} = D_t + a \quad (3.10)$$

where:

$D_{t+1}$  is the duty cycle for the next time-step

$D_t$  is the current duty cycle and

$a$  is the increment, decrement or no-change in duty cycle value.

As an increment, actions 1, 3 and 5 are considered. As a decrement, actions 2, 4 and 6 are considered. “No change” is the action 7. If duty cycle reaches its minimum value that was referred in paragraph 3.2.2 and actions 2, 4 or 6 are selected, then the duty cycle’s value will not change. Respectively, if the duty cycle’s value reaches its upper bound and the actions 1, 3 or 5 are selected there will be no change to its value. In every state that the agent is visited, all actions need to be selected in order to help the agent to learn the optimal path for that specific MPP. This is the reason that there is a gradation in selections for the increment and for the decrement. This happens because, under partial shading, there are many peaks that have to be explored. There is a probability of the existence of two peaks with not a great difference in their values. If these peaks are very close to each other, then a small change in duty cycle is necessary in order to help the agent understand which peak is the highest. On the other hand, if those peaks are not close enough to each other, the agent will need another action to move as close as possible to the other peak. If this action does not exist, then the agent will consider that the first peak is the highest because it did not manage to visit another higher peak.

### 3.2.4 Reward

The criterion for the selection of the most suitable action is the reward ( $R_t$ ). One pair of current-voltage values corresponds to each duty cycle value. For these pairs, a specific power value is generated by the PV array. This value helps the agent to decide the reward that will evaluate the action selection according to the following equation:

$$R_t = \begin{cases} +2, & \text{if } P(t+1) - P(t) > +1 \\ 0, & \text{if } -1 \leq P(t+1) - P(t) \leq +1 \\ -2, & \text{if } P(t+1) - P(t) < -1 \end{cases} \quad (3.11)$$

where:

$R_t$  is the immediate reward

$P(t)$  is the PV power value for the current time-step (W) and

$P(t+1)$  is the PV power value for the next time-step (W).

If the PV power value in time-step  $t+1$  minus the PV power value in time-step  $t$  is higher than 1, then a small positive reward will be given for that specific pair of state-action, in order to “encourage” the agent to keep this map-selection in the next time-step that will visit this state. If

the difference that was mentioned before is very small or even equal to zero, it is considered that there is no change in the generated power, and the reward will be equal to zero. Usually, this reward is referred to action 7, where the duty cycle remains constant. However, there is a small probability of the existence of points with different duty cycle and almost the same power. Lastly, if the difference of the generated power, for these two time-steps, is smaller than -1, then a small negative reward will be given to the agent. This happens, because a very small negative Q-value is necessary when there is a PV array output power decrease. In this way, in the next time-step that the agent will visit this state, the action that caused the power loss will have a small probability of reselection, according to (3.5). This will result in a selection of actions with power increment even during the exploration.

### 3.2.5 Discount factor, learning rate

In section 3.1, the discount factor was defined as a value which is used to balance immediate and future rewards. The agent is trying to maximize the total future rewards, instead of the immediate reward. It is used to adjust the weight of the reward of the current optimal value and indicate the significance of future rewards [25]. The discount factor is necessary in order to help the agent to find the optimal path of future actions and only not the current optimal action.

It has to be ensured that the Q-learning-based MPPT method will converge. In order to achieve this, a visit frequency criterion was considered. The agent has to be encouraged in order to consider that knowledge acquired from states with a small number of visits is more important than knowledge acquired from states with a higher number of visits. Thus, states that have been visited less times will have a higher learning rate than states that have been explored more times. This is accomplished by the following equation:

$$\alpha = \frac{k_1}{k_2 + k_3 \cdot N} \quad (3.12)$$

where:

$k_1$ ,  $k_2$  and  $k_3$  are factors which determine the initial learning rate value for a state with no visits

$N$  is the number of times that the agent visited the specific state and  $\alpha$  is the learning rate.

### 3.2.6 Q- learning-based algorithm process

The MPPT process is performed as follows:

**Step 1:** Q-table initialization. According to Section 3.2.2, a Q-table is created using four or three dimensions, according to (3.7) and (3.8), respectively. The dimensions' selection depends on the location that the photovoltaic installation will take place (advantages and disadvantages of each implementation of the Q-learning-based MPPT methods are being analyzed in Chapters 5 and 6). This table is filled with zeros because there is no prior knowledge. In addition, every parameter that was mentioned in Sections 3.2.1-3.2.5 is defined ( $\tau_{\min}$ ,  $\tau_{\max}$ , discount factor,  $k_1$ ,  $k_2$ ,  $k_3$ , maximum number of visits,  $D_{\max}$ ,  $D_{\min}$ , learning rate etc.). Lastly, a table which includes each state's number of visits is defined. The size of this table depends on the number of the equal quantized steps of the duty cycle of the current time-step and the power generated by the PV array, or the current duty cycle, the power generated and the duty cycle of a previous time-step, according to the case that is being studied.

**Step 2:** After the PV array voltage and current signals reach the steady-state, the PV array generated power is calculated. Then, the state is determined.

**Step 3:** The learning rate is calculated by (3.12).

**Step 4:** The temperature  $\tau$  is calculated by (3.6), and the probability of every possible action for the specific state is also calculated by (3.5).

**Step 5:** If the number of visits of this state is higher or equal to 20 and the Q-value for action 7 (constant duty cycle) is the maximum compared with the others, then the P&O algorithm is executed. In other words, if that state is the GMPP that the algorithm is trying to detect, the P&O algorithm is executed in order to help the agent to make a better approach of the that point. The execution of the P&O algorithm is also necessary when there is a steep change of solar irradiance or ambient temperature.

**Step 6:** If the previous decision is not true, a number in the range (0, 1) is randomly selected. Then, this number is compared with the probabilities which were mentioned in step 4. Each probability encodes a specific duty cycle change, as indicated in the following formulas:

- $0 < \text{Random number} \leq \text{Prob1}, \quad \text{action 1}$
- $\text{Prob1} < \text{Random number} \leq \text{Prob1} + \text{Prob2}, \quad \text{action 2}$

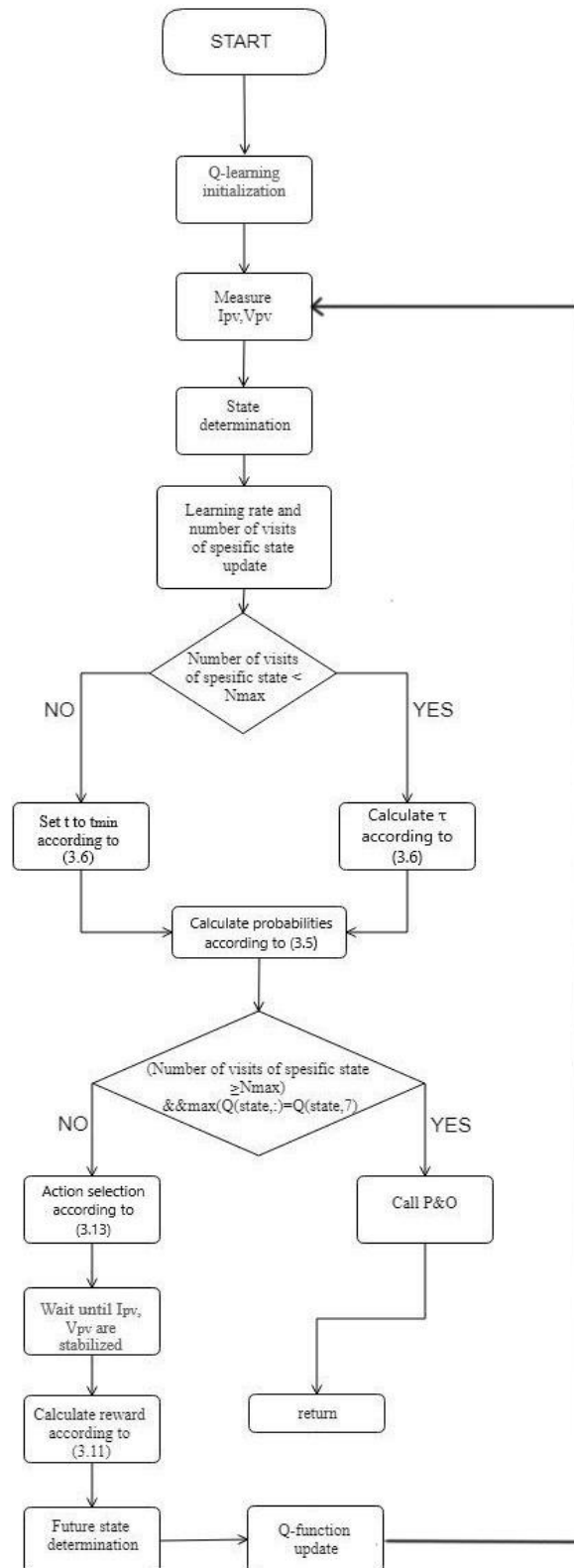
- $Prob1 + Prob2 < Random\ number \leq Prob1 + Prob2, + Prob3$ , action 3
  - $Prob1 + Prob2, + Prob3 < Random\ number \leq Prob1 + Prob2, + Prob3 + Prob4$ , action 4
  - $Prob1 + Prob2, + Prob3 + Prob4 < Random\ number \leq Prob1 + Prob2, + Prob3 + Prob4 + Prob5$ , action 5
  - $Prob1 + Prob2, + Prob3 + Prob4 + Prob5$ , action 5  $< Random\ number \leq Prob1 + Prob2, + Prob3 + Prob4 + Prob5 + Prob6$ , action 6
  - else, action 7
- (3.13)

**Step 7:** After the change of the action, the algorithm waits until the PV array output voltage and the current signals reach the steady-state. Then, the power difference between the generated power of the previous time-step and current time-step is calculated. According to (3.11) the reward is calculated.

**Step 8:** Future state is determined and the Q-function is updated, according to (3.4).

**Step 9:** Return to step 2.

The flowchart of the Q-learning-based MPPT algorithm is presented in Fig. 3-4.



**Fig. 3-4.** Q- learning-based MPPT algorithm flowchart.



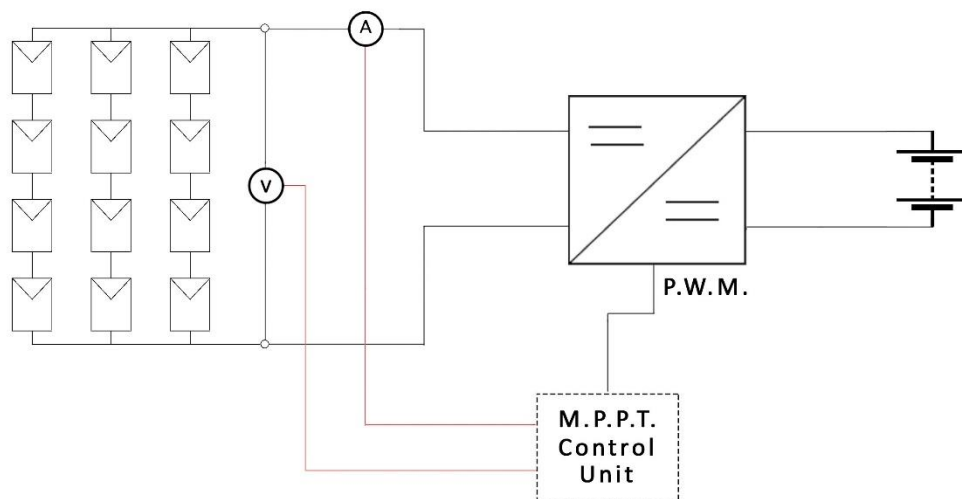
If the number of visits of a state is higher than or equal to 20, then the learned experience that is acquired by the agent seems to be mature and it can be exploited. During the exploration, the agent knows which action selection will provide the highest reward and this is the criterion by which, the agent will choose its future action. If the number of visits of a state is below 20, then it is considered that the agent is still trying to learn which action will provide to it the highest reward. This learning period is being called exploration period. During exploration, the agent is trying to learn the action selection policy that it will lead it to the GMPP. Practically, there is no termination condition and each step is continuously executed. In a real PV system, there must be a real period of time, in which the agent will learn how to react under real environmental conditions. After that period, the agent will have been trained and it is assumed that this method will have a high MPPT efficiency in a small number of time-steps.

# 4.

## The Simulink model of the PV system

---

The photovoltaic system under study in this diploma thesis consists of the following components: (a) a PV array connected according to the series-parallel topology, (b) an MPPT control unit, (c) a DC/DC Boost-type power converter and (d) a battery, which is connected to the power converter output. Each component was designed and implemented in the MATLAB<sup>TM</sup>/Simulink software. In Fig., 4-1 the overall PV system under study is presented.

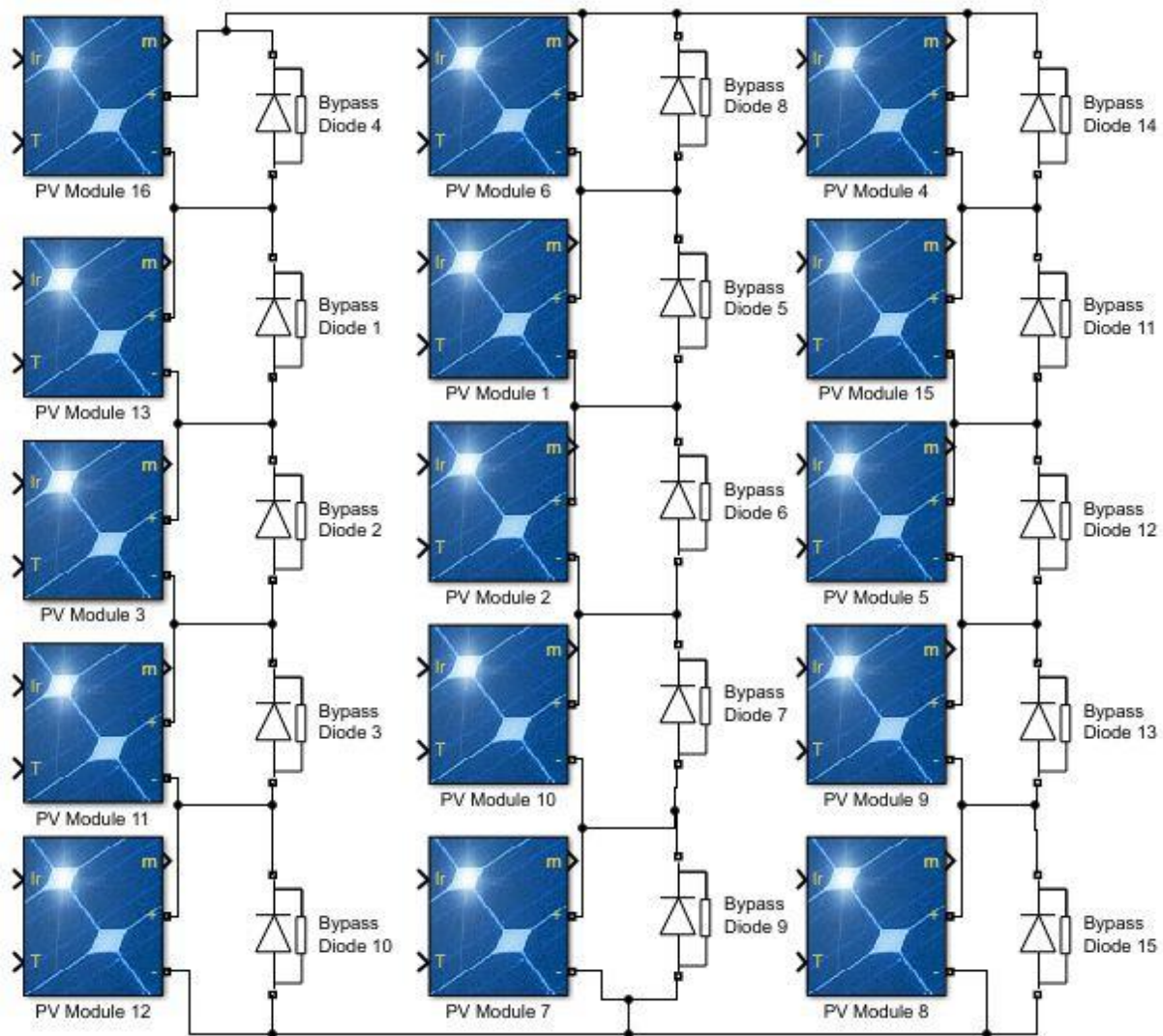


**Fig. 4-1.** The PV system under study.

## 4.1 Photovoltaic arrays

In Section 2.6, it was mentioned that a series-parallel topology is obtained, by connecting solar modules in series and creating a string. Then, many strings are connected parallel in order to increase the total output current of the PV array. In addition, bypass diodes are wired in parallel with each solar panel for protection of the solar panels during partial shading and during normal operation they are reversed biased. The PV topology that was designed includes three parallel PV strings consisting of five PV modules, which are in series connected.

The “PV array” block of “Simscape/Power systems/Specialized technology/Renewables/Solar” library was used. In Fig. 4-2, the topology that was designed in Simulink software is presented. The bypass diodes parameters are presented in Table 4-1 and the PV module data are presented in Table 4-2.



**Fig. 4-2.** The PV array topology that was developed in Matlab™/Simulink.

**Table 4-1. Bypass diodes operational parameters.**

Resistance $R_{ON}$ (Ohms)	0.001
Forward voltage $V_f$ (V)	0.8
Snubber resistance $R_S$ (Ohms)	500
Snubber capacitance (F)	$250^{-9}$

**Table 4-2. Operational characteristics of the photovoltaic modules.**

Maximum Power (W)	11.6
Open-circuit voltage, $V_{oc}$ (V)	7.25
Voltage at maximum power point $V_{MP}$ (V)	5.75
Temperature coefficient of $V_{oc}$ (%/deg.C)	-0.322
Cells per module ( $N_{cell}$ )	24
Short-circuit current $I_{sc}$ (A)	2.204
Current at maximum power point $I_{mp}$ (A)	2.016
Temperature coefficient of $I_{sc}$ (%/deg.C)	0.071996

Two arrays of temperature and solar irradiation values, respectively, were given as inputs for every solar module. These values are presented in Chapter 5.

## 4.2 DC/DC Boost-type power converter

A Boost-type power converter is used because the intrinsic impedance of PV panels has to be matched with the impedance of load. In this way, maximum power transfer can occur between them. MPPT is achieved by manipulating the duty cycle ratio of the converter, such that the PV array operates at its maximum power point.

The Boost converter was designed separately in order to ensure its correct operation. A DC voltage source was used as an input and a resistance was used as its output. In Table 4-3, the Boost converter operating parameters are presented. All of these parameters, were extracted from

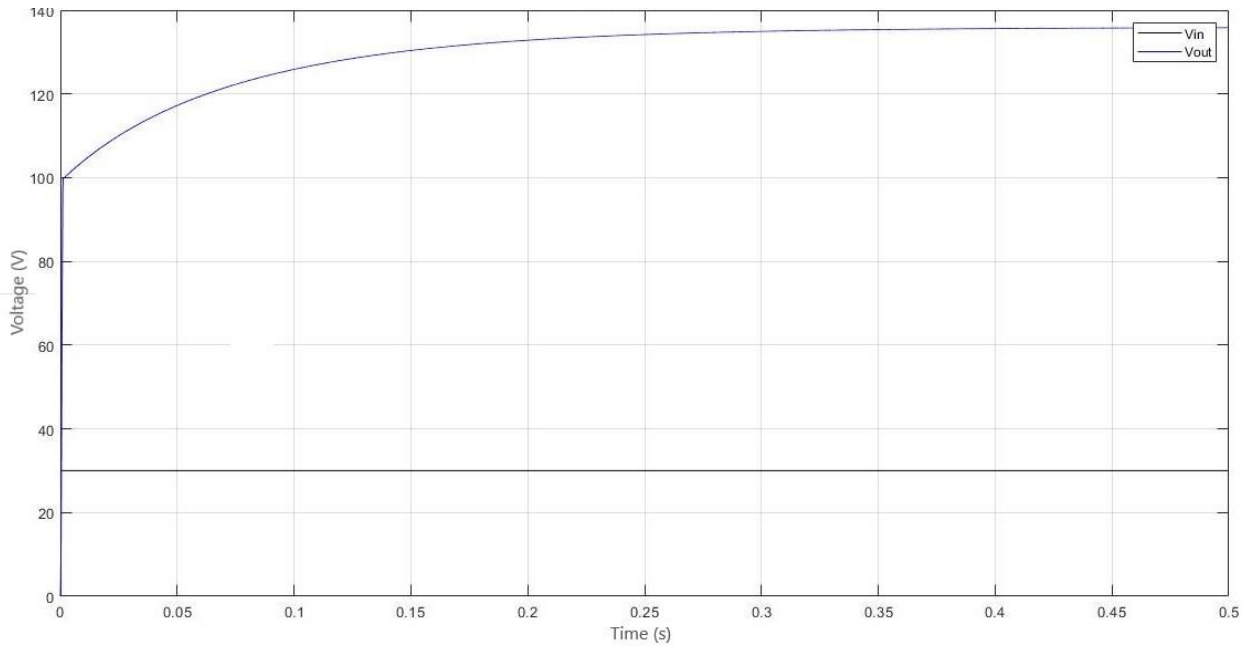
the equations that were analyzed in Section 2.5.1. The Boost converter diode parameters are presented in Table 4-4, while the Mosfet parameters are presented in Table 4-5.

<b>Table 4-3. Boost converter operating parameters.</b>	
Inductor ( $\mu\text{H}$ )	250
Capacitor ( $\mu\text{F}$ )	200
Load (Ohm)	1000
DC input voltage (V)	30

<b>Table 4-4. Boost converter diode parameters.</b>	
Resistance $R_{\text{ON}}$ (Ohm)	0.001
Forward voltage $V_{\text{F}}$ (V)	0.8
Snubber resistance $R_{\text{S}}$ (Ohms)	500
Snubber capacitance (nF)	250

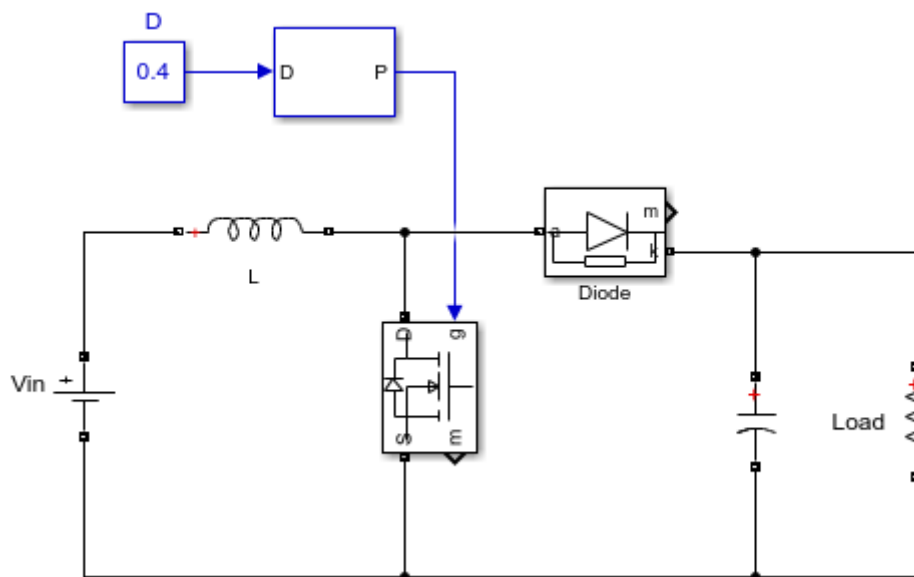
<b>Table 4-5. Boost converter Mosfet parameters.</b>	
FET resistance $R_{\text{ON}}$ (Ohm)	0.001
Internal diode forward resistance $R_{\text{d}}$ (Ohm)	0.001

An example of the Boost-type power converter input and output voltage is shown in Fig 4-3, demonstrating how much higher the output voltage is compared to the input voltage.



**Fig. 4-3.** An example of the Boost converter input and output voltage.

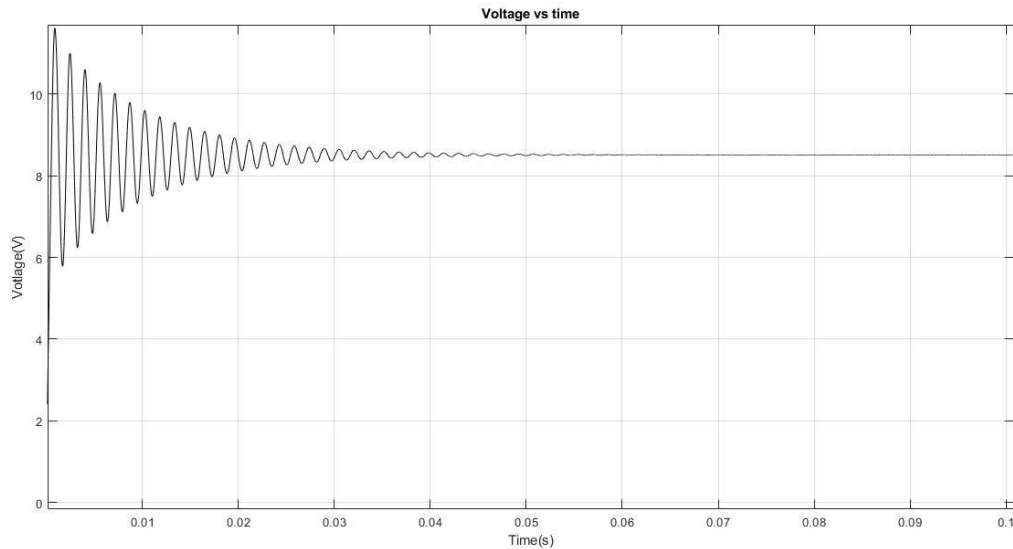
The circuit of the Boost-type power converter is presented in Fig. 4-4



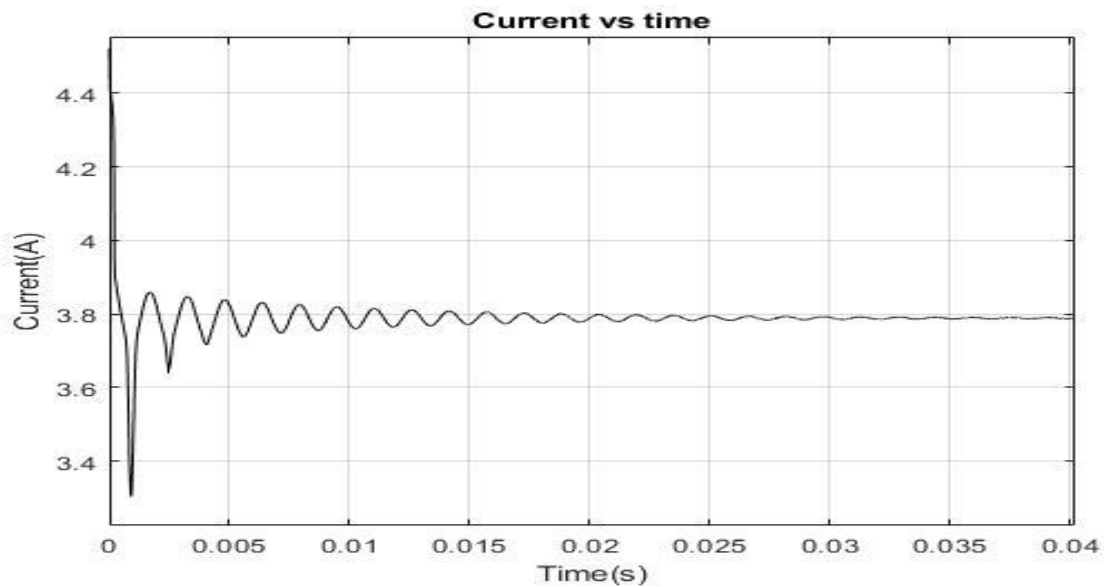
**Fig. 4-4.** The DC/DC Boost converter circuit diagram.

Nevertheless, the converter needed some changes, in order to connect it to the PV array output terminals. First of all, an input capacitor was necessary in order to filter the current generated by the PV arrays. The capacitor was defined to be equal to  $250\mu\text{F}$ . Secondly, the ohmic load was replaced by another DC voltage source, which represents the battery that was mentioned in the

beginning. This source was also connected in series with a resistance, which is equal to 1 Ohm. The switching frequency of the PWM signal, which controls the DC/DC Boost-type power converter was set equal to 100 kHz. At steady-state the input voltage and current ripple is <1% according to Figs. 4-5 and 4-6, respectively.



**Fig. 4-5.** An example of the DC/DC converter input voltage.



**Fig 4-6.** An example of the DC/DC converter input current.

The efficiency of the power converter is defined by (4.1).  $P_{in}$  and  $P_{out}$  were measured by changing the duty cycle in the range of (0.2-0.8) using a step-size equal to 0.1. The results are presented in Table 4-6.

$$\text{Power converter's efficiency} = \frac{P_{in}}{P_{out}} \quad (4.1)$$

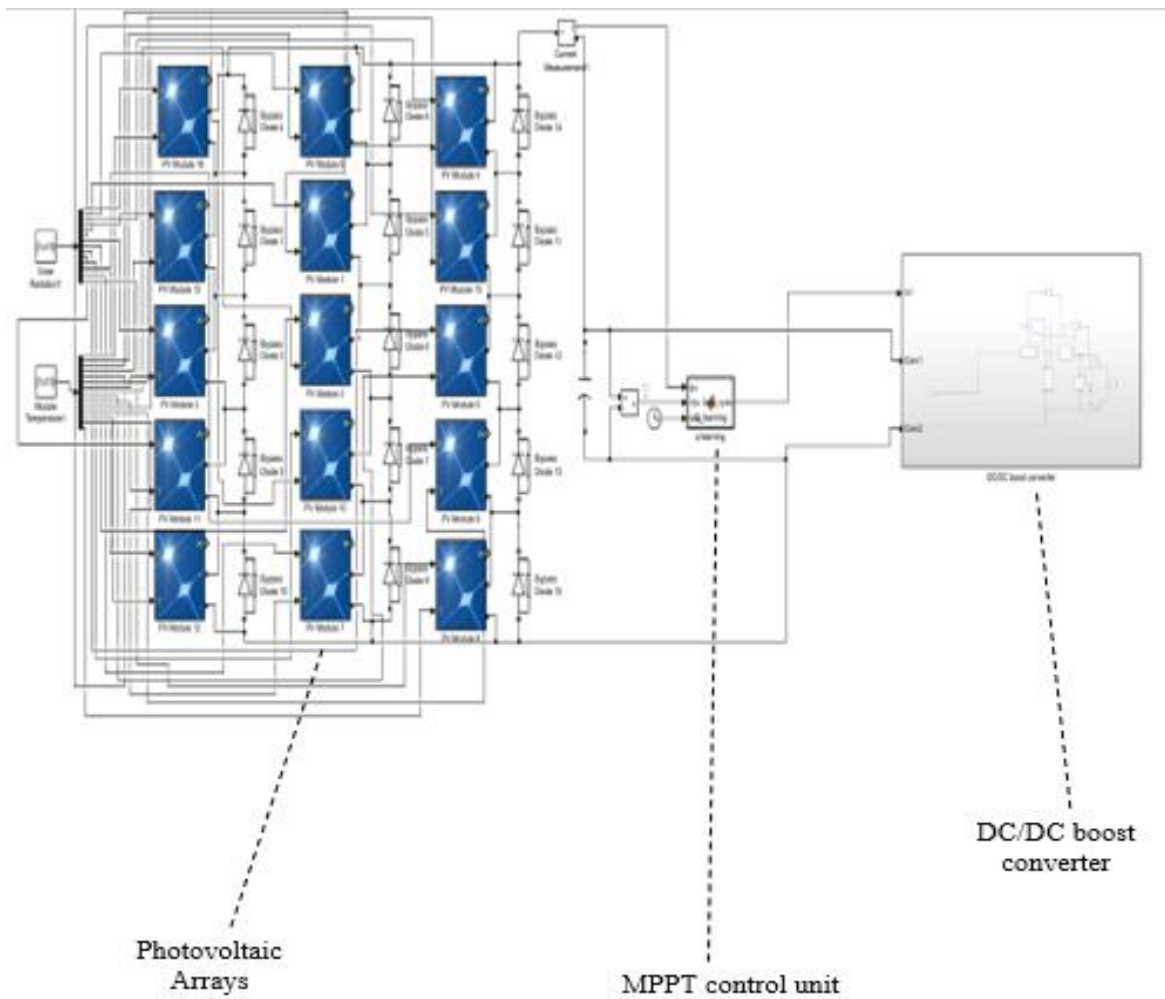
where:

$P_{in}$  is the power converter input power (W) and

$P_{out}$  is the power converter output power (W).

<b>Table 4-6. Boost converter efficiency.</b>		
$P_{in}$ (W)	$P_{out}$ (W)	Power converter's efficiency
43.28	41.8	96%
38.85	37.3	96%
65.19	63.2	96.9%
55.62	53.65	96.4%
54.155	52.3	96.5%
41.5	40.2	96.8%
32.25	31.2	96.7%





**Fig. 4-7.** Simulink design schematic diagram.

Then, the MPPT control unit was used to control the DC/DC converter. This unit was coded by using the MATLAB function block, which is located in the “Simulink/User-defined functions” library and corresponds to the microcontroller, which would be used in a real practical implementation. In Fig. 4-7, the overall design is presented.

# 5.

## Simulation results

---

In order to evaluate the Q-learning-based MPPT method, four different implementations were simulated. Two of them contained the value of the duty cycle in a previous time-step and the rest did not. In addition, a different number of equal quantized steps of the PV array output power constitute every case. More specifically:

### **First implementation of the Q-learning-based MPPT method:**

- The duty cycle (D) is quantized in 18 equal steps.
- The PV array output power (P) is quantized in 15 equal steps.
- The duty cycle of a previous time-step (DO) is quantized in 9 equal steps.

### **Second implementation of the Q-learning-based MPPT method:**

- The duty cycle (D) is quantized in 18 equal steps.
- The PV array output power (P) is quantized in 30 equal steps.
- The duty cycle of a previous time-step (DO) is quantized in 9 equal steps.

### **Third implementation of the Q-learning-based MPPT method:**

- The duty cycle (D) is quantized in 18 equal steps.
- The PV array output power (P) is quantized in 15 equal steps.
- The duty cycle of a previous time-step (DO) is not used for the determination of the Q-learning method state.

**Fourth implementation of the Q-learning-based MPPT method:**

- The duty cycle (D) is quantized in 18 equal steps.
- The PV array output power (P) is quantized in 30 equal steps.

The duty cycle of a previous time-step (DO) is not used for the determination of the Q-learning method state.

Each Q-learning-based MPPT method under study was simulated for 14 different shading patterns. These shading patterns were tested from shading pattern 1 to shading pattern 14 in sequential order. Thus, in each implementation of the Q-learning-based method, it is clear how the MPPT efficiency and the convergence time were affected by the order in which the shading patterns were applied. 9 of them were under non-uniform incident solar irradiation conditions and 5 of them were under uniform incident solar irradiation conditions. Some of them were used during the learning process and the rest were used during the MPPT execution process. Depending on the implemented Q-learning-based MPPT method, the agent needs to learn more or less shading patterns in to order to succeed an efficient MPPT execution process. These are described in detail next. Subsequently, a PSO-based MPPT method was also implemented and simulated in four representative shading patterns [24]. The PSO-based MPPT method was simulated only in four shading patterns, because it takes a specific amount of time in order to detect the GMPP in most cases. Thus, 1 shading pattern under uniform solar irradiation and 3 shading patterns under non-uniform solar irradiation were selected. Shading patterns 1-9 are related with non-uniform incident solar irradiation conditions and shading patterns 10-14 are related with uniform incident solar irradiation conditions. Sections 5.1-5.13 present the analysis of the simulation results of each Q-learning-based MPPT method for every shading pattern separately. The simulation results of the PSO-based MPPT method is presented separately in Section 5.14. In the end, the Q-learning-based MPPT methods performance was compared to that of the PSO-based MPPT method. It must be mentioned that the simulations stop when the agent converges to a value of the PV array output power. The MPPT efficiency is defined by (5.1).

$$n = \frac{P_{calculated}}{P_{theoretical}} \quad (5.1)$$

where:

$P_{calculated}$  is the calculated MPP using the MPPT methods (W) and

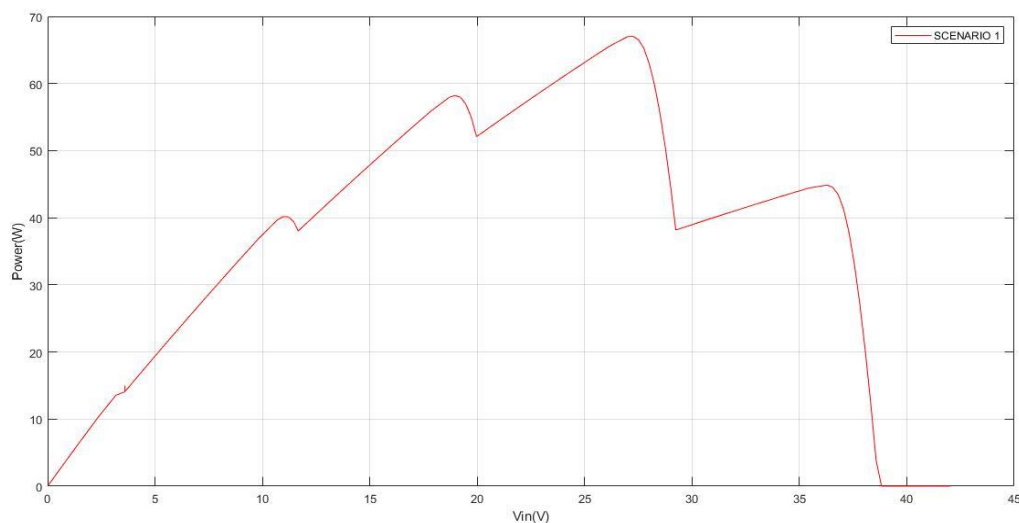
$P_{theoretical}$  is the theoretical MPP (W).

The parameter settings of the Q-learning-based methods are presented in Table 5.1.

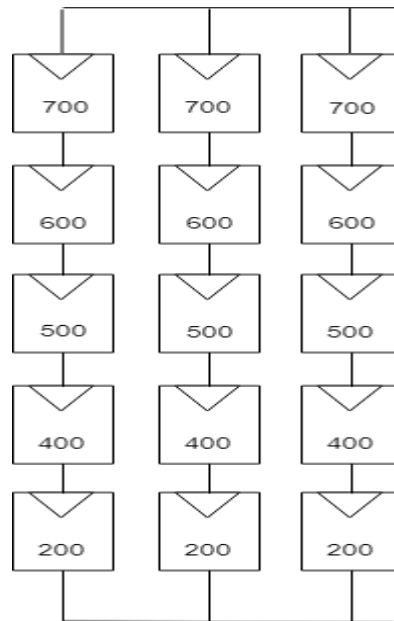
<b>Table 5-1. Parameter settings of the Q-learning-based methods.</b>			
Initial duty cycle	0.71	Minimum DC	0.19
$\tau_{\min}$	0.08	Maximum DC	0.83
$\tau_{\max}$	0.8	$k_1, k_2, k_3$	10, 25, 0.6
Discount factor	0.9	$N_{\max}$	20

## 5.1 Analysis of the simulation results for shading pattern 1

In Fig. 5-1 the PV array output power-voltage curve is presented and in Fig. 5-2 the distribution of incident solar irradiation over each PV module is presented. The numbers above each PV module represent the solar irradiation intensity, which is measured in  $W/m^2$ . Shading pattern 1 is a non-uniform incident solar irradiance pattern.

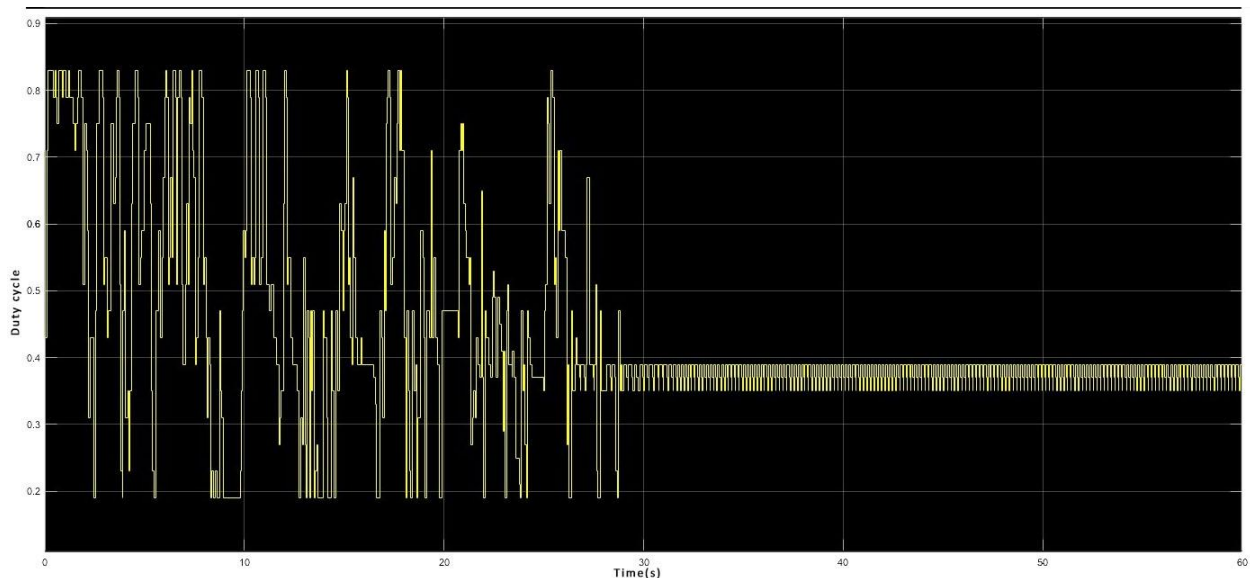


**Fig. 5-1.** The PV array output power-voltage curve for shading pattern 1.

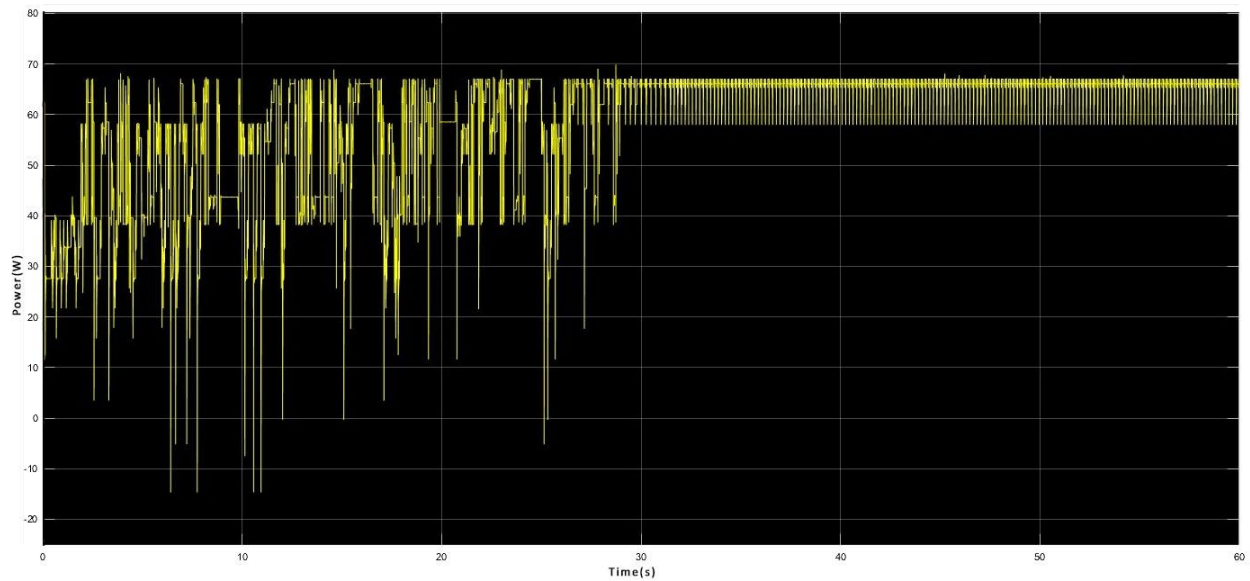


**Fig. 5-2.** The distribution of incident solar irradiation over each PV module for shading pattern 1.

Eight plots were exported, the PV array output power versus time plot and the duty cycle versus time plot in the four aforementioned implementations of the Q-learning-based MPPT method. Figs. 5-3 and 5-4 present the results of the simulations of the duty cycle and the PV array output power versus time for the first implementation of the Q-learning-based MPPT method for shading pattern 1.



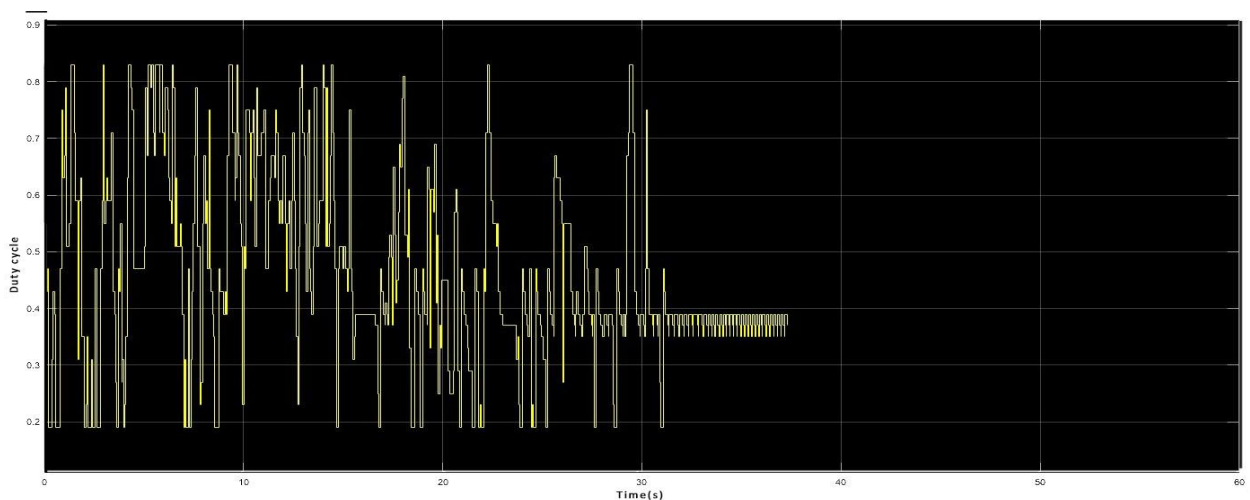
**Fig. 5-3.** Duty cycle versus time for shading pattern 1 of the first implementation of the Q-learning-based MPPT method.



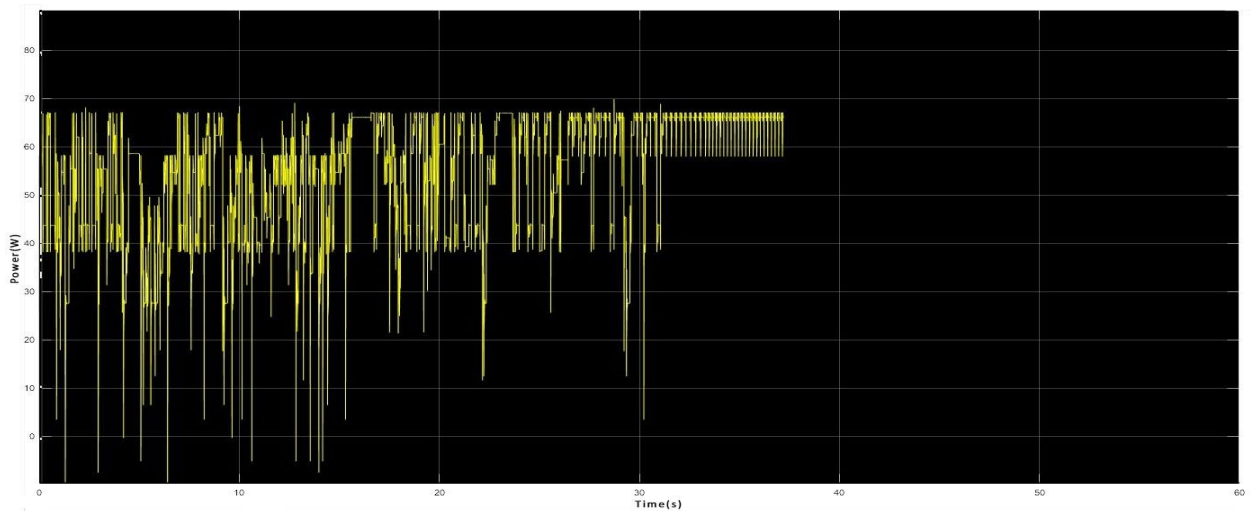
**Fig. 5-4.** PV array output power versus time for shading pattern 1 of the first implementation of the Q-learning-based MPPT method.

It is observed that the algorithm needs a relatively long amount of time in order to converge. This is due to the fact that the agent has no prior knowledge of where the GMPP may be located. Thus, it has to visit many states until it converges to a specific duty cycle value. It is considered that the exploration is over, when the oscillations start. These oscillations are due to the P&O method, which is used at that time.

Figs. 5-5 and 5-6 present the results of the simulations of duty cycle and PV array output power versus time of the second implementation of the Q-learning-based MPPT method for shading pattern 1.



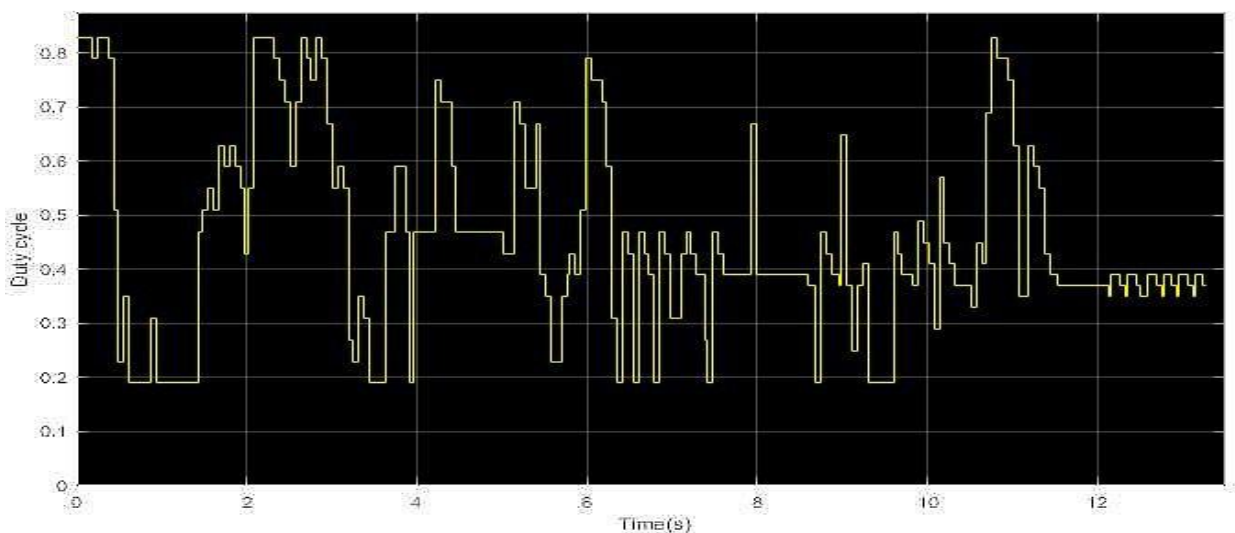
**Fig. 5-5.** Duty cycle versus time for shading pattern 1 of the second implementation of the Q-learning-based MPPT method.



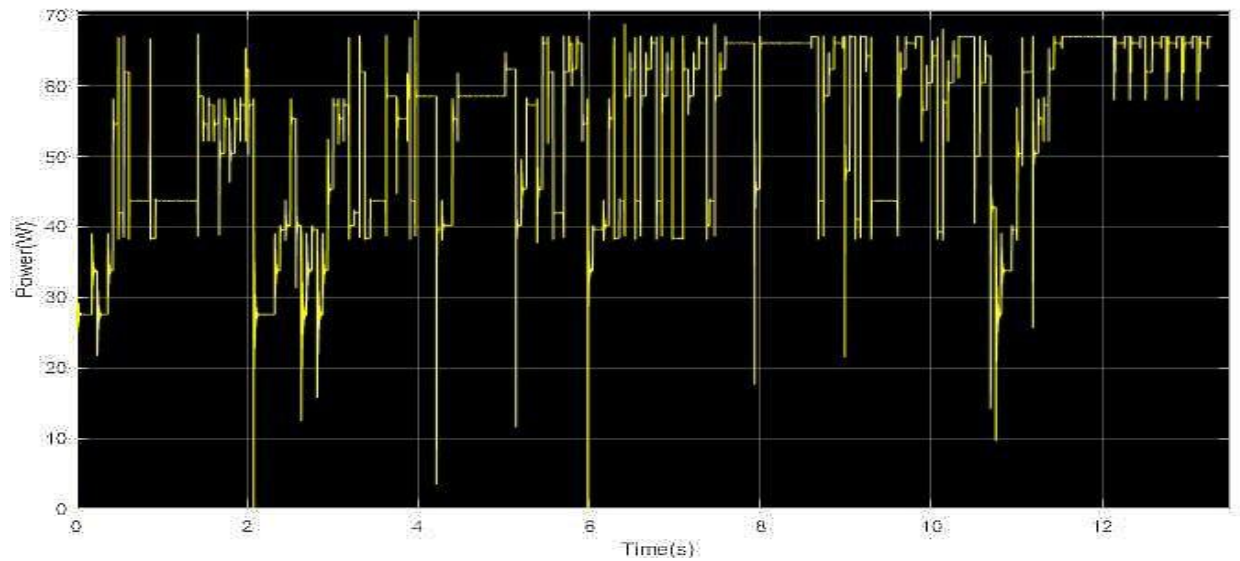
**Fig. 5-6.** PV array output power versus time for shading pattern 1 of the second implementation of the Q-learning-based MPPT method.

It is observed that the agent needs approximately the same amount of time that it took in the first implementation of the Q-learning-based MPPT method, in order to locate the MPP. This happens due to the fact that this is the first shading pattern that the agent is trying to learn where the MPP is located. The oscillations are due to the P&O method, which is used at that time.

Figs. 5-7 and 5-8 present the results of the simulations of duty cycle and PV array output power versus time of the third implementation of the Q-learning-based MPPT method for shading pattern 1.



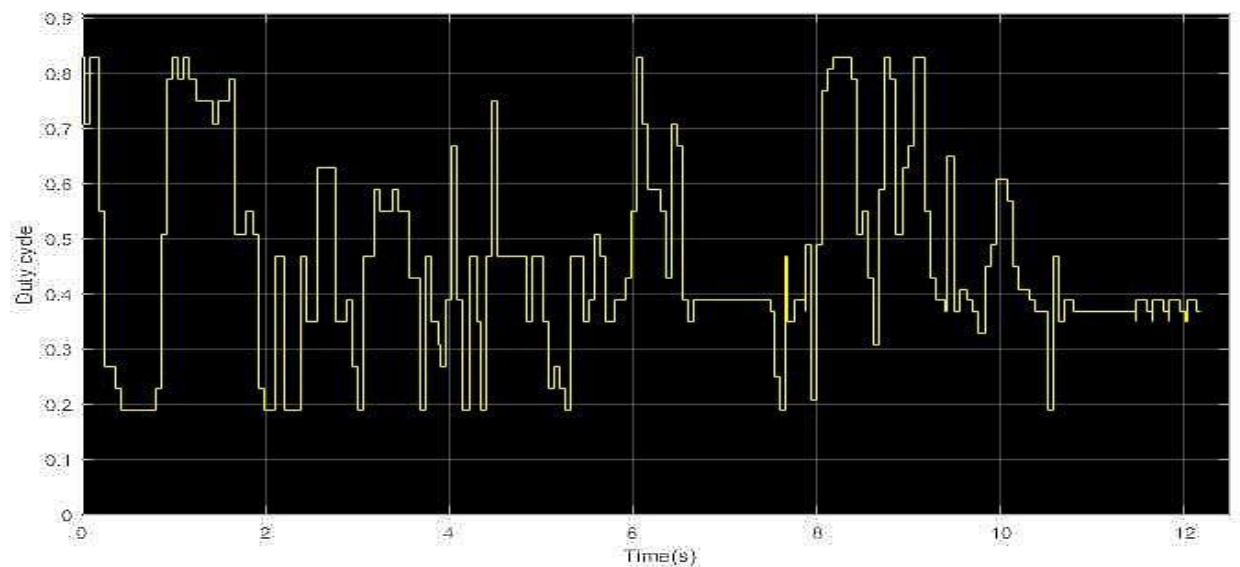
**Fig. 5-7.** Duty cycle versus time for shading pattern 1 of the third implementation of the Q-learning-based MPPT method.



**Fig. 5-8.** PV array output versus time for shading pattern 1 of the third implementation of the Q-learning-based MPPT method.

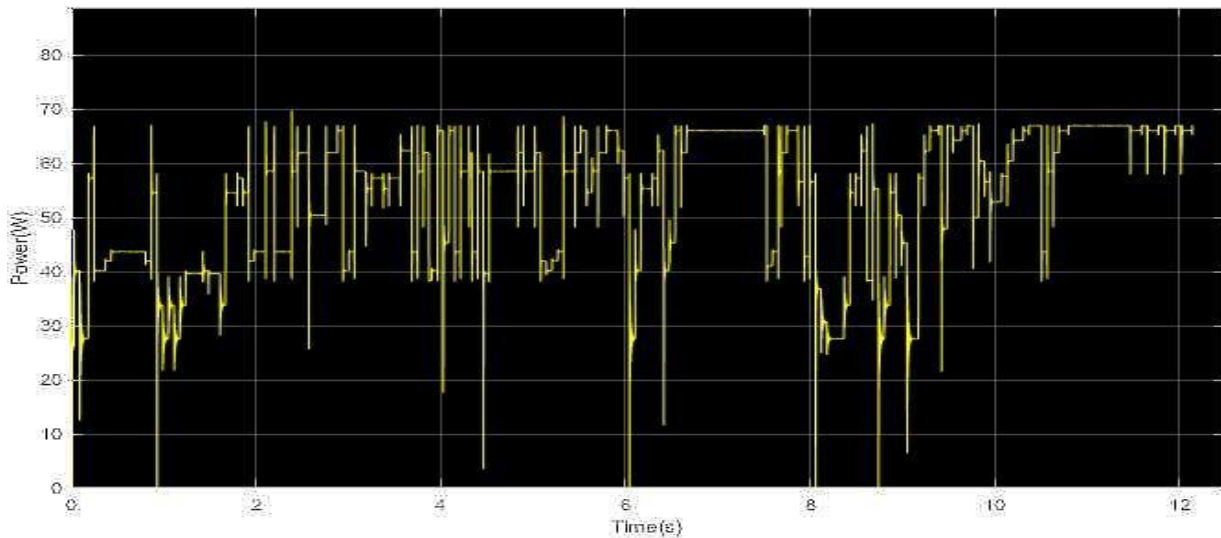
The shorter convergence time is due to the fact that in the third implementation of the Q-learning-based MPPT method, the agent does not need to know the duty cycle value in the previous time-step. This results in a significantly smaller number of time steps until the agent converges, compared with the first and second implementation of the Q-learning-based MPPT method.

Figs. 5-9 and 5-10 present the results of the simulations of duty cycle and PV array output power versus time of the fourth implementation of the Q-learning-based MPPT method for shading pattern 1.



**Fig. 5-9.** Duty cycle versus time for shading pattern 1 of the fourth implementation of the Q-learning-based MPPT method.



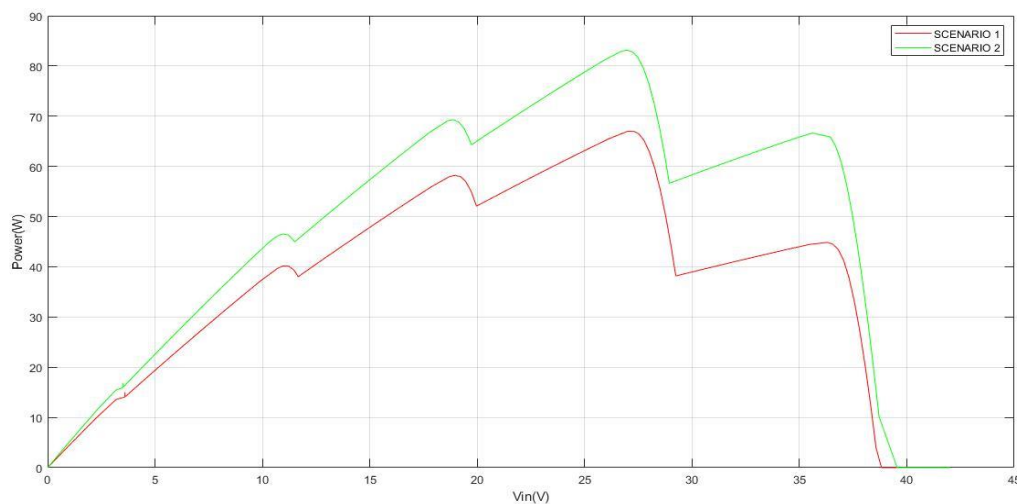


**Fig. 5-10.** PV array output power versus time for shading pattern 1 of the fourth implementation of the Q-learning-based MPPT method.

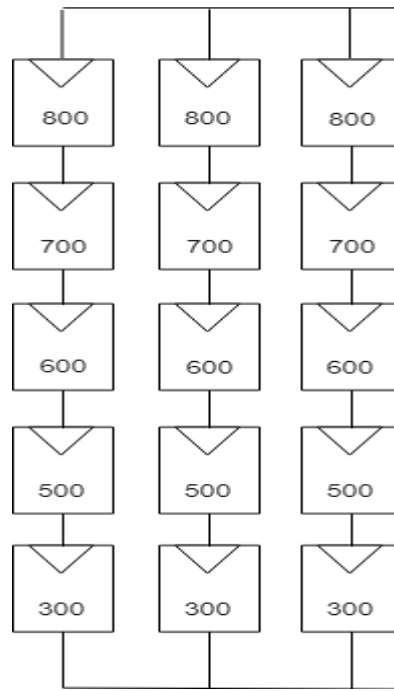
It is observed that the agent needs approximately the same amount of time that it took in the third implementation of the Q-learning-based MPPT method, in order to locate the GMPP. This happens due to the fact that this is the first shading pattern that the agent is trying to learn.

## 5.2 Analysis of the simulation results for shading pattern 2

In Fig. 5-11 the PV array output power-voltage curve is presented and in Fig. 5-12 the distribution of incident solar irradiation over each PV module is presented. The numbers above each PV module represent the incident solar irradiation intensity, which is measured in  $W/m^2$ . Shading pattern 2 is a non-uniform solar irradiance pattern.

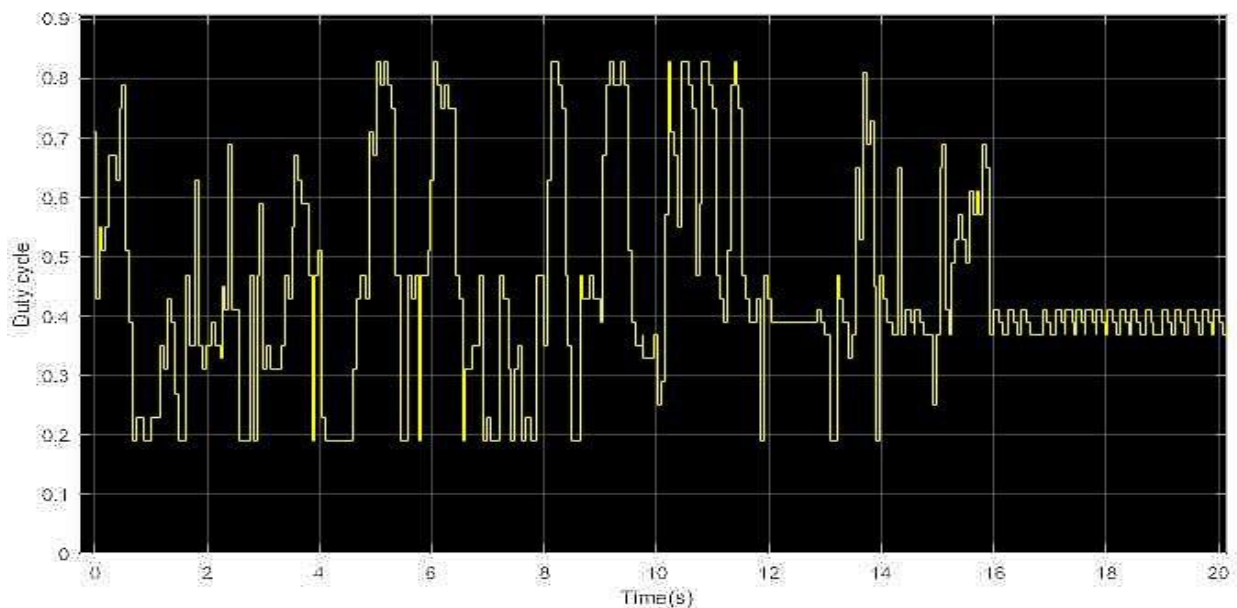


**Fig. 5-11.** The PV array output power-voltage curve for shading pattern 2.

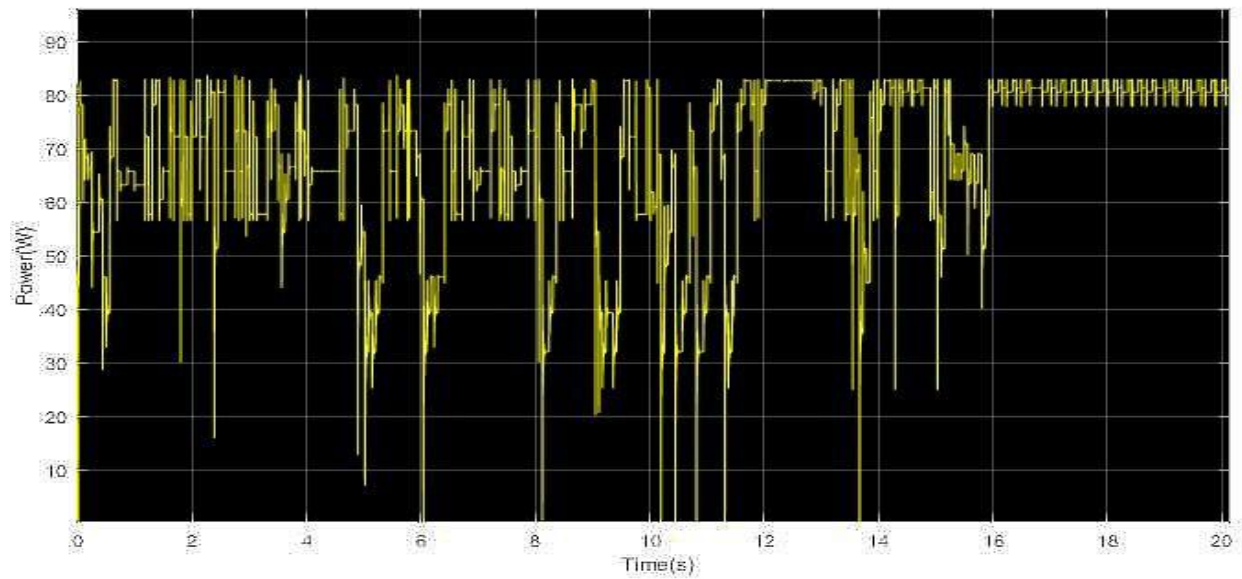


**Fig. 5-12.** The distribution of incident solar irradiation over each PV module for shading pattern 2.

Eight plots were exported, the PV array output power versus time plot and the duty cycle versus time plot in the four aforementioned implementations of the Q-learning-based MPPT method. Figs. 5-13 and 5-14 present the results of the simulations of the duty cycle and the PV array output power versus time for the first implementation of the Q-learning-based MPPT method for shading pattern 2.



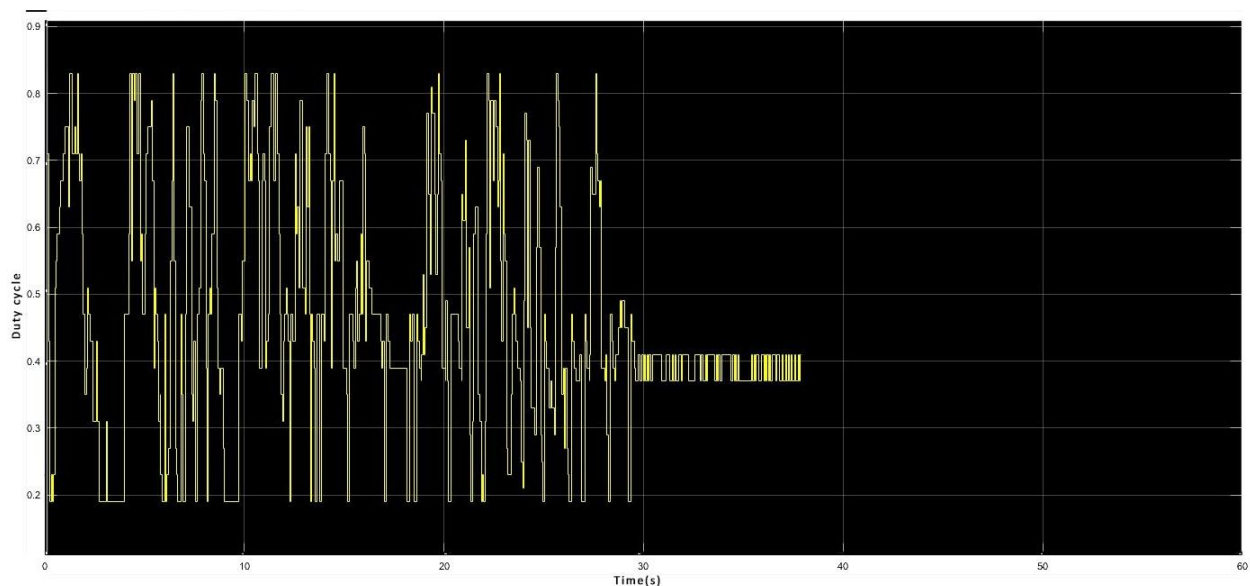
**Fig. 5-13.** Duty cycle versus time for shading pattern 2 of the first implementation of the Q-learning-based MPPT method.



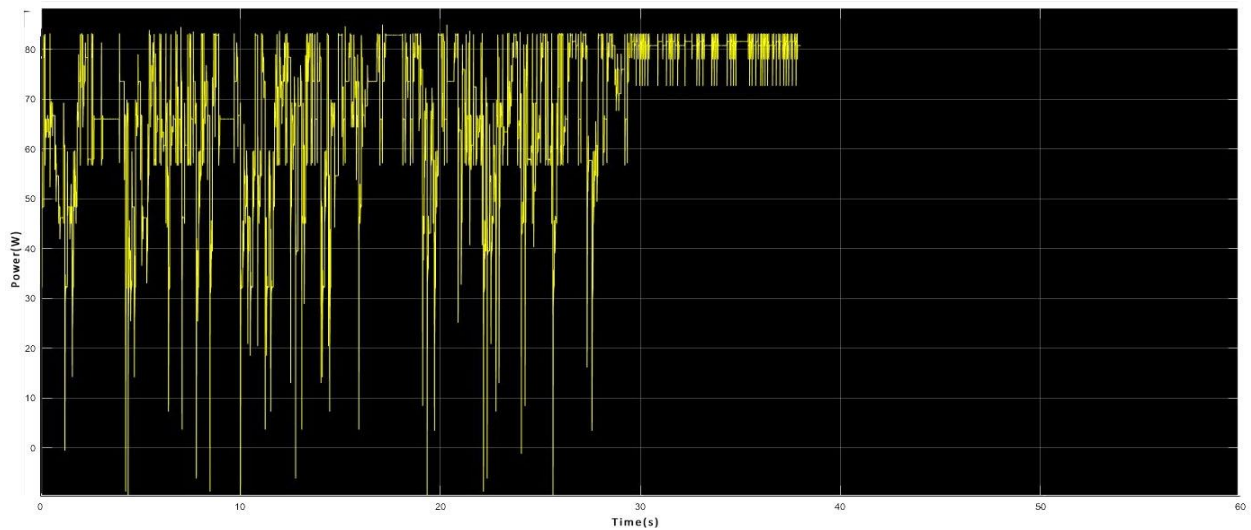
**Fig. 5-14.** PV array output power versus time for shading pattern 2 of the first implementation of the Q-learning-based MPPT method.

It is observed that the knowledge gained by the agent during the MPPT process in the first implementation of the Q-learning-based MPPT method for shading pattern 1, helped the agent to locate the GMPP in a smaller amount of time.

Figs. 5-15 and 5-16 present the results of the simulations of duty cycle and PV array output power versus time of the second implementation of the Q-learning-based MPPT method for shading pattern 2.



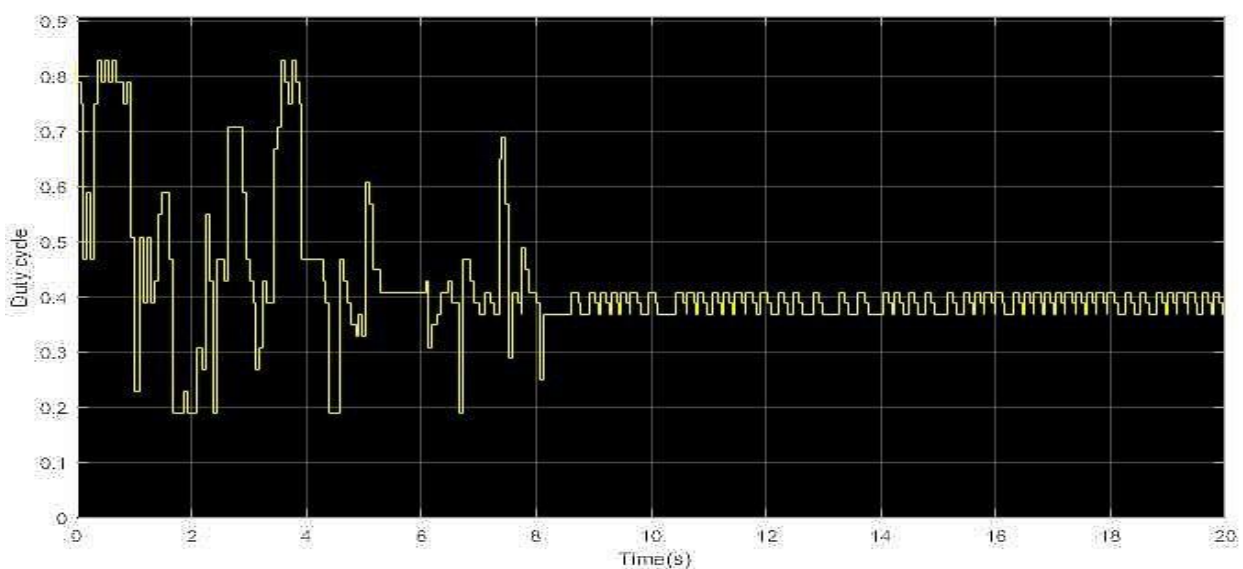
**Fig. 5-15.** Duty cycle versus time for shading pattern 2 of the second implementation of the Q-learning-based MPPT method.



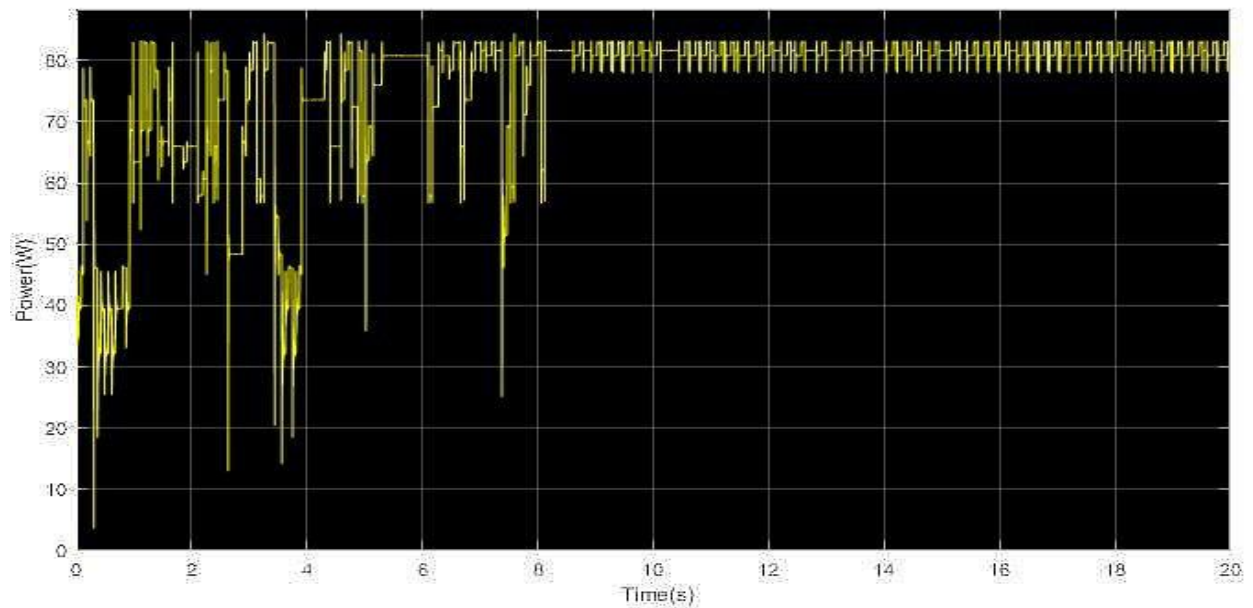
**Fig. 5-16.** PV array output power versus time for shading pattern 2 of the second implementation of the Q-learning-based MPPT method.

It is observed that the knowledge gained by the agent during the MPPT process in the second implementation of the Q-learning-based MPPT method for shading pattern 1 did not affect the MPPT process for locating the MPP for shading pattern 2. This is due to the fact that in the second implementation of the Q-learning based MPPT method the PV array output power is quantized in 30 equal steps, while in the first implementation the PV array output power is quantized in 15 equal steps.

Figs. 5-17 and 5-18 present the results of the simulations of duty cycle and PV array output power versus time of the third implementation of the Q-learning-based MPPT method for shading pattern 2.



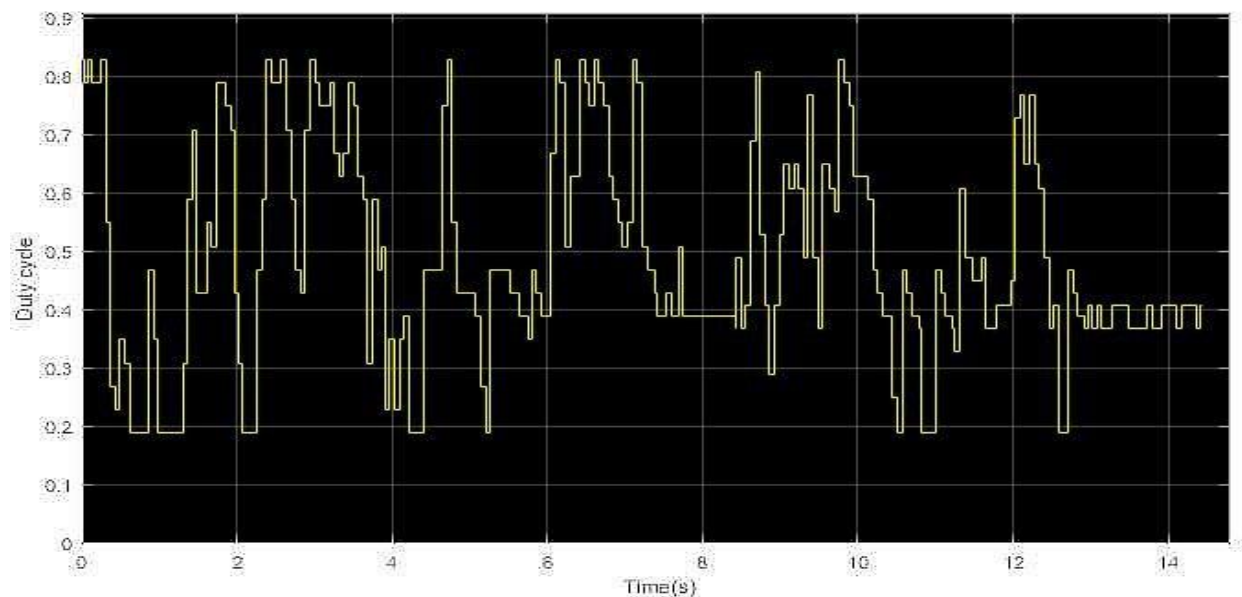
**Fig. 5-17.** Duty cycle versus time for shading pattern 2 of the third implementation of the Q-learning-based MPPT method.



**Fig. 5-18.** PV array output power versus time for shading pattern 2 of the third implementation of the Q-learning-based MPPT method.

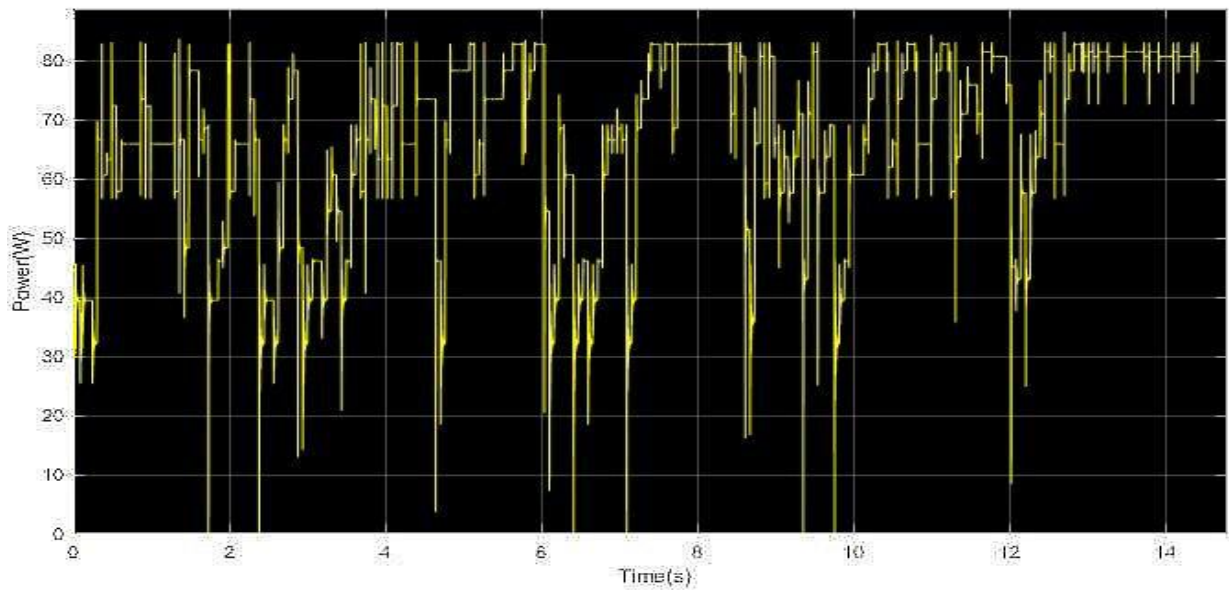
It is observed that the knowledge gained by the agent during the MPPT process in the third implementation of the Q-learning method for shading pattern 1, helped the agent to locate the GMPP in a smaller amount of time.

Figs. 5-19 and 5-20 present the results of the simulations of duty cycle and PV array output power versus time of the fourth implementation of the Q-learning-based MPPT method for shading pattern 2.



**Fig. 5-19.** Duty cycle versus time for shading pattern 2 of the fourth implementation of the Q-learning-based MPPT method.



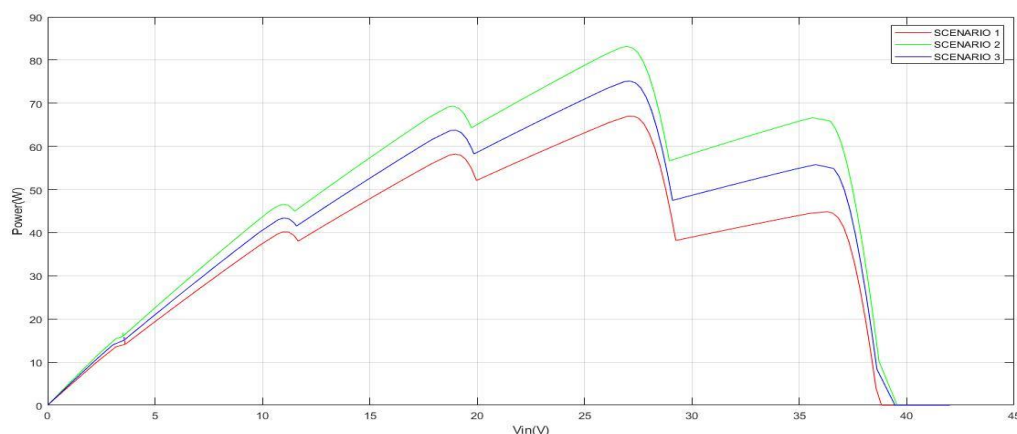


**Fig. 5-20.** PV array output power versus time for shading pattern 2 of the fourth implementation of the Q-learning-based MPPT method.

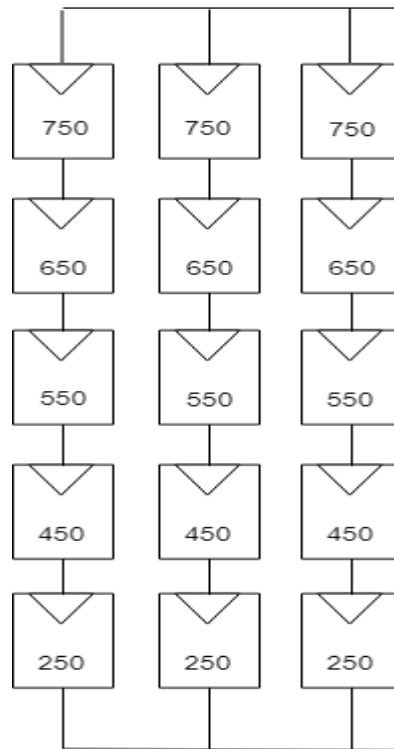
It is observed that the knowledge gained by the agent during the MPPT process in the fourth implementation of the Q-learning method did not affect the MPPT process for locating the MPP for shading pattern 2. This is due to the fact that in the fourth implementation of the Q-learning based MPPT method the PV array output power is quantized in 30 equal steps, while in the third implementation the PV array output power is quantized in 15 equal steps.

### 5.3 Analysis of the simulation results for shading pattern 3

In Fig. 5-21 the PV array output power-voltage curve is presented and in Fig. 5-22 the distribution of incident solar irradiation over each PV module is presented. The numbers above each PV module represent the solar irradiation intensity, which is measured in  $W/m^2$ . Shading pattern 3 is a non-uniform solar irradiance pattern.

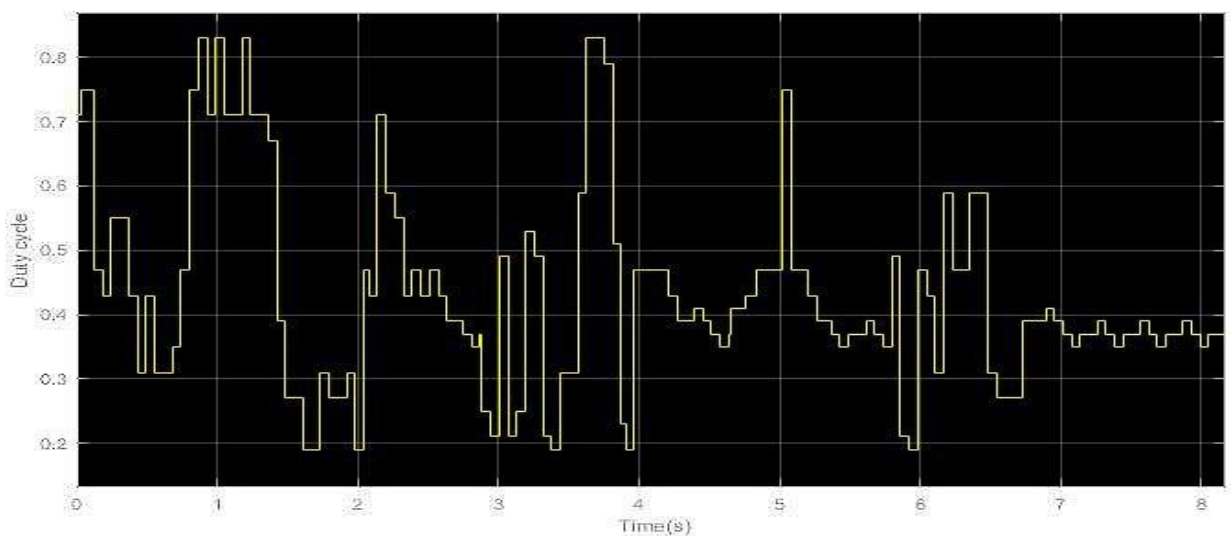


**Fig. 5-21.** The PV array output power-voltage curve for shading pattern 3.

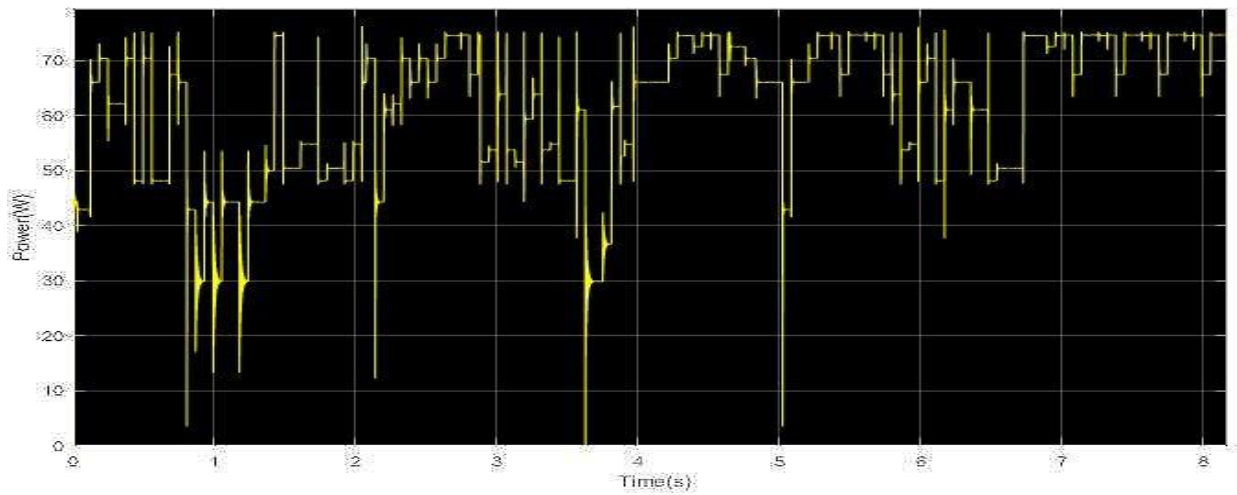


**Fig. 5-22.** The distribution of incident solar irradiation over each PV module for shading pattern 3.

Eight plots were exported, the PV array output power versus time plot and the duty cycle versus time plot in the four aforementioned implementations of the Q-learning-based MPPT method. Figs. 5-23 and 5-24 present the results of the simulations of the duty cycle and the PV array output power versus time for the first implementation of the Q-learning-based MPPT method for shading pattern 3.



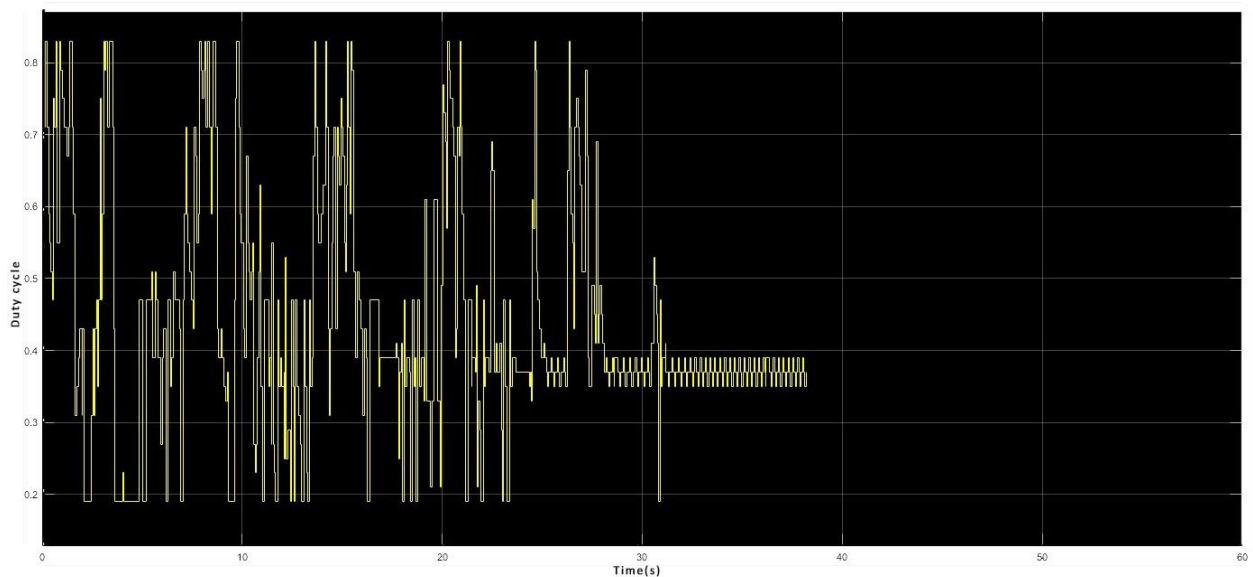
**Fig. 5-23.** Duty cycle versus time for shading pattern 3 of the first implementation of the Q-learning-based MPPT method.



**Fig. 5-24.** PV array output power versus time for shading pattern 3 of the first implementation of the Q-learning-based MPPT method.

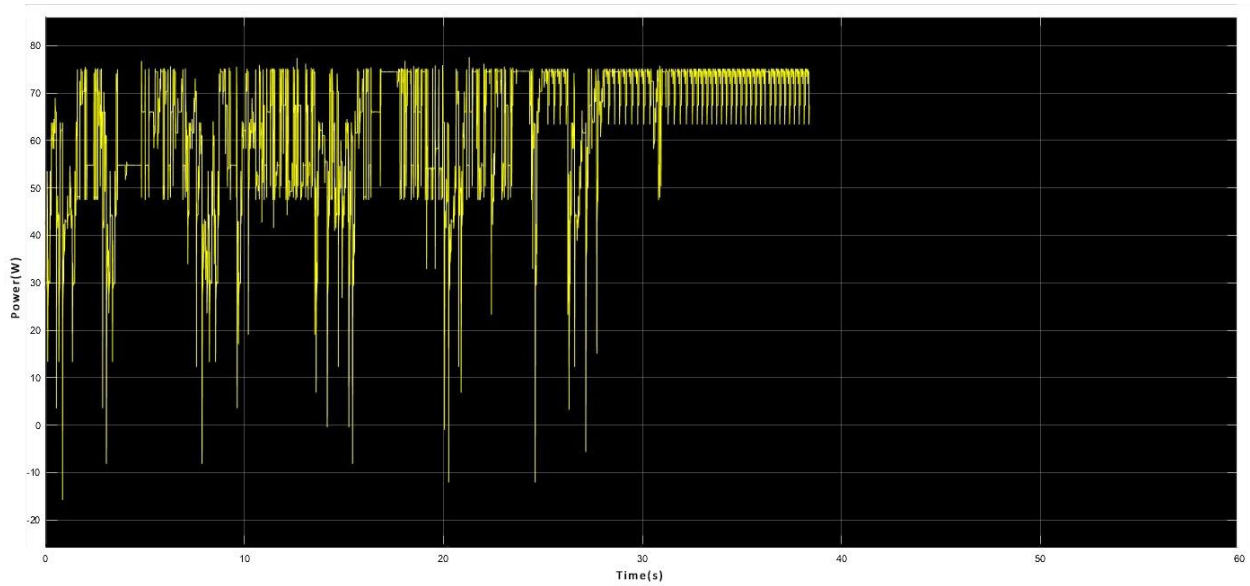
It is observed that the knowledge gained by the agent during the MPPT process in the first implementation of the Q-learning method for shading patterns 1 and 2, helped the agent to reduce even further the required time to locate the MPP.

Figs. 5-25 and 5-26 present the results of the simulations of duty cycle and PV array output power versus time of the second implementation of the Q-learning-based MPPT method for shading pattern 3.



**Fig. 5-25.** Duty cycle versus time for shading pattern 3 of the second implementation of the Q-learning-based MPPT method.

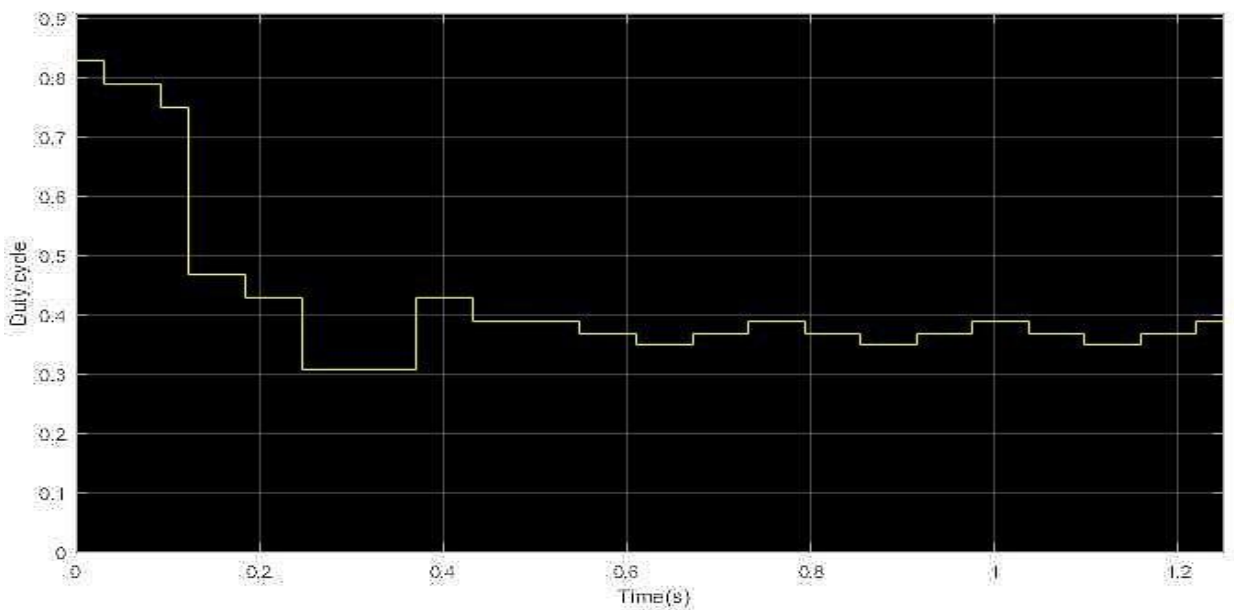




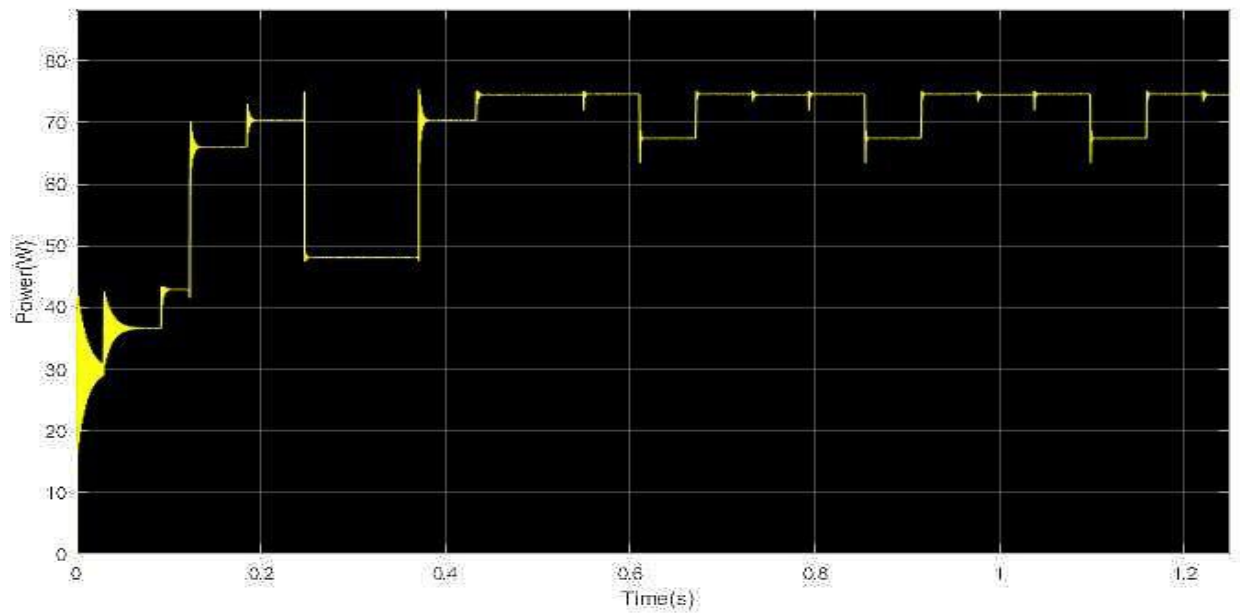
**Fig. 5-26.** PV array output power versus time for shading pattern 3 of the second implementation of the Q-learning-based MPPT method.

It is observed that the knowledge gained by the agent during the MPPT process in the second implementation of the Q-learning method for shading pattern 1 and 2, did not affect the MPPT process for locating the MPP for shading pattern 3.

Figs. 5-27 and 5-28 present the results of the simulations of duty cycle and PV array output power versus time of the third implementation of the Q-learning-based MPPT method for shading pattern 3.



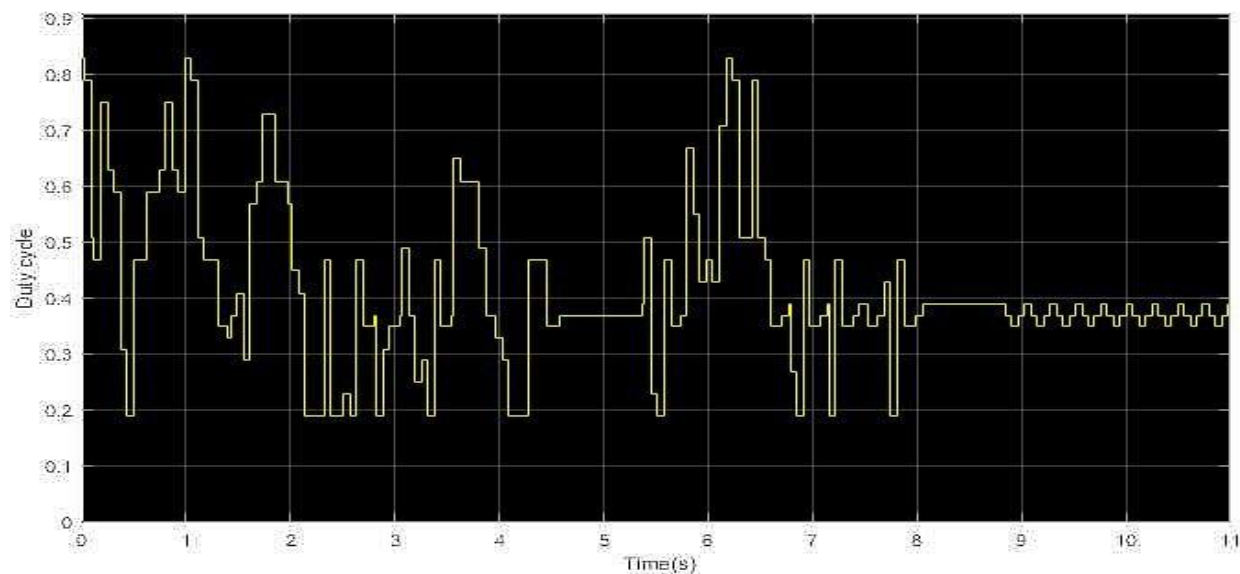
**Fig. 5-27.** Duty cycle versus time for shading pattern 3 of the third implementation of the Q-learning-based MPPT method.



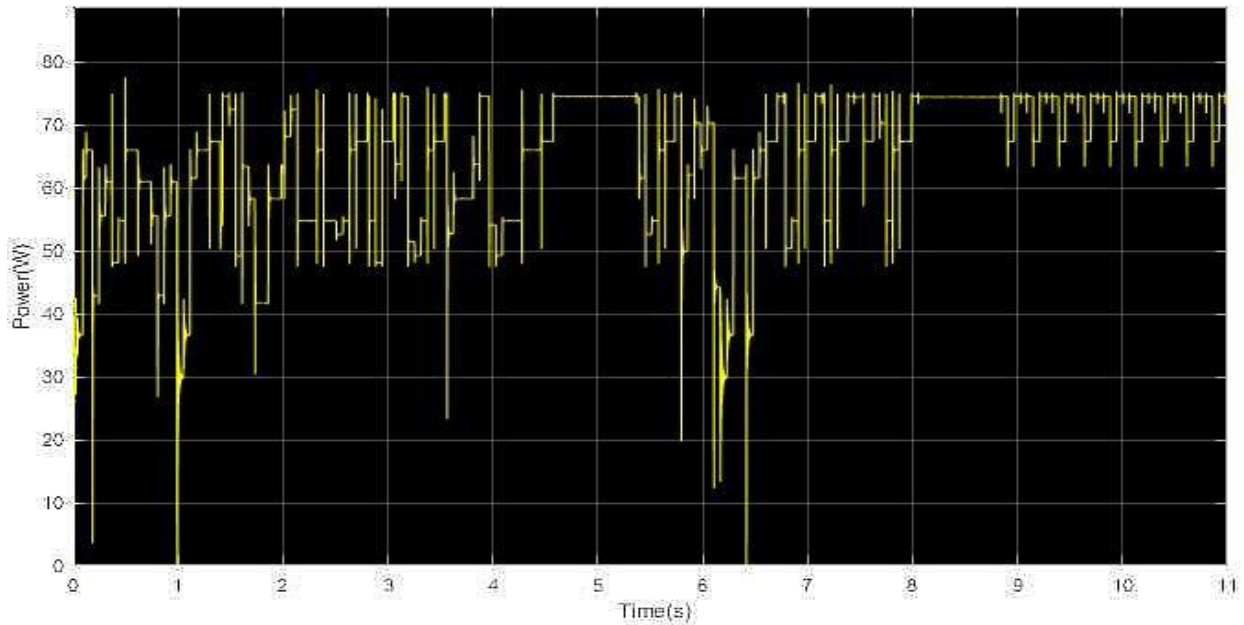
**Fig. 5-28.** PV array output power versus time for shading pattern 3 of the third implementation of the Q-learning-based MPPT method.

It is observed that the knowledge gained by the agent during the MPPT process in the third implementation of the Q-learning method for shading pattern 1 and 2, helped the agent to locate the MPP in only 6 time-steps.

Figs. 5-29 and 5-30 present the results of the simulations of duty cycle and PV array output power versus time of the fourth implementation of the Q-learning-based MPPT method for shading pattern 3.



**Fig. 5-29.** Duty cycle versus time for shading pattern 3 of the fourth implementation of the Q-learning-based MPPT method.

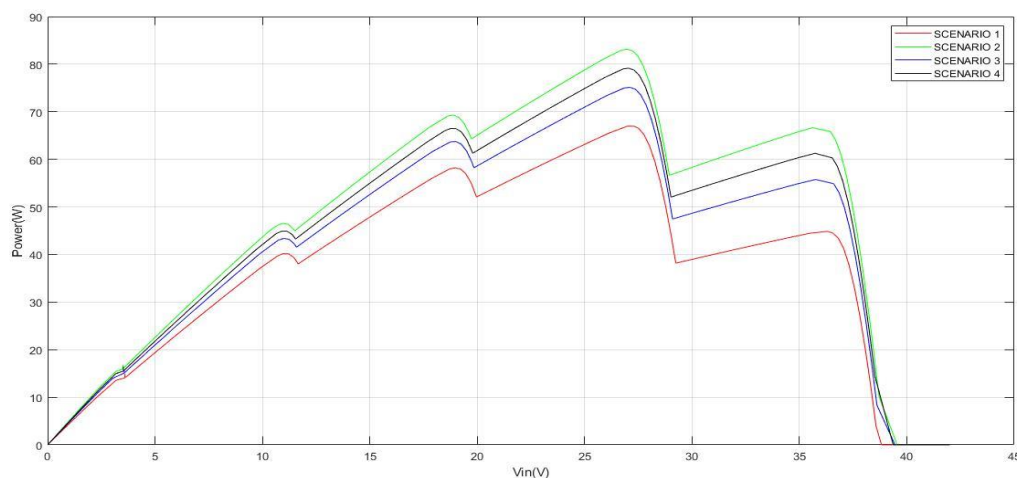


**Fig. 5-30** PV array output power versus time for shading pattern 3 of the fourth implementation of the Q-learning-based MPPT method.

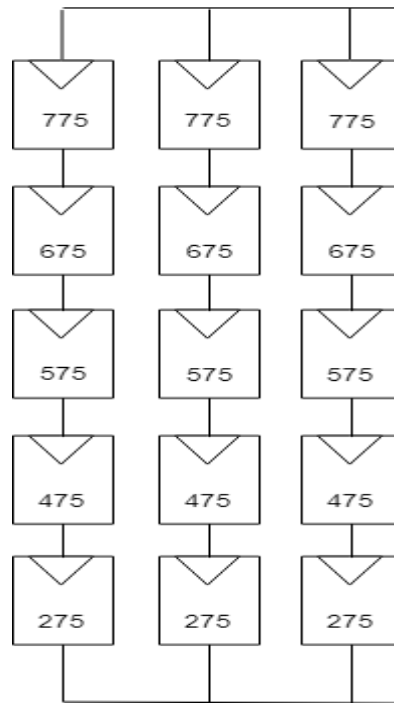
It is observed that the knowledge gained by the agent during the MPPT process in the fourth implementation of the Q-learning method for shading pattern 1 and 2, helped the agent to reduce the required time-steps until it converges to the GMPP.

#### 5.4 Analysis of the simulation results for shading pattern 4

In Fig. 5-31 the PV array output power-voltage curve is presented and in Fig. 5-32 the distribution of incident solar irradiation over each PV module is presented. The numbers above each PV module represent the solar irradiation intensity, which is measured in  $W/m^2$ . Shading pattern 4 is a non-uniform solar irradiance pattern.

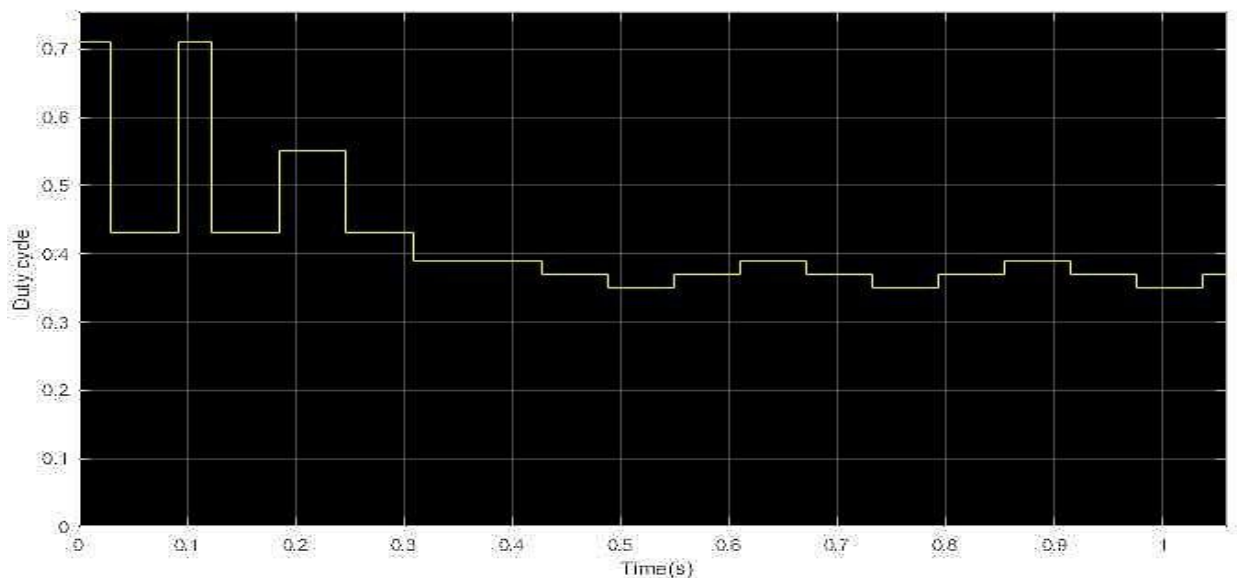


**Fig. 5-31.** The PV array output power-voltage curve for shading pattern 4.

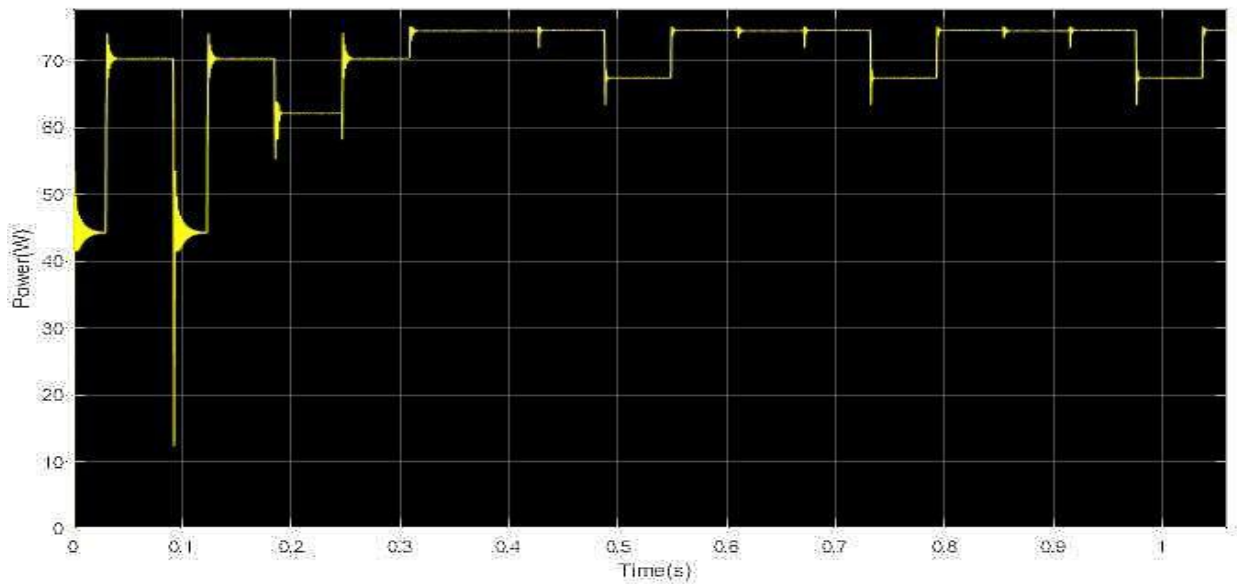


**Fig. 5-32.** The distribution of incident solar irradiation over each PV module for shading pattern 4.

Eight plots were exported, the PV array output power versus time plot and the duty cycle versus time plot in the four aforementioned implementations of the Q-learning-based MPPT method. Figs. 5-33 and 5-34 present the results of the simulations of the duty cycle and the PV array output power versus time for the first implementation of the Q-learning-based MPPT method for shading pattern 4.



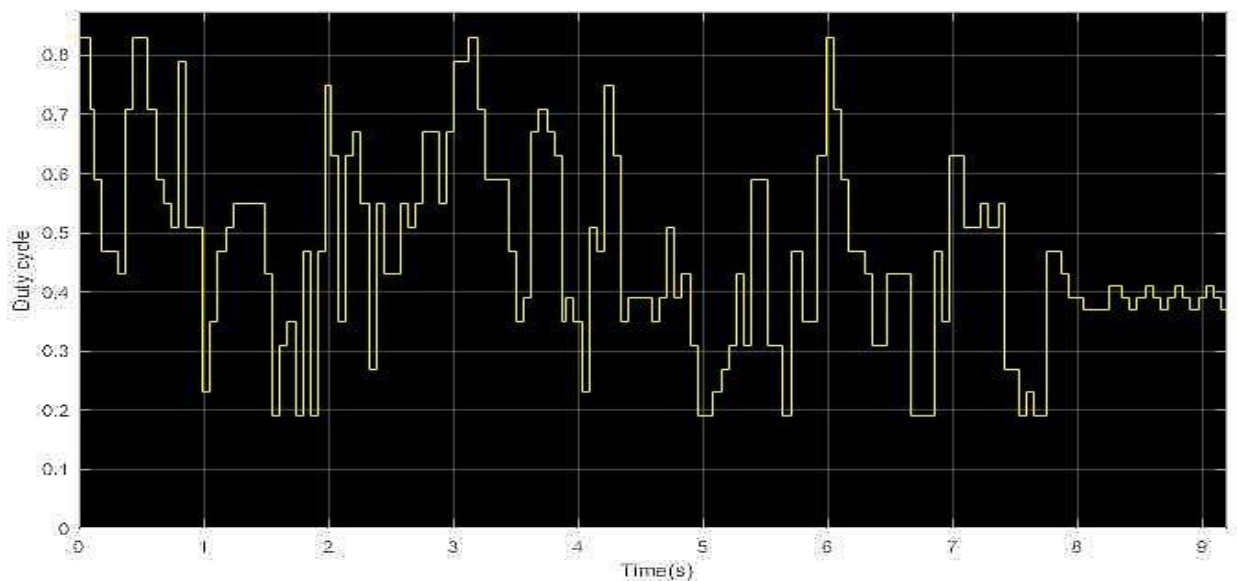
**Fig. 5-33.** Duty cycle versus time for shading pattern 4 of the first implementation of the Q-learning-based MPPT method.



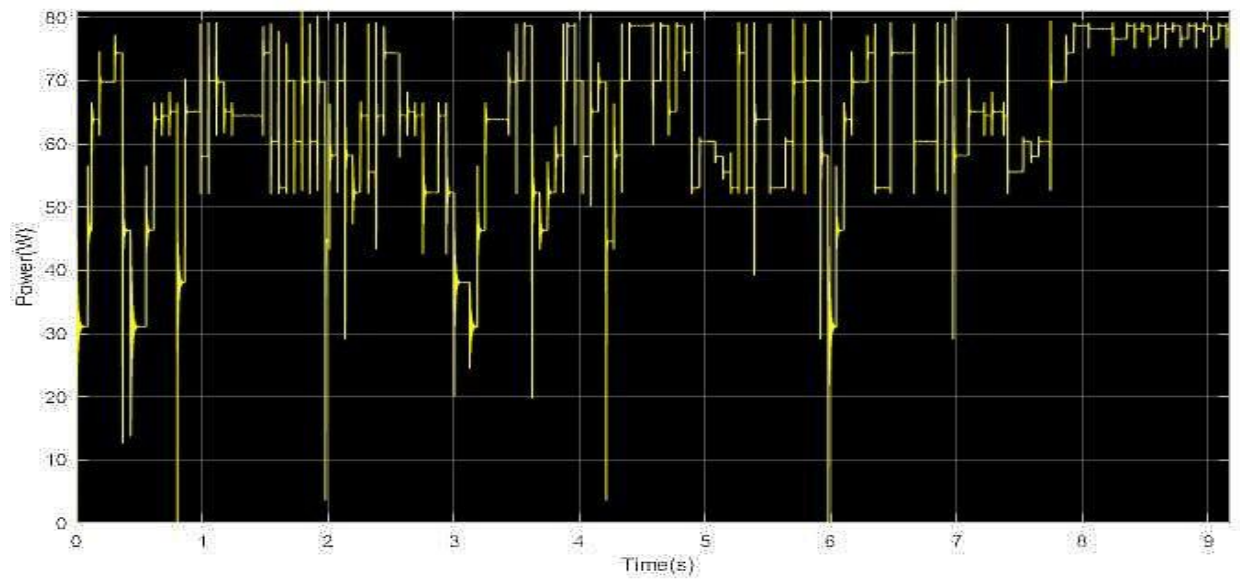
**Fig. 5-34.** PV array output power versus time for shading pattern 4 of the first implementation of the Q-learning-based MPPT method.

It is observed that in the first implementation of the Q-learning-based method the agent is able to detect the GMPP in unknown intermediate conditions of solar irradiation of shading patterns 1 and 2 in a few time-steps.

Figs. 5-35 and 5-36 present the results of the simulations of duty cycle and PV array output power versus time of the second implementation of the Q-learning-based MPPT method for shading pattern 4.



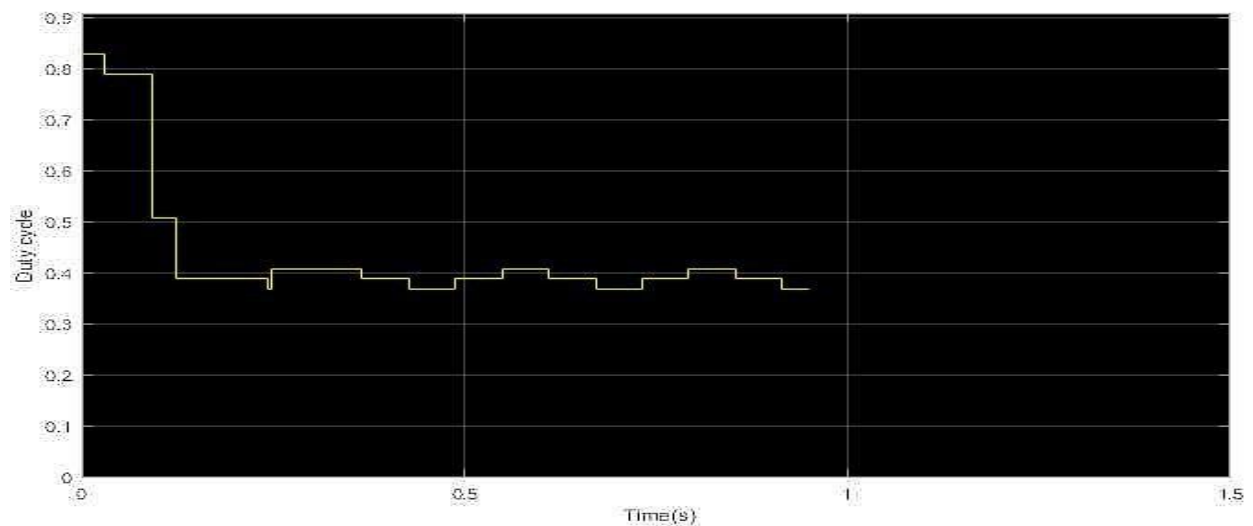
**Fig. 5-35.** Duty cycle versus time for shading pattern 4 of the second implementation of the Q-learning-based MPPT method.



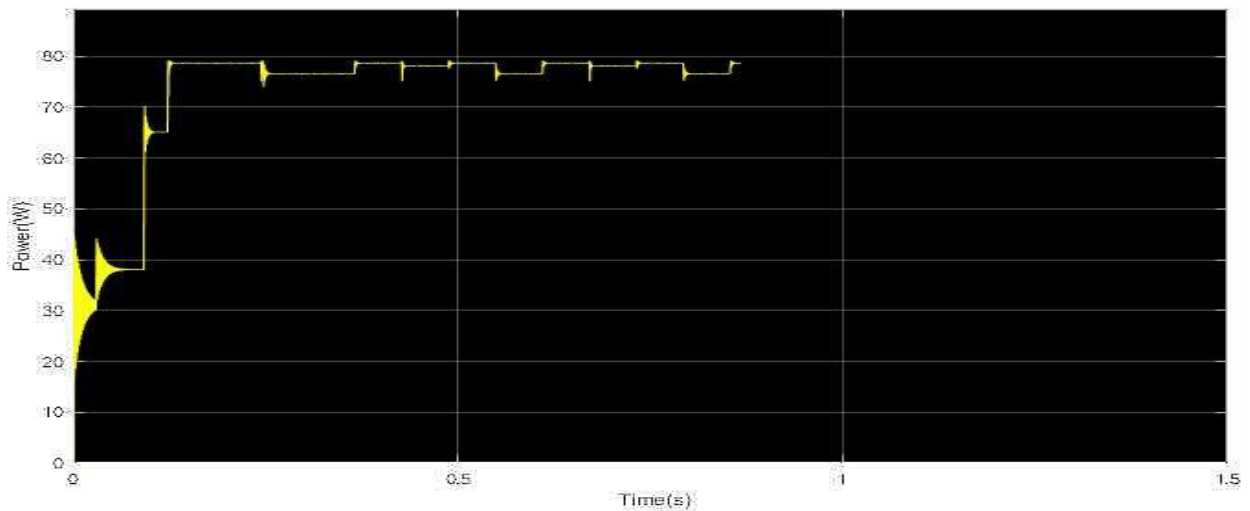
**Fig. 5-36.** PV array output power versus time for shading pattern 4 of the second implementation of the Q-learning-based MPPT method.

It is observed that after the MPPT process for shading patterns 1, 2 and 3, the agent is able to detect the GMPP in a smaller amount of time in intermediate conditions of solar irradiation, similar to the shading pattern 4.

Figs. 5-37 and 5-38 present the results of the simulations of duty cycle and PV array output power versus time of the third implementation of the Q-learning-based MPPT method for shading pattern 4.



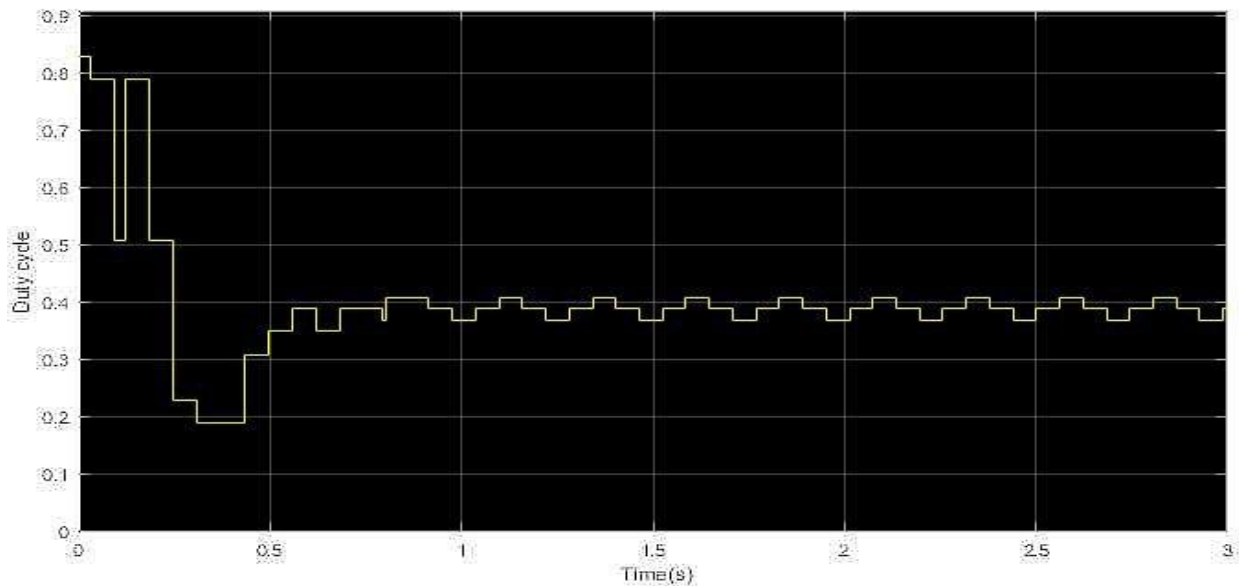
**Fig. 5-37.** Duty cycle versus time for shading pattern 4 of the third implementation of the Q-learning-based MPPT method.



**Fig. 5-38.** PV array output power versus time for shading pattern 4 of the third implementation of the Q-learning-based MPPT method.

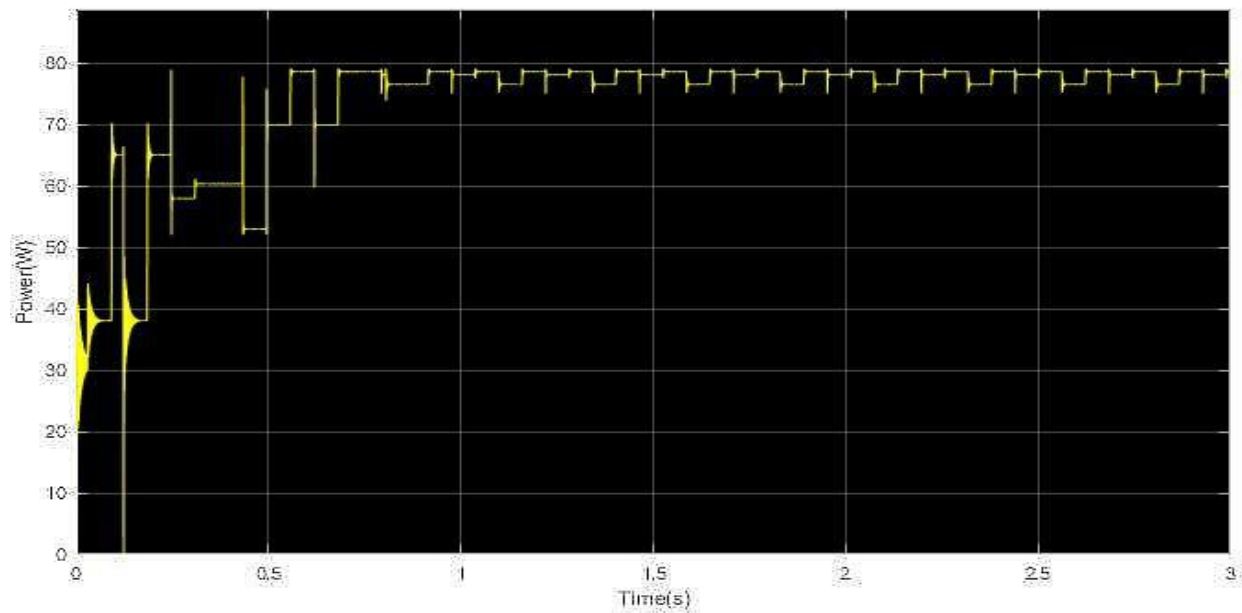
It is observed that in the third implementation of the Q-learning-based method the agent is able to detect the GMPP in unknown intermediate conditions of solar irradiation of shading patterns 1 and 2 in a few time-steps.

Figs. 5-39 and 5-40 present the results of the simulations of duty cycle and PV array output power versus time of the fourth implementation of the Q-learning-based MPPT method for shading pattern 4.



**Fig. 5-39.** PV array output power versus time for shading pattern 4 of the fourth implementation of the Q-learning-based MPPT method.



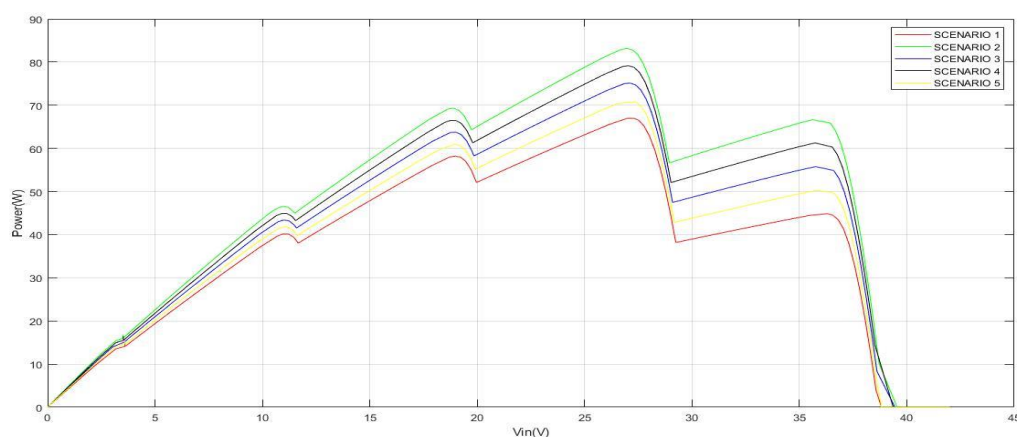


**Fig. 5-40.** PV array output power versus time for shading pattern 4 of the fourth implementation of the Q-learning-based MPPT method.

It is observed that after the MPPT process for shading patterns 1,2 and 3 of the fourth implementation of the Q-learning-based method, the agent is able to detect the GMPP in a smaller amount of time in intermediate conditions of solar irradiation, similar to the shading pattern 4.

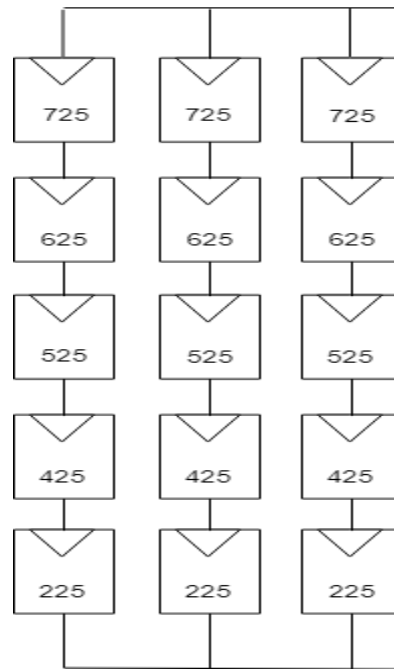
## 5.5 Analysis of the simulation results for shading pattern 5

In Fig. 5-41 the PV array output power-voltage curve is presented and in Fig. 5-42 the distribution of incident solar irradiation over each PV module is presented. The numbers above each PV module represent the solar irradiation intensity, which is measured in  $W/m^2$ . Shading pattern 5 is a non-uniform solar irradiance pattern.



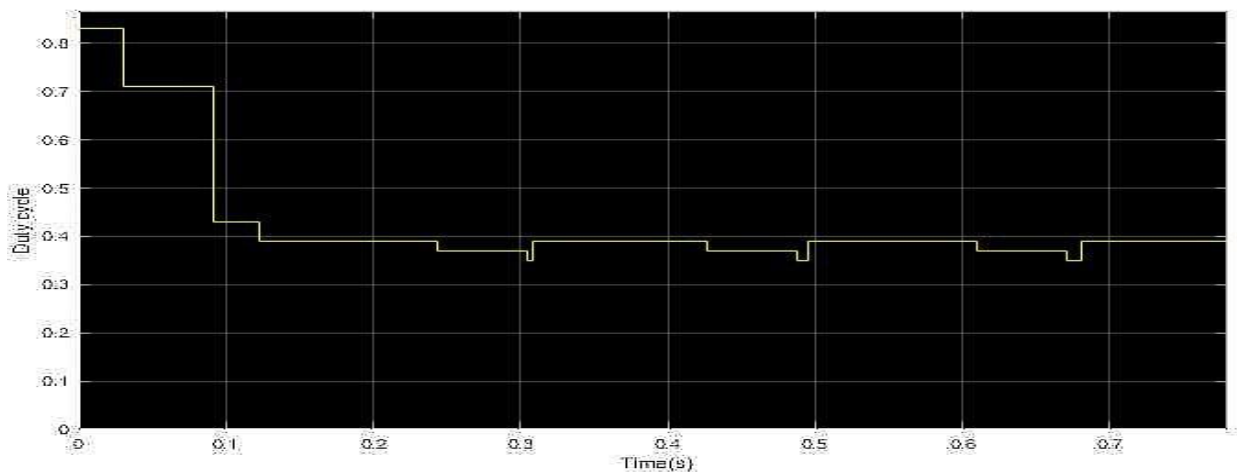
**Fig. 5-41.** The PV array output power-voltage curve for shading pattern 5.



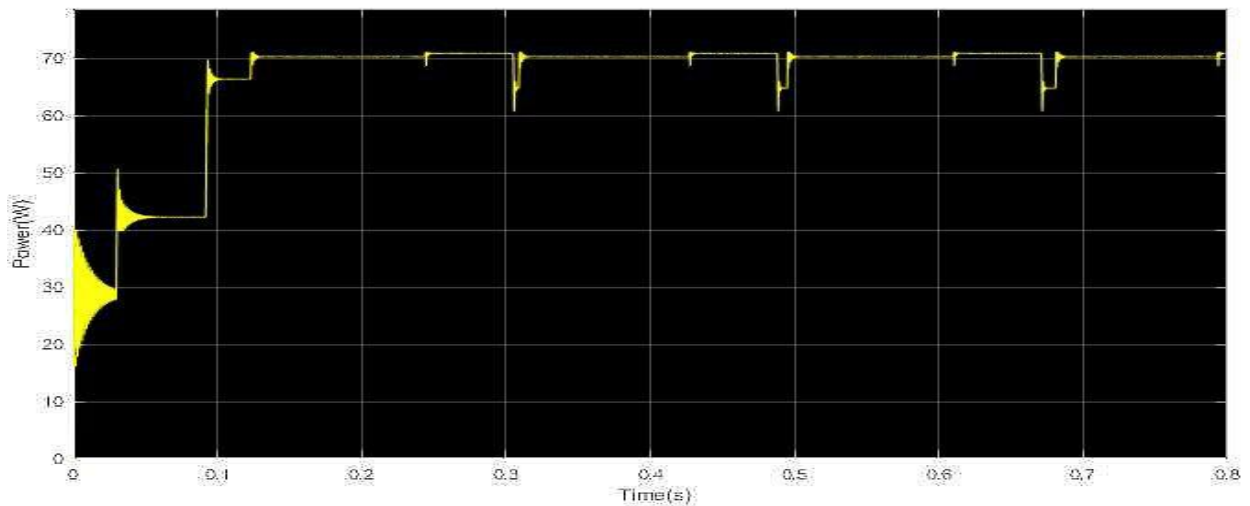


**Fig. 5-42.** The distribution of incident solar irradiation over each PV module for shading pattern 5.

Eight plots were exported, the PV array output power versus time plot and the duty cycle versus time plot in the four aforementioned implementations of the Q-learning-based MPPT method. Figs. 5-43 and 5-44 present the results of the simulations of duty cycle and PV array output power versus time for the first implementation of the Q-learning-based MPPT method for shading pattern 5.



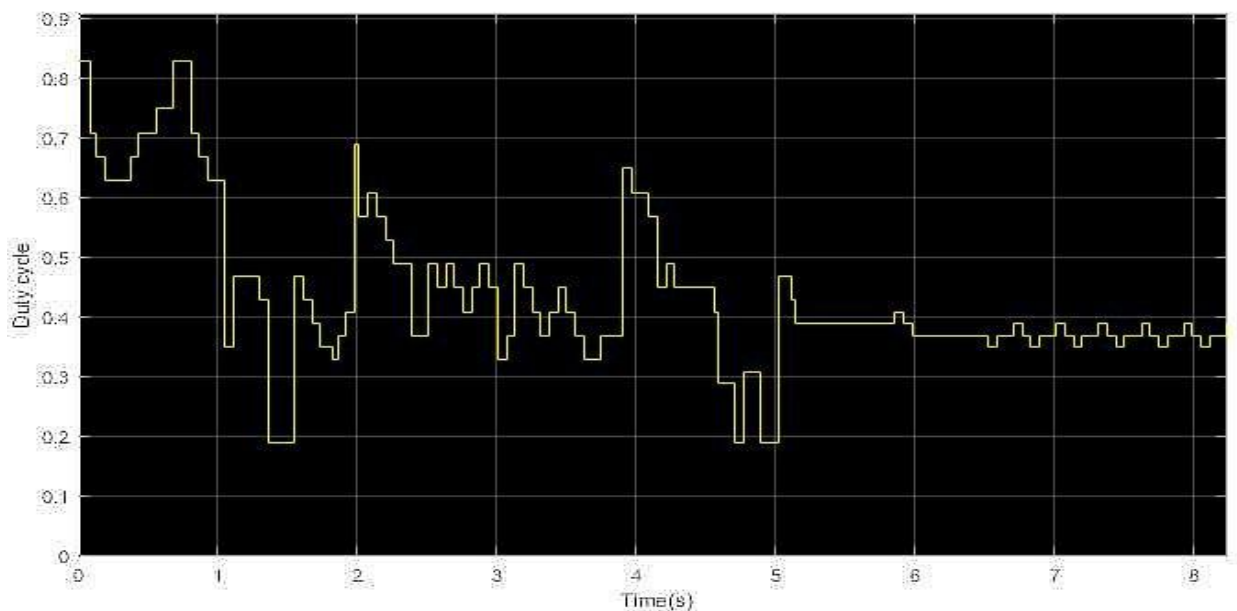
**Fig. 5-43.** Duty cycle versus time for shading pattern 5 of the first implementation of the Q-learning-based MPPT method.



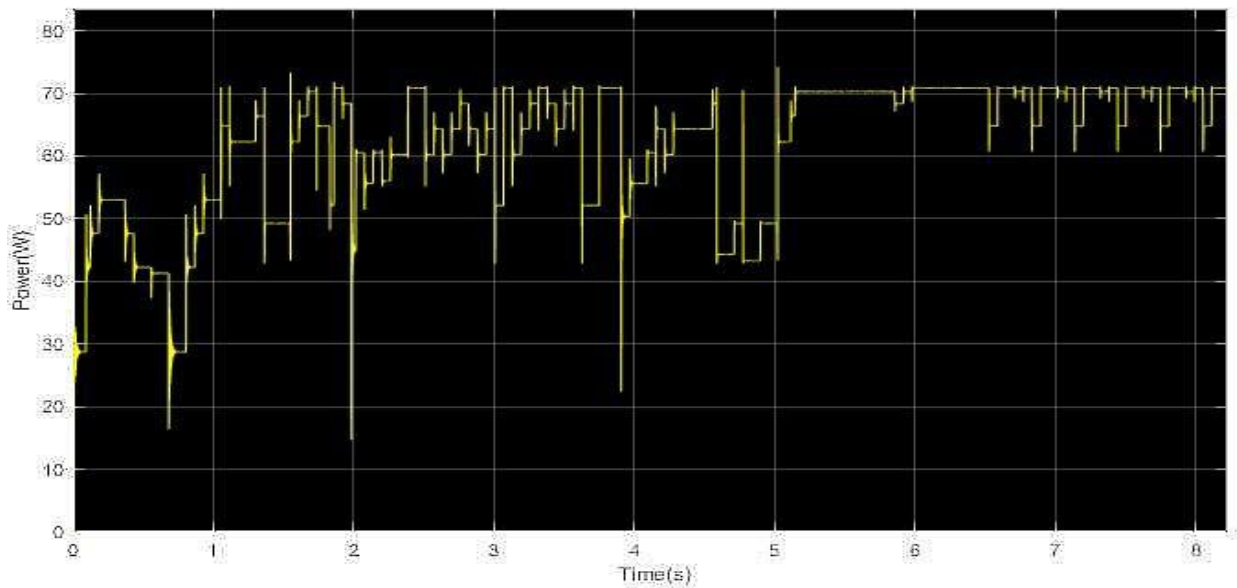
**Fig. 5-44.** PV array output power versus time for shading pattern 5 of the first implementation of the Q-learning-based MPPT method.

It is observed that in the first implementation of the Q-learning-based method the agent is able to detect the GMPP in unknown intermediate conditions of solar irradiation of shading patterns 1 and 2 in a few time-steps.

Figs. 5-45 and 5-46 present the results of the simulations of duty cycle and PV array output power versus time of the second implementation of the Q-learning-based MPPT method for shading pattern 5.



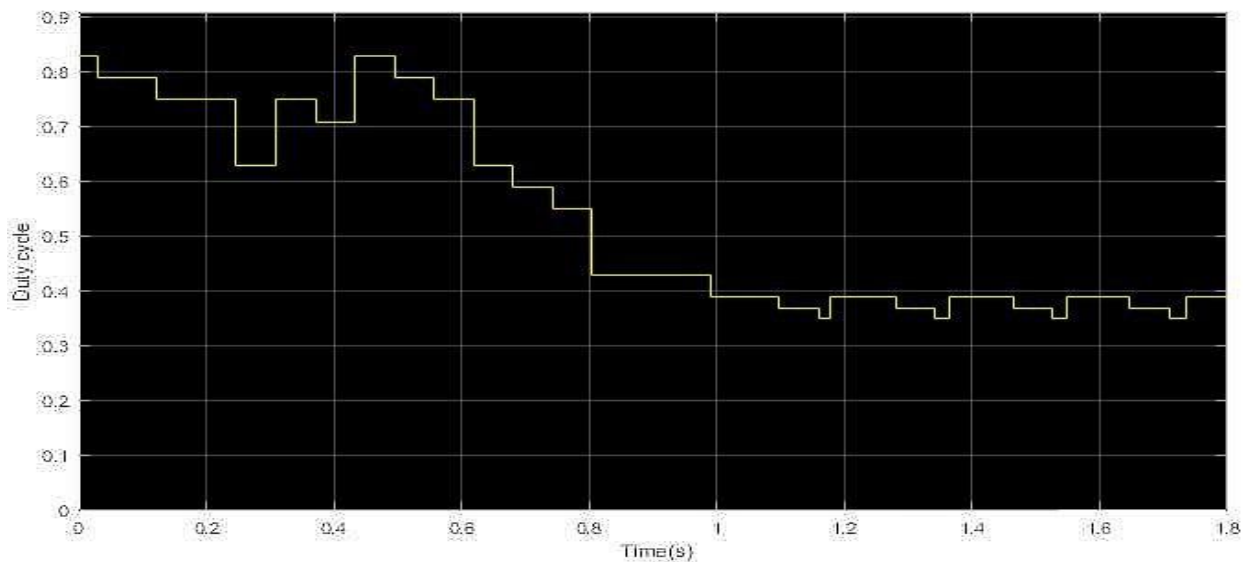
**Fig. 5-45.** Duty cycle versus time for shading pattern 5 of the second implementation of the Q-learning-based MPPT method.



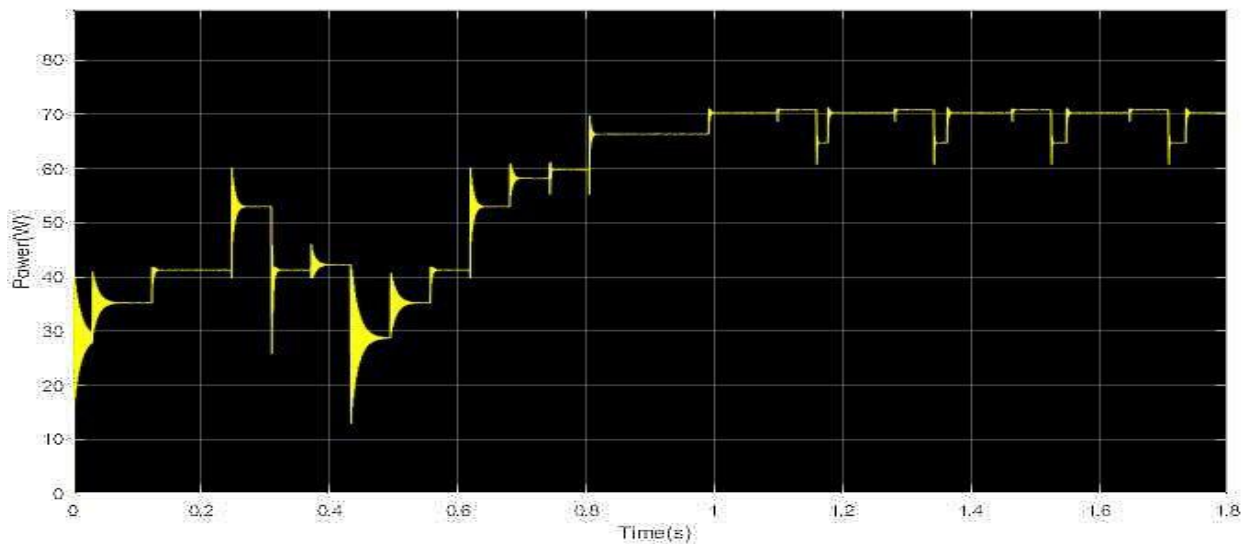
**Fig. 5-46.** PV array output power versus time for shading pattern 5 of the second implementation of the Q-learning-based MPPT method.

It is observed that in the second implementation of the Q-learning-based method the required time for the MPP detection for shading pattern 5 is even further reduced.

Figs. 5-47 and 5-48 present the results of the simulations of duty cycle and PV array output power versus time of the third implementation of the Q-learning-based MPPT method for shading pattern 5.



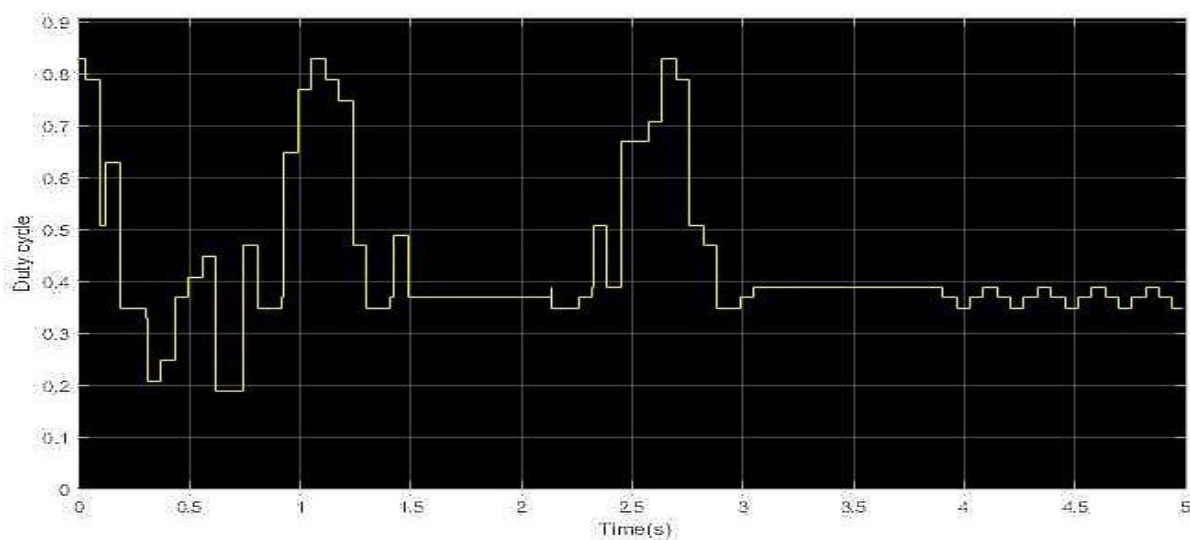
**Fig. 5-47.** Duty cycle versus time for shading pattern 5 of the third implementation of the Q-learning-based MPPT method.



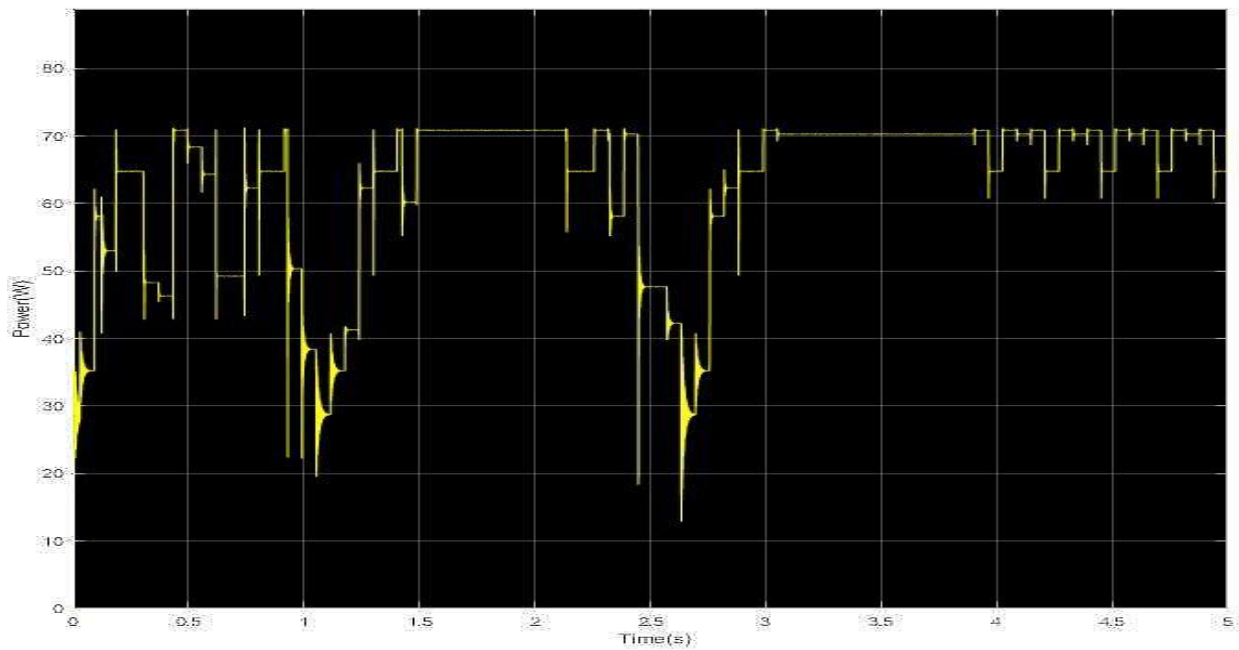
**Fig. 5-48.** Duty cycle versus time for shading pattern 5 of the third implementation of the Q-learning-based MPPT method.

It is observed that in the third implementation of the Q-learning-based method the agent is able to detect the MPP in unknown intermediate conditions of solar irradiation of shading patterns 1 and 2 in a few time-steps.

Figs. 5-49 and 5-50 present the results of the simulations of duty cycle and PV array output power versus time of the fourth implementation of the Q-learning-based MPPT method for shading pattern 5.



**Fig. 5-49.** Duty cycle versus time for shading pattern 5 of the fourth implementation of the Q-learning-based MPPT method.

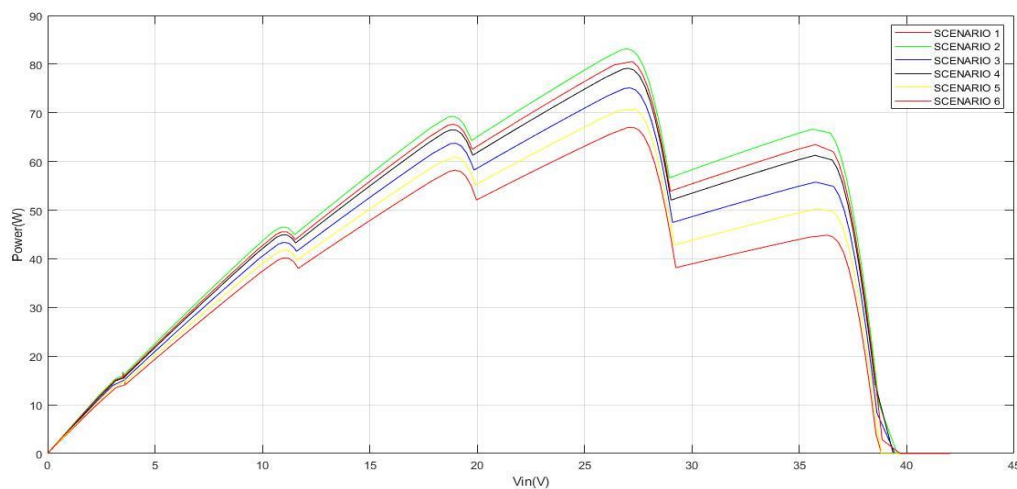


**Fig. 5-50.** PV array output power versus time for shading pattern 5 of the fourth implementation of the Q-learning-based MPPT method.

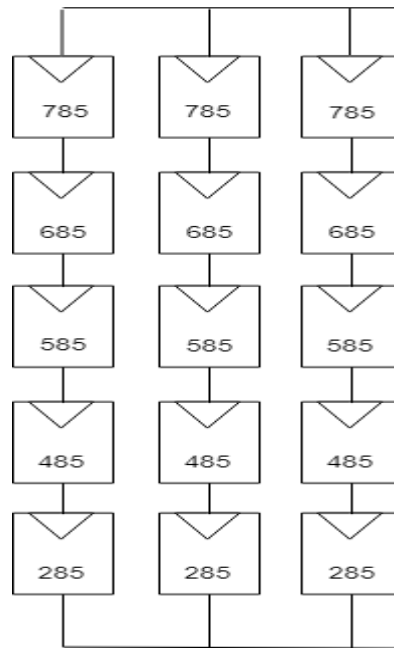
It is observed that in the fourth implementation of the Q-learning-based method the required time for the MPP detection for shading pattern 5 is even further reduced.

## 5.6 Analysis of the simulation results for shading pattern 6

In Fig. 5-51 the PV array output power-voltage curve is presented and in Fig. 5-52 the distribution of incident solar irradiation over each PV module is presented. The numbers above each PV module represent the solar irradiation intensity, which is measured in  $W/m^2$ . Shading pattern 6 is a non-uniform solar irradiance pattern.

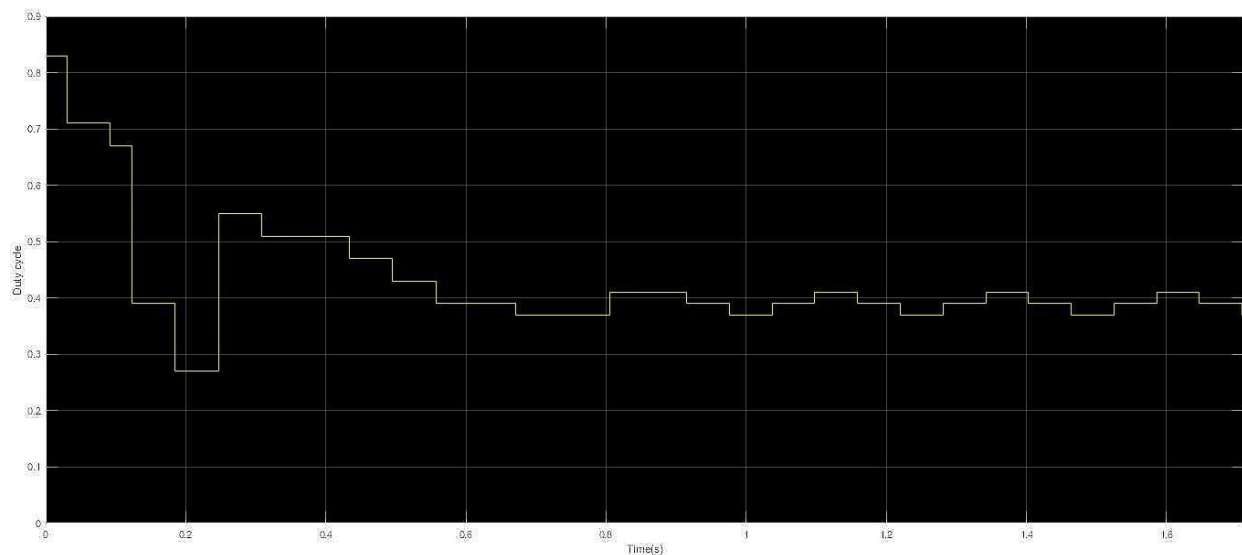


**Fig. 5-51.** The PV array output power-voltage curve for shading pattern 6.

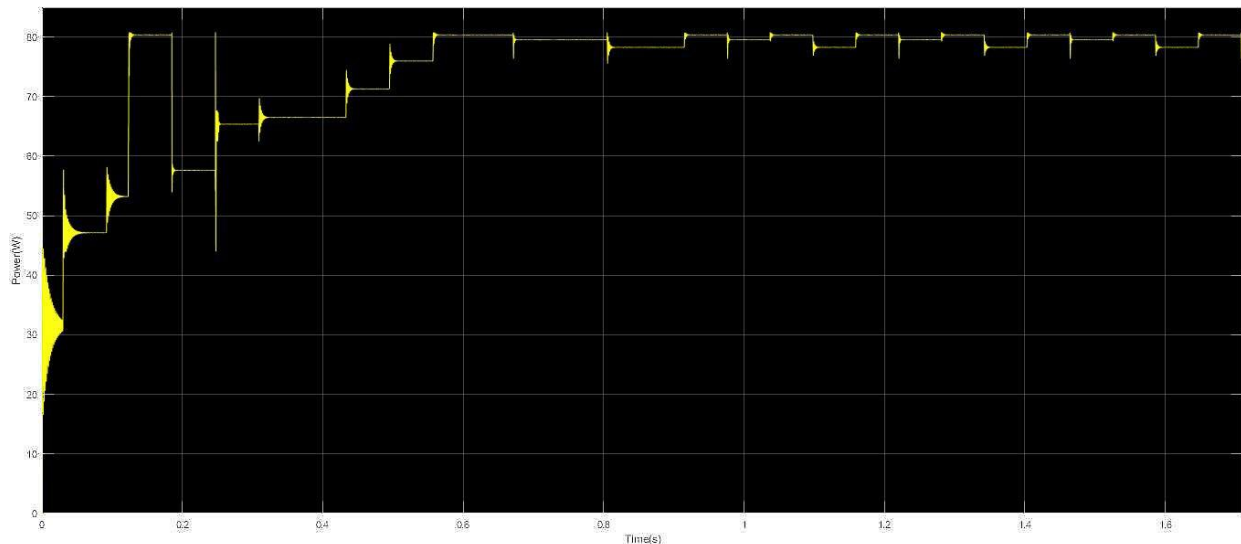


**Fig. 5-52.** The distribution of incident solar irradiation over each PV module for shading pattern 6.

Eight plots were exported, the PV array output power versus time plot and the duty cycle versus time plot in the four aforementioned implementations of the Q-learning-based MPPT method. Figs. 5-53 and 5-54 present the results of the simulations of the duty cycle and the PV array output power versus time for the first implementation of the Q-learning-based MPPT method for shading pattern 6.



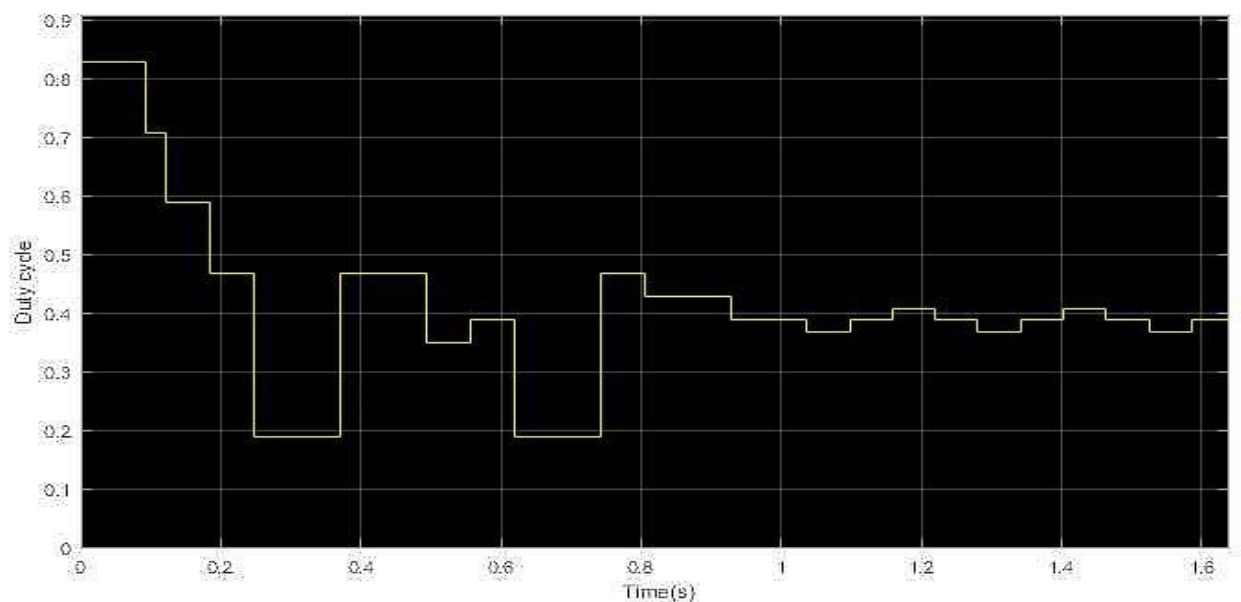
**Fig. 5-53.** Duty cycle versus time for shading pattern 6 of the first implementation of the Q-learning-based MPPT method.



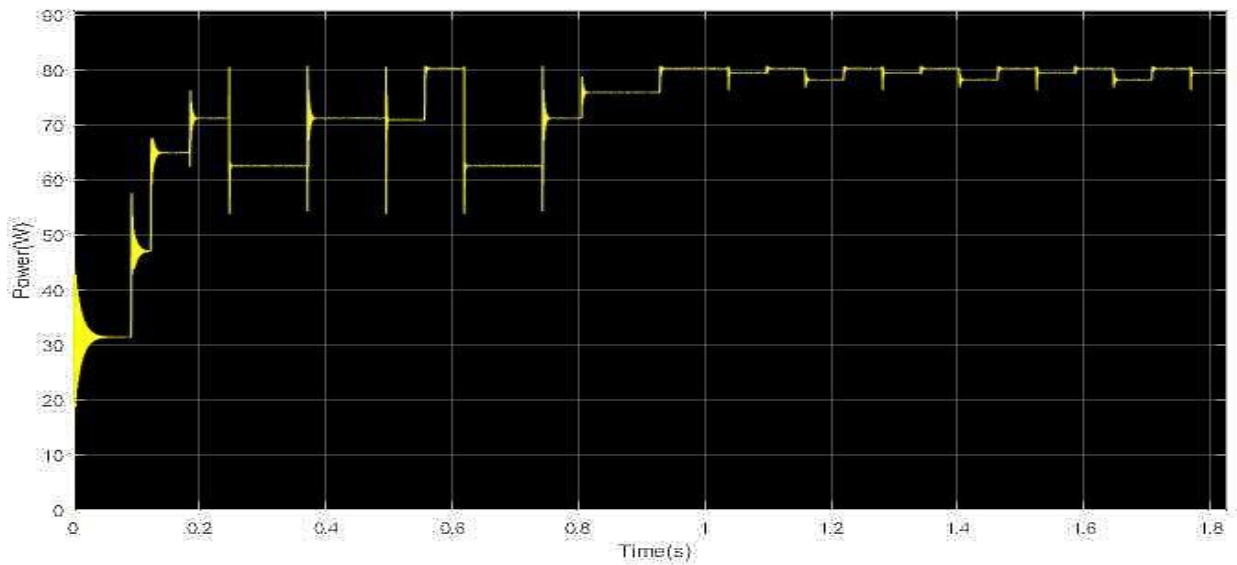
**Fig. 5-54.** PV array output power versus time for shading pattern 6 of the first implementation of the Q-learning-based MPPT method.

It is observed that in the first implementation of the Q-learning-based method the agent is able to detect the MPP in unknown intermediate conditions of solar irradiation of shading patterns 1 and 2 in a few time-steps.

Figs. 5-55 and 5-56 present the results of the simulations of duty cycle and PV array output power versus time of the second implementation of the Q-learning-based MPPT method for shading pattern 6.



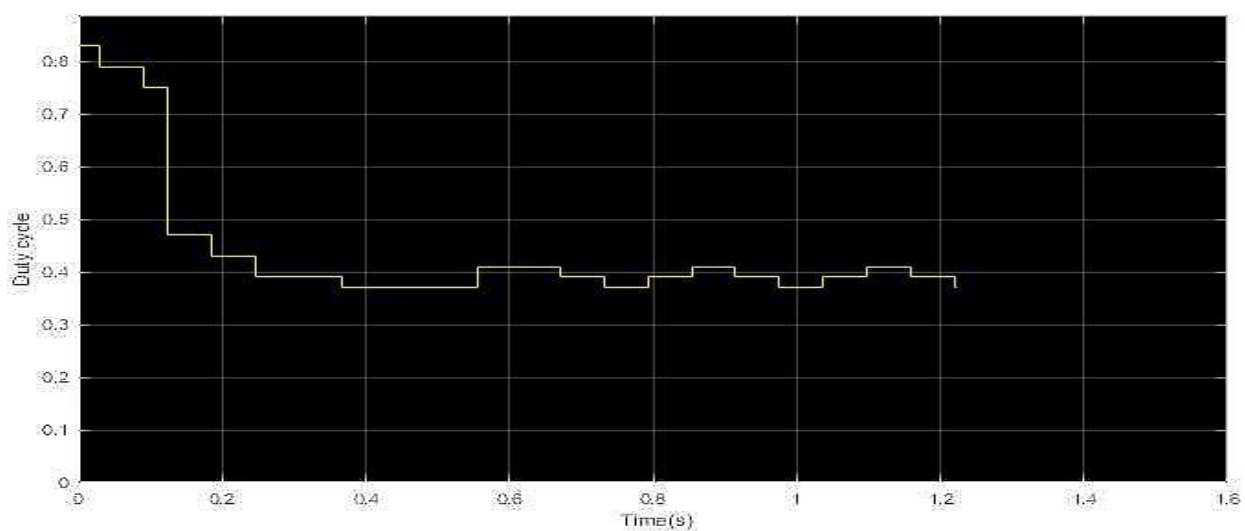
**Fig. 5-55.** Duty cycle versus time for shading pattern 6 of the second implementation of the Q-learning-based MPPT method.



**Fig. 5-56.** PV array output power versus time for shading pattern 6 of the second implementation of the Q-learning-based MPPT method.

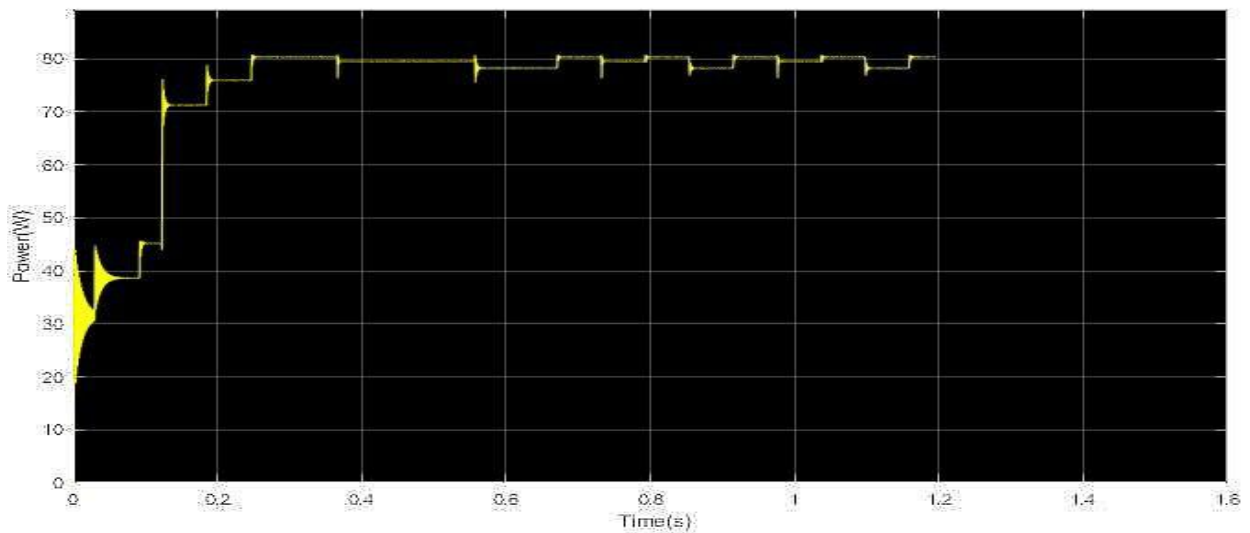
It is observed that from now onwards the agent of the second implementation of the Q-learning-based MPPT method is able to detect the GMPP in unknown intermediate conditions of solar irradiation of shading patterns 1 and 2 in a few time-steps. The time required for the MPPT process is equal to the required time for the MPPT process of the first implementation of the Q-learning-based MPPT method.

Figs. 5-57 and 5-58 present the results of the simulations of duty cycle and PV array output power versus time of the third implementation of the Q-learning-based MPPT method for shading pattern 6.



**Fig. 5-57.** Duty cycle versus time for shading pattern 6 of the third implementation of the Q-learning-based MPPT method.

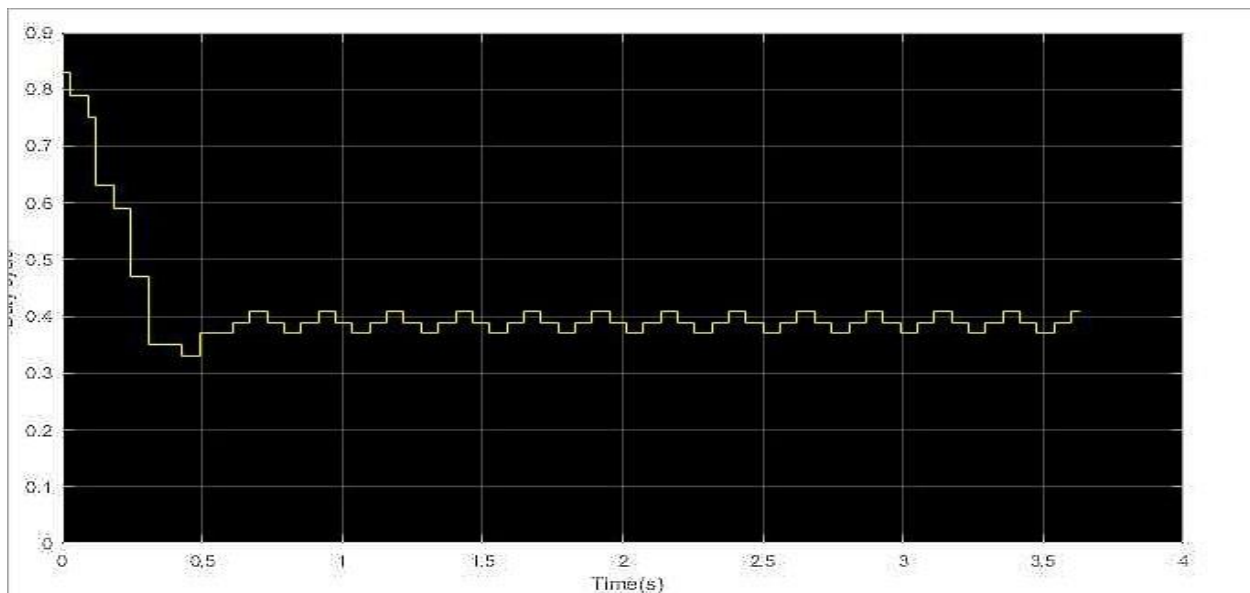




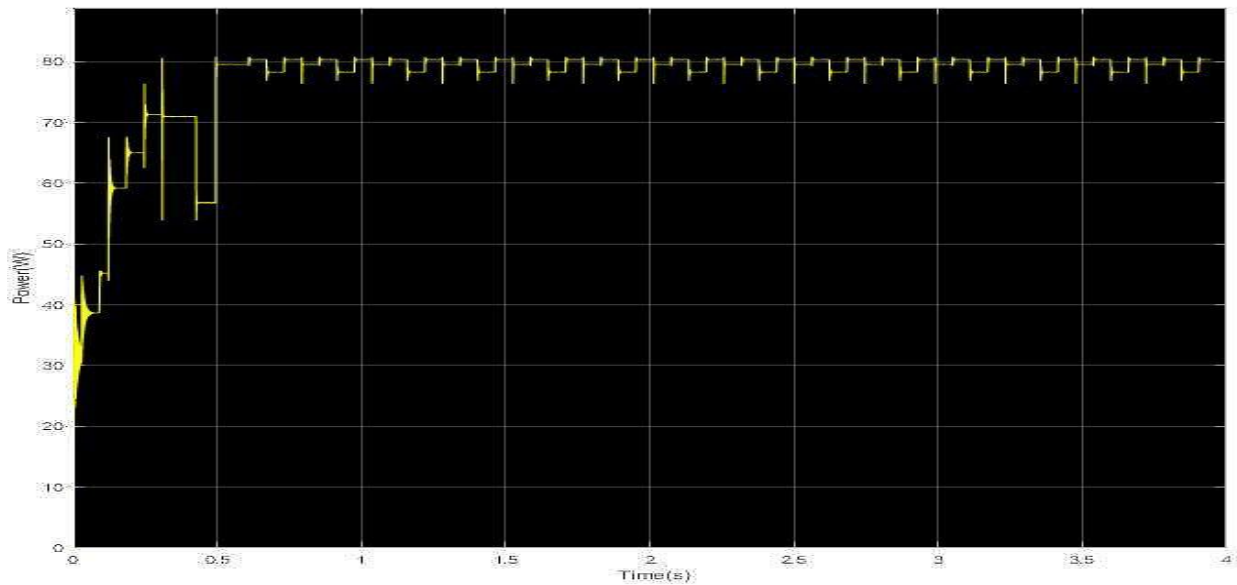
**Fig. 5-58.** PV array output power versus time for shading pattern 6 of the third implementation of the Q-learning-based MPPT method.

It is observed that in the third implementation of the Q-learning-based method the agent is able to detect the GMPP in unknown intermediate conditions of solar irradiation of shading patterns 1 and 2 in a few time-steps.

Figs. 5-59 and 5-60 present the results of the simulations of duty cycle and PV array output power versus time of the fourth implementation of the Q-learning-based MPPT method for shading pattern 6.



**Fig. 5-59.** Duty cycle versus time for shading pattern 6 of the fourth implementation of the Q-learning-based MPPT method.

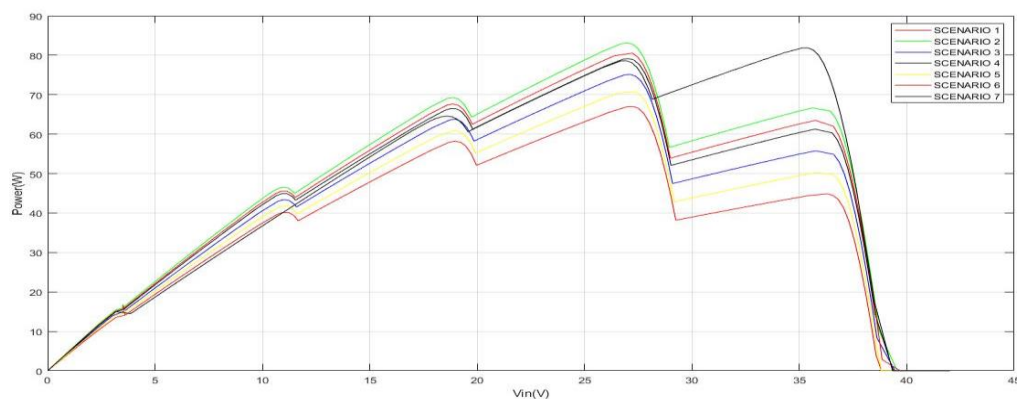


**Fig. 5-60.** PV array output power versus time for shading pattern 6 of the fourth implementation of the Q-learning-based MPPT method.

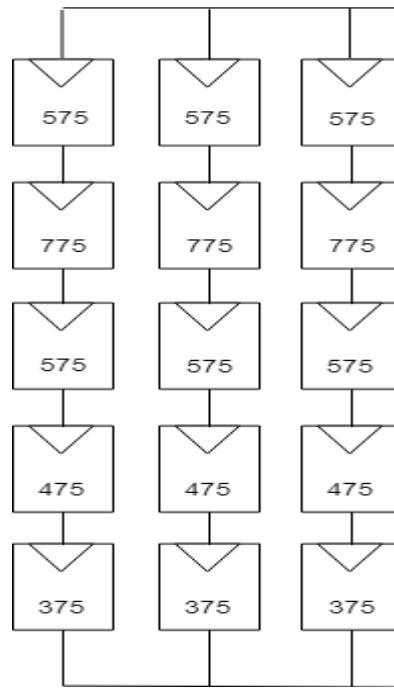
It is observed that from now onwards the fourth implementation of the Q-learning-based MPPT method is able to detect the MPP in unknown intermediate conditions of solar irradiation of shading patterns 1 and 2 in a few time-steps. The time required for the MPPT process is equal to the required time for the MPPT process of the third implementation of the Q-learning-based MPPT method.

## 5.7 Analysis of the simulation results for shading pattern 7

In Fig. 5-61 the PV array output power-voltage curve is presented and in Fig. 5-62 the distribution of incident solar irradiation over each PV module is presented. The numbers above each PV module represent the solar irradiation intensity, which is measured in  $W/m^2$ . Shading pattern 7 is a non-uniform solar irradiance pattern.

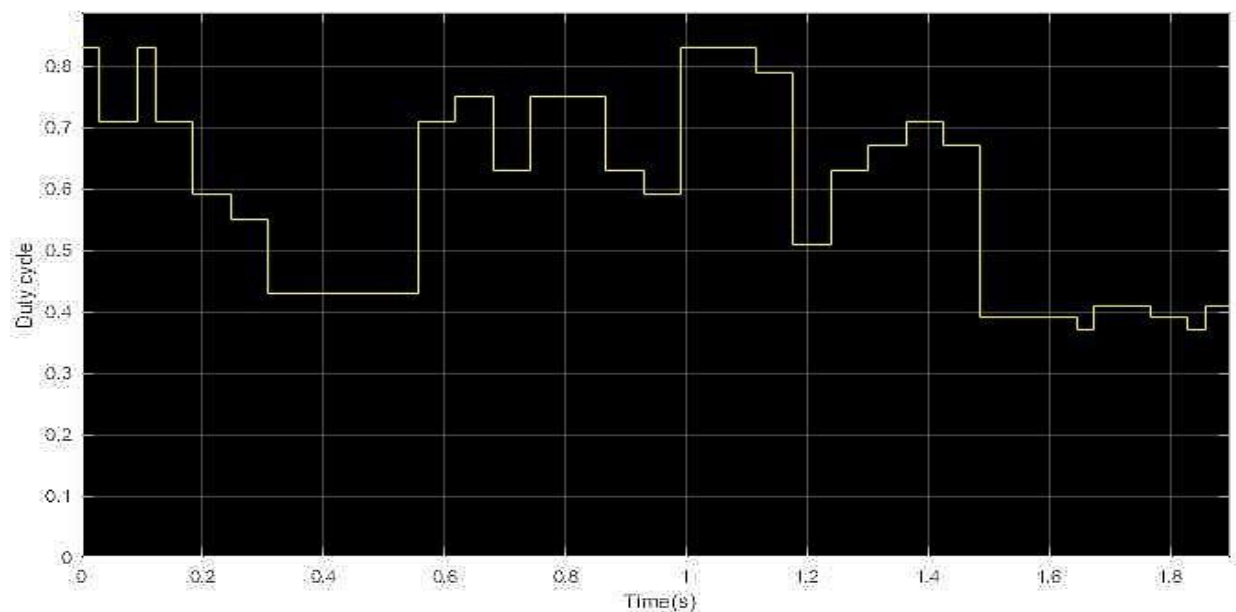


**Fig. 5-61.** The PV array output power-voltage curve for shading pattern 7.

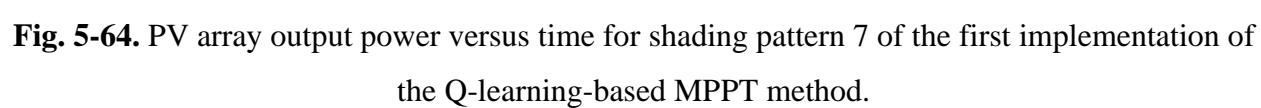


**Fig. 5-62.** The distribution of incident solar irradiation over each PV module for shading pattern 7.

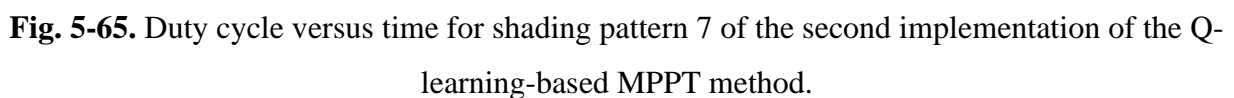
Eight plots were exported, the PV array output power versus time plot and the duty cycle versus time plot in the four aforementioned implementations of the Q-learning-based MPPT method. Figs. 5-63 and 5-64 present the results of the simulations of the duty cycle and the PV array output power versus time for the first implementation of the Q-learning-based MPPT method for shading pattern 7.

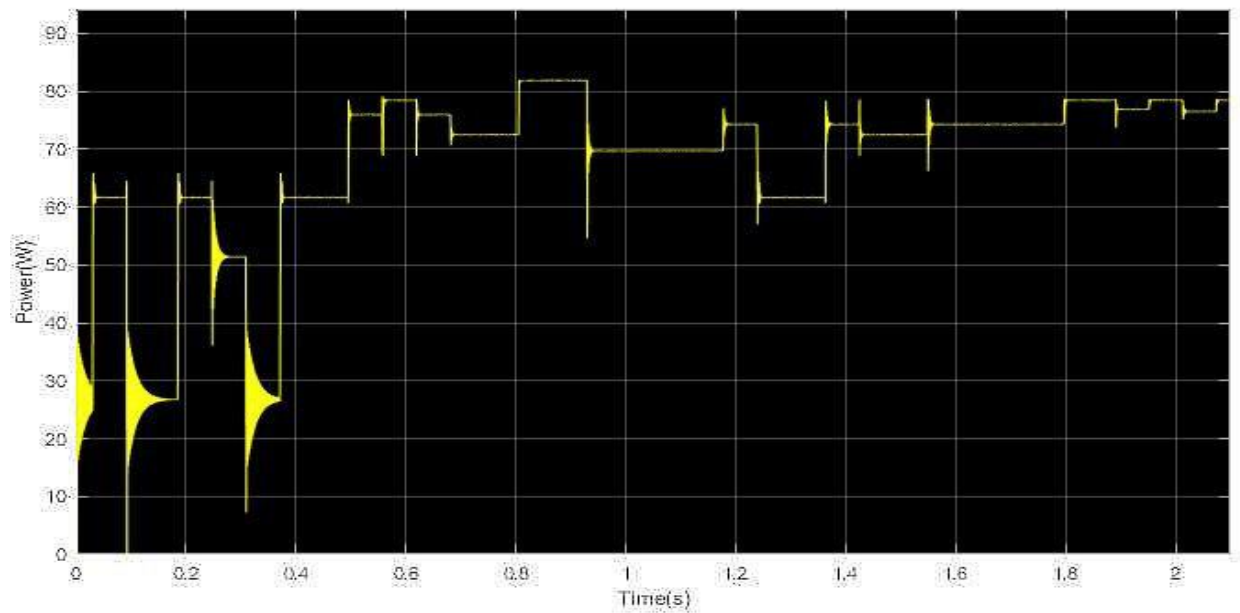


**Fig. 5-63.** Duty cycle versus time for shading pattern 7 of the first implementation of the Q-learning-based MPPT method.



Figs. 5-65 and 5-66 present the results of the simulations of duty cycle and PV array output power versus time of the second implementation of the Q-learning-based MPPT method for shading pattern 7.

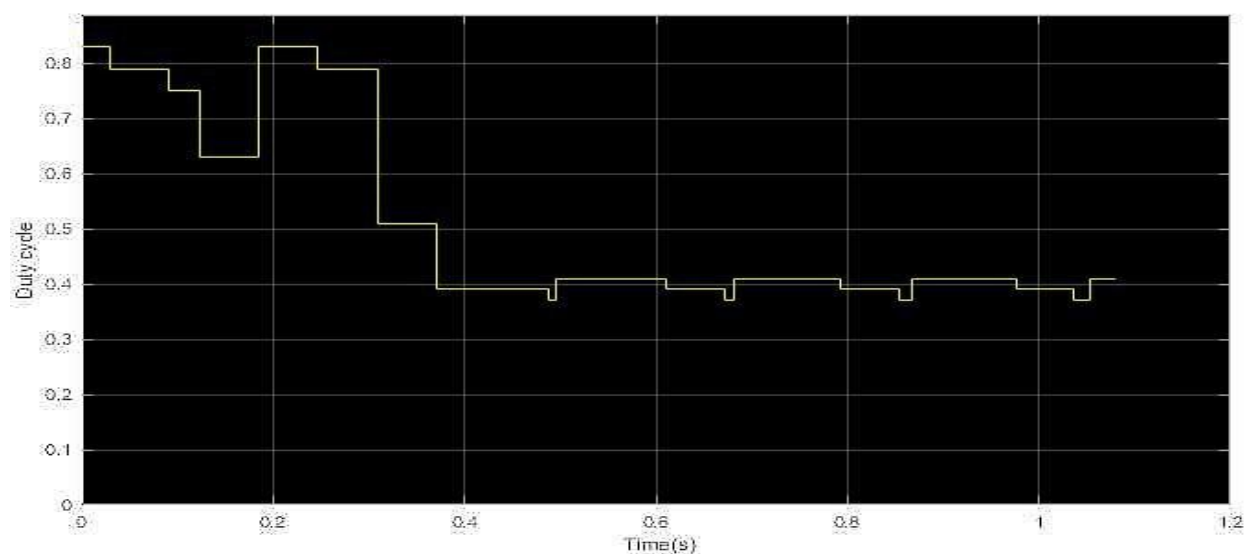




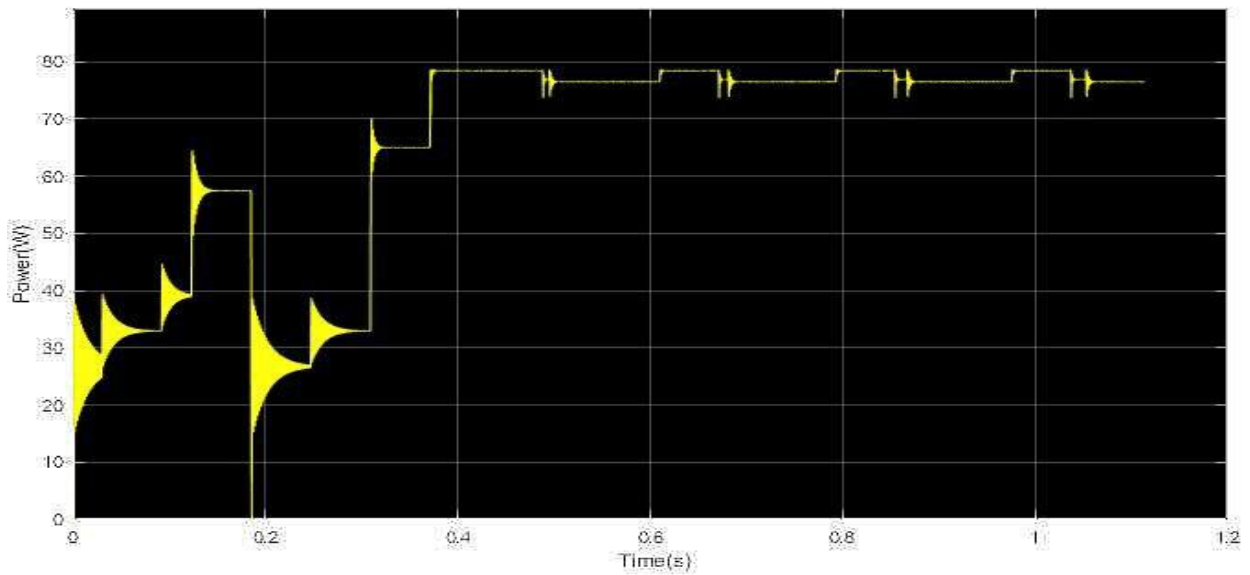
**Fig. 5-66.** PV array output power versus time for shading pattern 7 of the second implementation of the Q-learning-based MPPT method.

The GMPP of the PV array output power for shading pattern 7 is equal to 81 W, however the agent converges to a point that the PV array output power is less than 80 W. In this shading pattern, the prior knowledge leads the agent to a point, which is far from the MPP. Nevertheless, the MPP was effectively approached in a small amount of time.

Figs. 5-67 and 5-68 present the results of the simulations of duty cycle and PV array output power versus time of the third implementation of the Q-learning-based MPPT method for shading pattern 7.



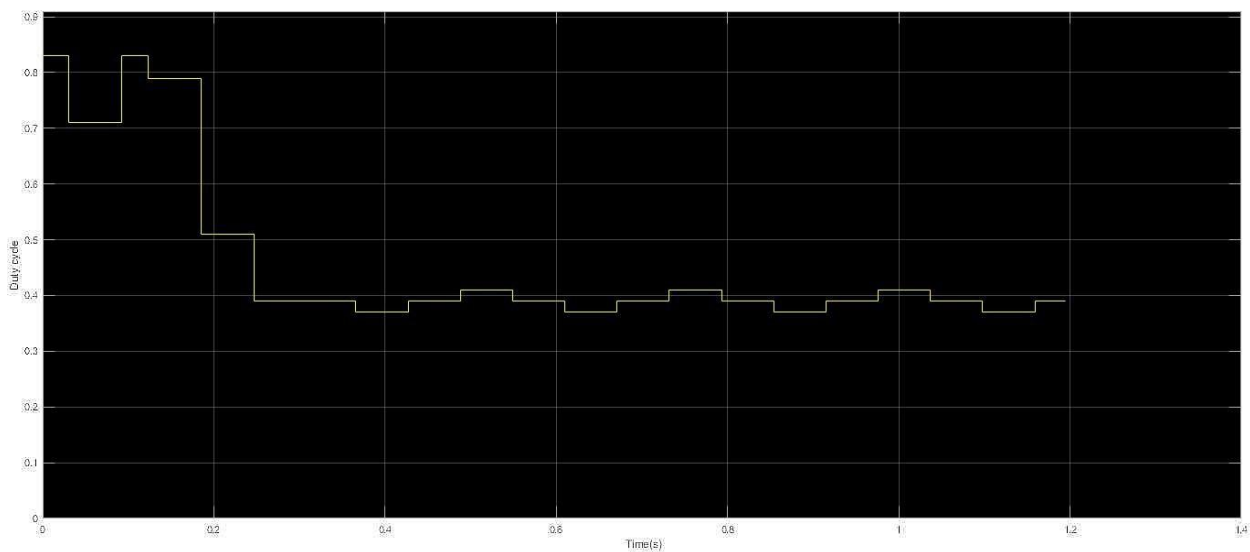
**Fig. 5-67.** Duty cycle versus time for shading pattern 7 of the third implementation of the Q-learning-based MPPT method.



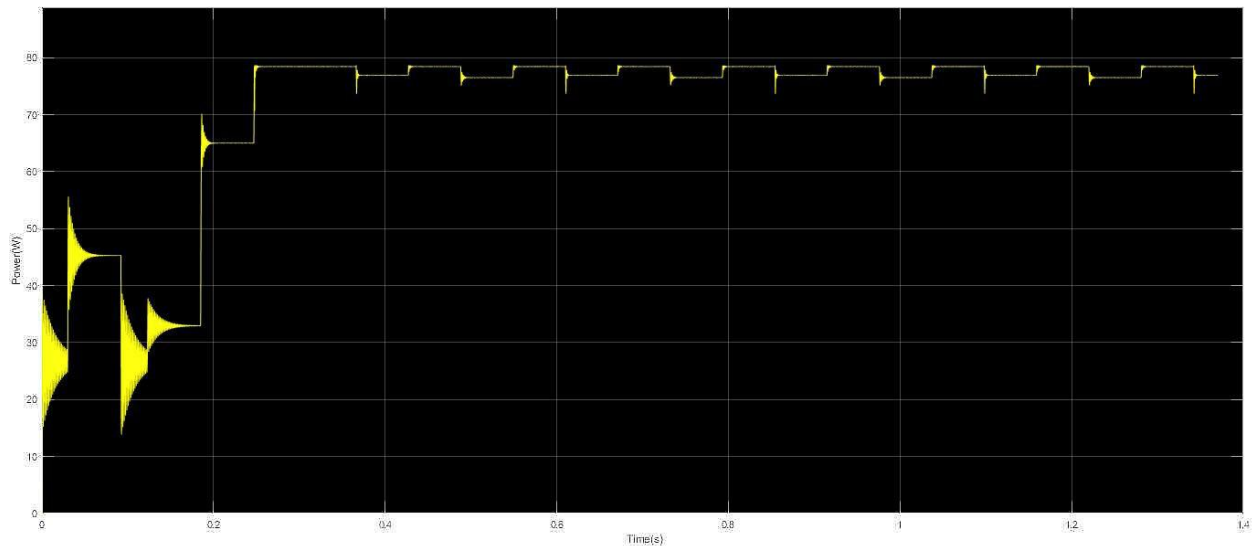
**Fig. 5-68.** PV array output power versus time for shading pattern 7 of the third implementation of the Q-learning-based MPPT method.

The GMPP of the PV array output power for shading pattern 7 is equal to 81 W, however the agent converges to a point that the PV array output power is less than 80 W. In this shading pattern, the prior knowledge leads the agent to a point far from the MPP. Nevertheless, the MPP was effectively approached in a small amount of time.

Figs. 5-69 and 5-70 present the results of the simulations of duty cycle and PV array output power versus time of the fourth implementation of the Q-learning-based MPPT method for shading pattern 7.



**Fig. 5-69.** Duty cycle versus time for shading pattern 7 of the fourth implementation of the Q-learning-based MPPT method.

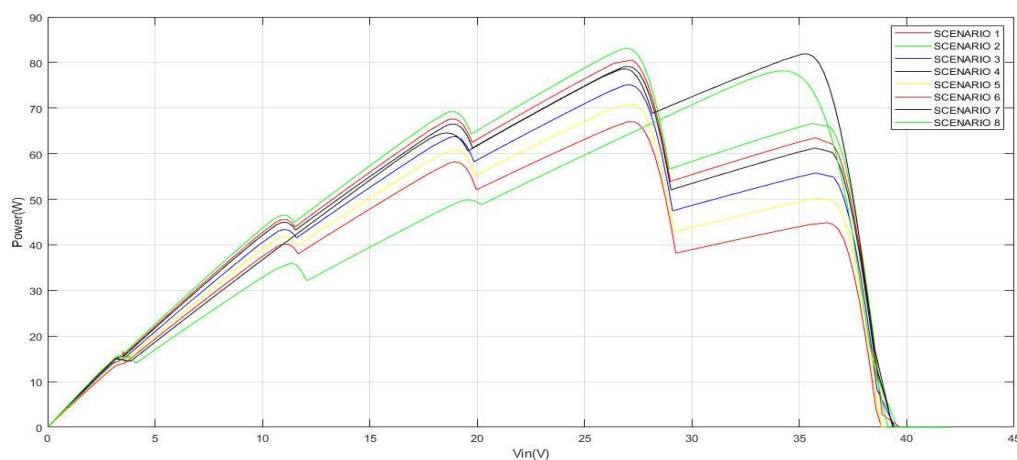


**Fig. 5-70.** PV array output power versus time for shading pattern 7 of the fourth implementation of the Q-learning-based MPPT method.

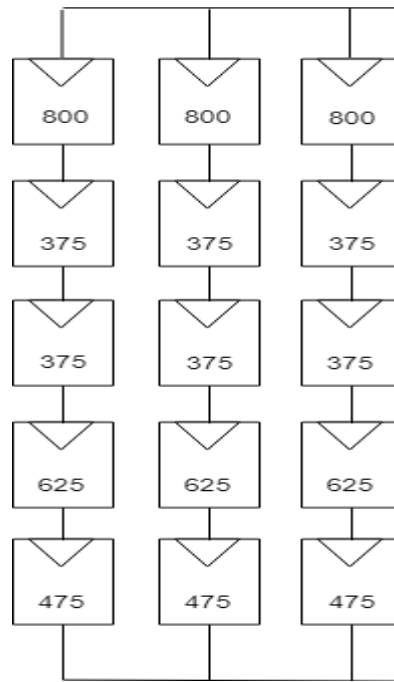
The GMPP of the PV array output power for shading pattern 7 is equal to 81 W, however the agent converges to a point that the PV array output power is less than 80 W. In this shading pattern, the prior knowledge leads the agent to a point far from the MPP. Nevertheless, the MPP was effectively approached in a small amount of time.

## 5.8 Analysis of the simulation results for shading pattern 8

In Fig. 5-71 the PV array output power-voltage curve is presented and in Fig. 5-72 the distribution of incident solar irradiation over each PV module is presented. The numbers above each PV module represent the solar irradiation intensity, which is measured in  $W/m^2$ . Shading pattern 8 is a non-uniform solar irradiance pattern.

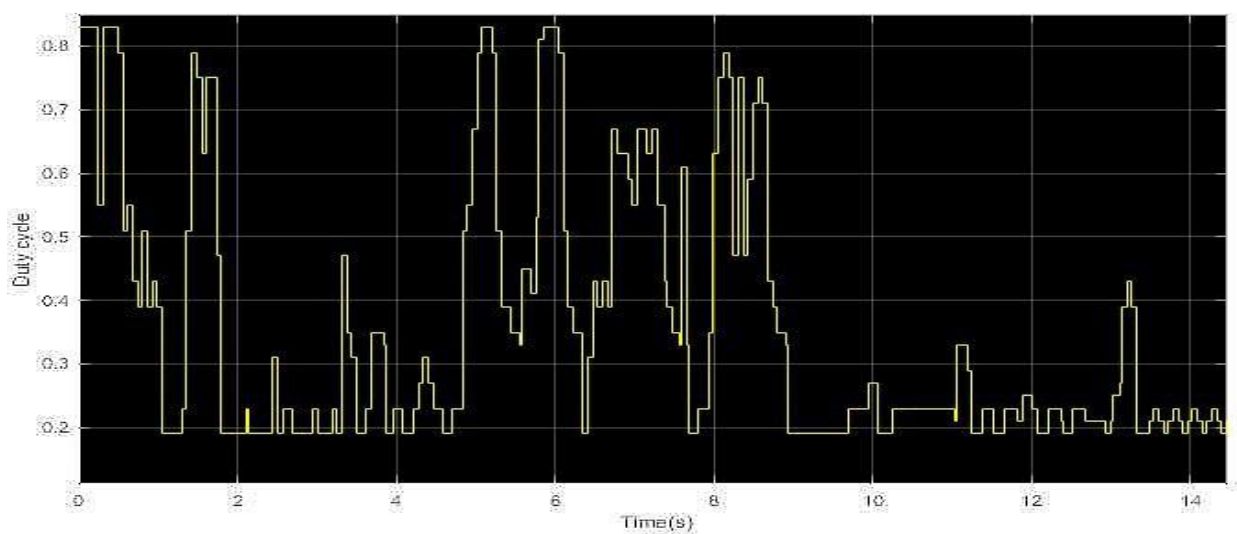


**Fig. 5-71.** The PV array output power-voltage curve for shading pattern 8.



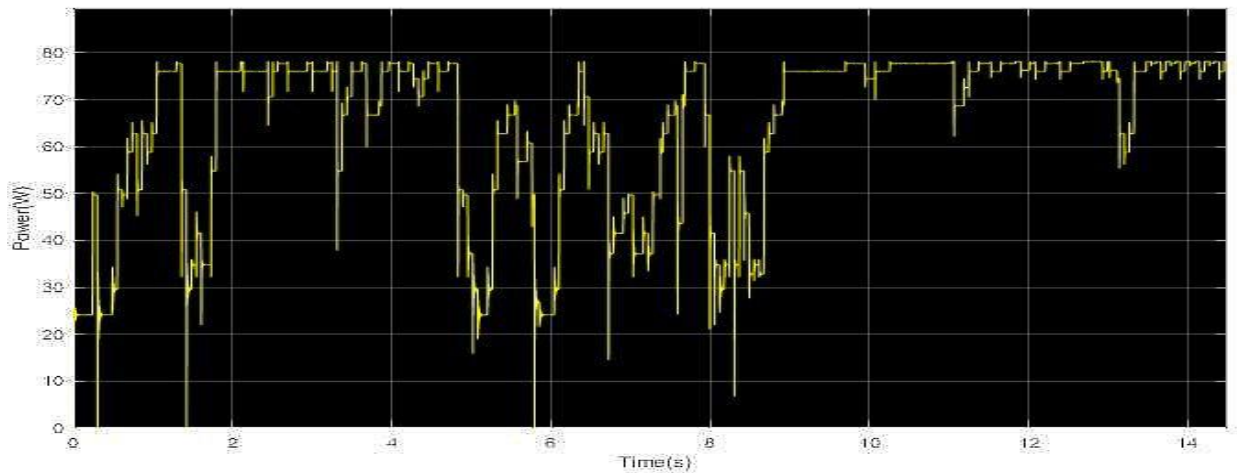
**Fig. 5-72.** The distribution of incident solar irradiation over each PV module for shading pattern 8.

Eight plots were exported, the PV array output power versus time plot and the duty cycle versus time plot in the four aforementioned implementations of the Q-learning-based MPPT method. Figs. 5-73 and 5-74 present the results of the simulations of the duty cycle and the PV array output power versus time for the first implementation of the Q-learning-based MPPT method for shading pattern 8. In shading pattern 8, GMPP is beyond the boundaries of the power-voltage curves that the agent had previously explored.



**Fig. 5-73.** Duty cycle versus time for shading pattern 8 of the first implementation of the Q-learning-based MPPT method.

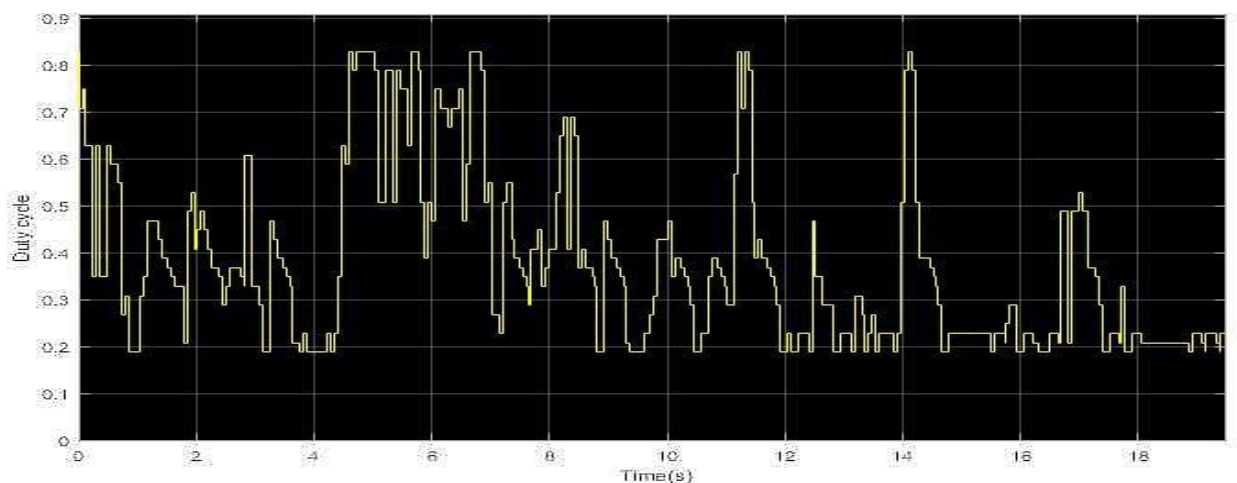




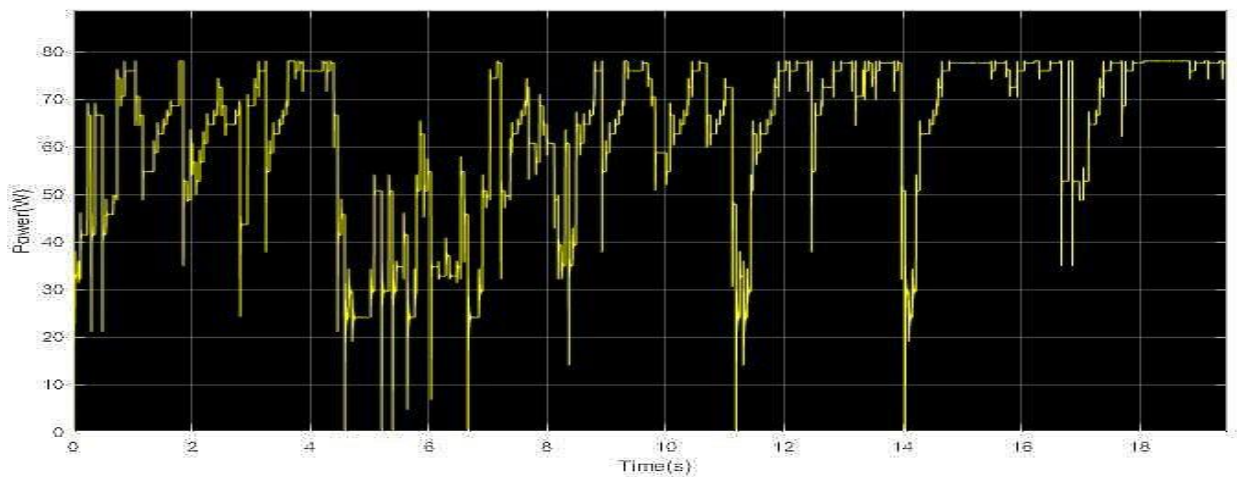
**Fig. 5-74.** Duty cycle versus time for shading pattern 8 of the first implementation of the Q-learning-based MPPT method.

It is observed that in the first implementation of the Q-learning-based MPPT method the agent needs more time to locate the GMPP for shading pattern 8 compared with the required time to locate the GMPP in the previous shading pattern. However, the MPPT process lasts a smaller amount of time compared with shading pattern 1, in which the agent had no prior knowledge of the location of the GMPP. This leads to the conclusion that the agent was able to use its previous knowledge from the GMPP detection in unknown intermediate conditions of solar irradiation of shading patterns 1 and 2. The agent explored only the states that it visited for the first time.

Figs. 5-75 and 5-76 present the results of the simulations of duty cycle and PV array output power versus time of the second implementation of the Q-learning-based MPPT method for shading pattern 8.



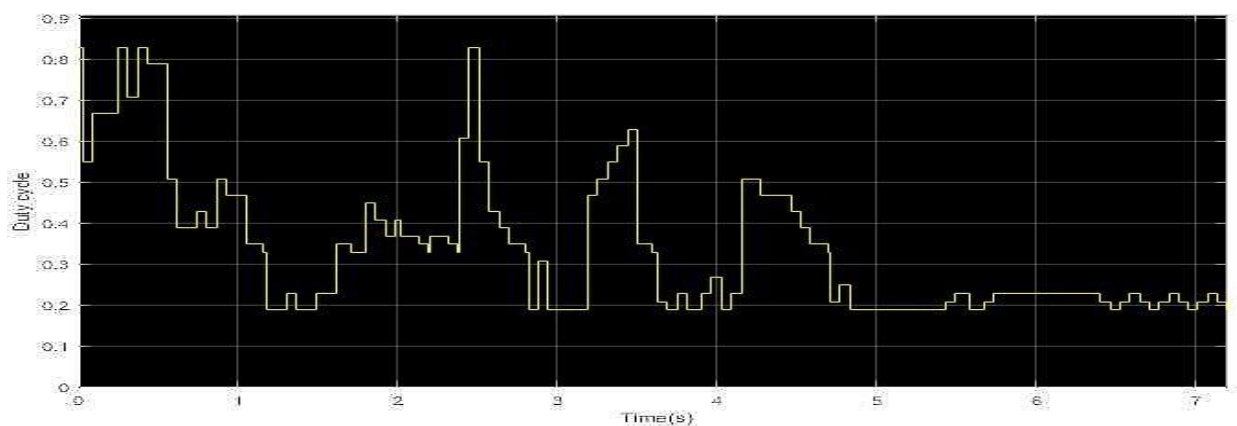
**Fig. 5-75.** Duty cycle versus time for shading pattern 8 of the second implementation of the Q-learning-based MPPT method.



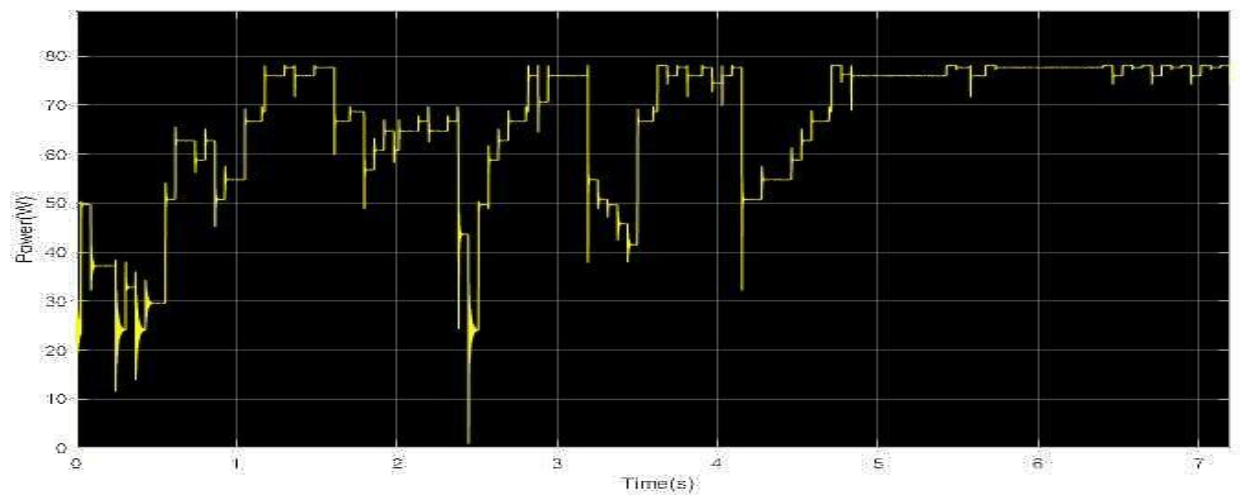
**Fig. 5-76.** PV array output power versus time for shading pattern 8 of the second implementation of the Q-learning-based MPPT method.

It is observed that in the first implementation of the Q-learning-based MPPT method the agent needs a larger amount of time to detect the location of the MPP for shading pattern 8 compared with the required time to locate the MPP in the previous shading pattern. However, the MPPT process lasts a smaller amount of time compared with shading pattern 1, in which the agent had no prior knowledge of the location of the MPP. This leads to the conclusion that the agent was able to use its previous knowledge from the MPP detection in unknown intermediate conditions of solar irradiation of shading patterns 1 and 2. The agent explored only the states that it visited for the first time.

Figs. 5-77 and 5-78 present the results of the simulations of duty cycle and PV array output power versus time of the third implementation of the Q-learning-based MPPT method for shading pattern 8.



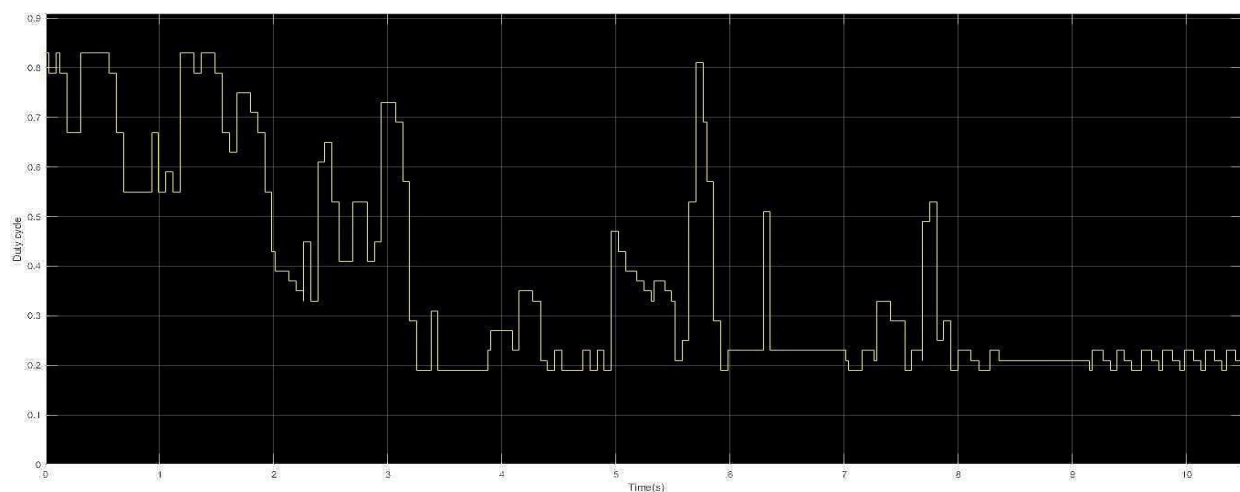
**Fig. 5-77.** Duty cycle versus time for shading pattern 8 of the third implementation of the Q-learning-based MPPT method.



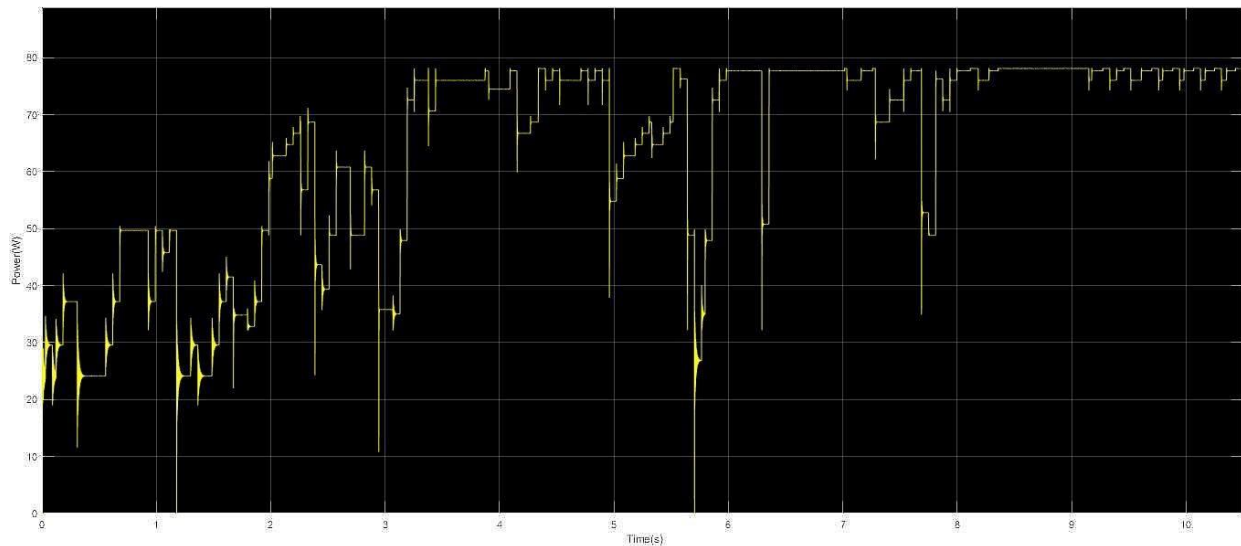
**Fig. 5-78.** PV array output power versus time for shading pattern 8 of the third implementation of the Q-learning-based MPPT method.

It is observed that in the third implementation of the Q-learning-based MPPT method the agent needs more time to detect the location of the GMPP for shading pattern 8 compared with the required time to locate the GMPP in the previous shading pattern. However, the MPPT process lasts a smaller amount of time compared with shading pattern 1, in which the agent had no prior knowledge of the location of the GMPP. This leads to the conclusion that the agent was able to use its previous knowledge from the GMPP detection in unknown intermediate conditions of solar irradiation of shading patterns 1 and 2. The agent explored only the states that it visited for the first time.

Figs. 5-79 and 5-80 present the results of the simulations of duty cycle and PV array output power versus time of the fourth implementation of the Q-learning-based MPPT method for shading pattern 8.



**Fig. 5-79.** Duty cycle versus time for shading pattern 8 of the fourth implementation of the Q-learning-based MPPT method.

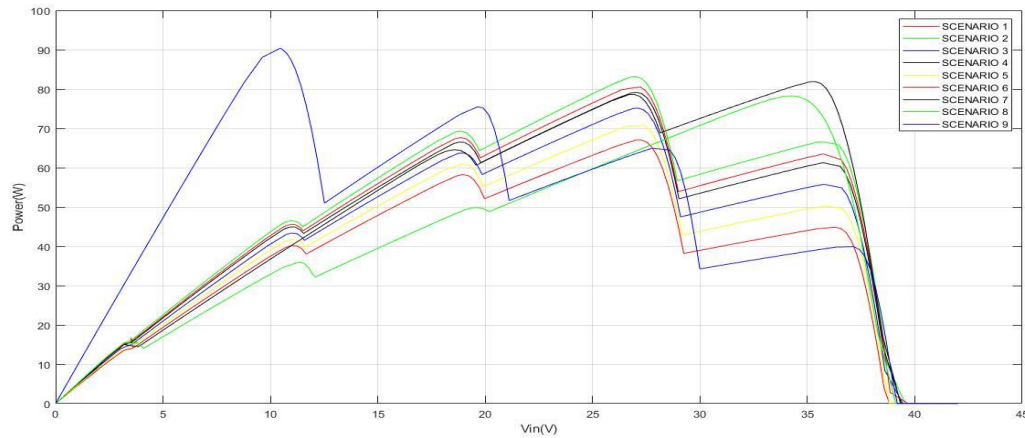


**Fig. 5-80.** PV array output power versus time for shading pattern 8 of the fourth implementation of the Q-learning-based MPPT method.

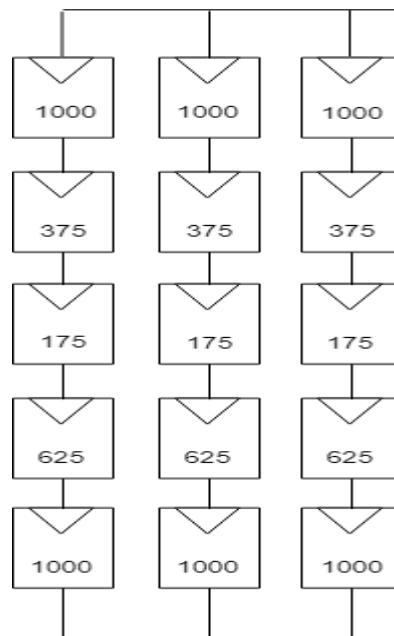
It is observed that in the fourth implementation of the Q-learning-based MPPT method the agent needs more time to detect the location of the GMPP for shading pattern 8 compared with the required time to locate the GMPP in the previous shading pattern. However, the MPPT process lasts a smaller amount of time compared with shading pattern 1, in which the agent had no prior knowledge of the location of the MPP. This leads to the conclusion that the agent was able to use its previous knowledge from the MPP detection in unknown intermediate conditions of solar irradiation of shading patterns 1 and 2. The agent explored only the states that it visited for the first time.

## 5.9 Analysis of the simulation results for shading pattern 9

In Fig. 5-81 the PV array output power-voltage curve is presented and in Fig. 5-82 the distribution of incident solar irradiation over each PV module is presented. The numbers above each PV module represent the solar irradiation intensity, which is measured in  $W/m^2$ . Shading pattern 9 is a non-uniform solar irradiance pattern.

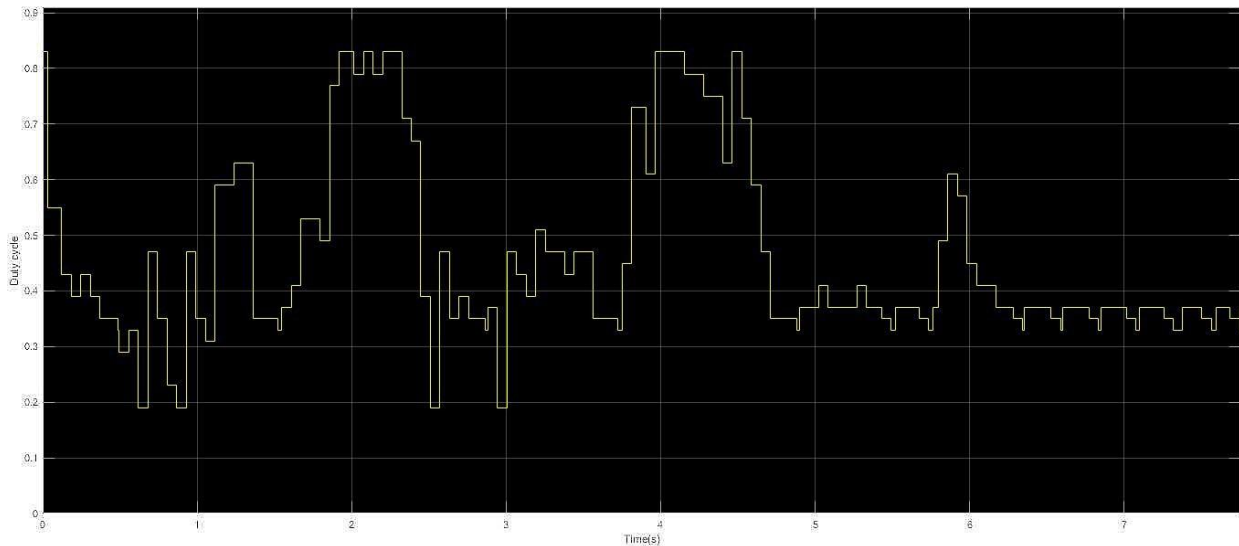


**Fig. 5-81.** The PV array output power-voltage curve for shading pattern 9.

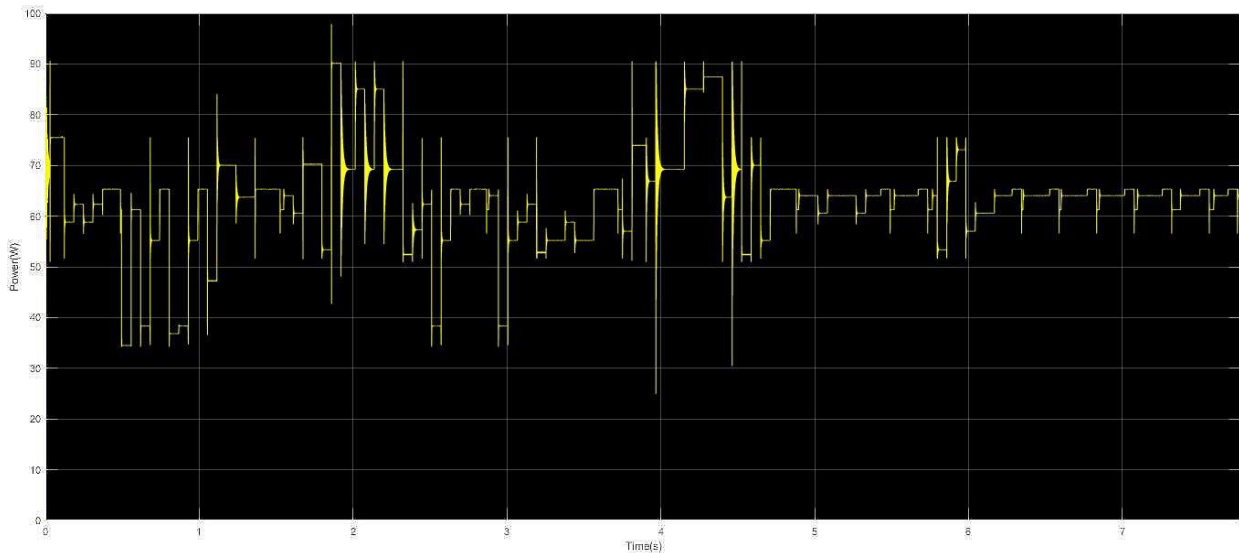


**Fig. 5-82.** The distribution of incident solar irradiation over each PV module for shading pattern 9.

Eight plots were exported, the PV array output power versus time plot and the duty cycle versus time plot in the four aforementioned implementations of the Q-learning-based MPPT method. Figs. 5-83 and 5-84 present the results of the simulations of the duty cycle and the PV array output power versus time for the first implementation of the Q-learning-based MPPT method for shading pattern 9.



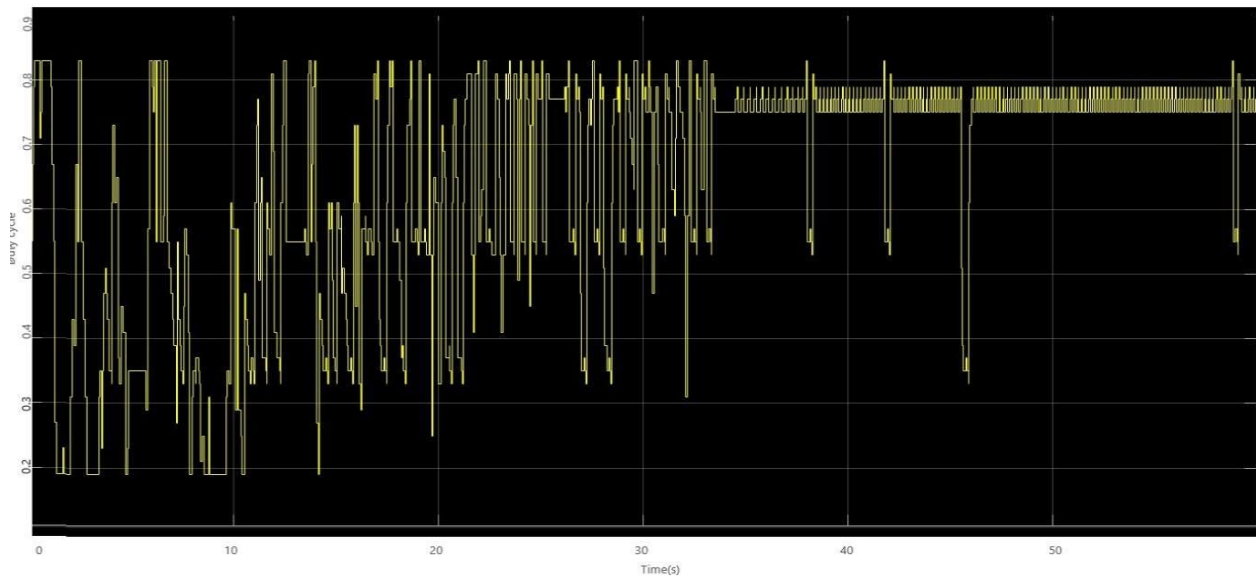
**Fig. 5-83.** Duty cycle versus time for shading pattern 9 of the first implementation of the Q-learning-based MPPT method.



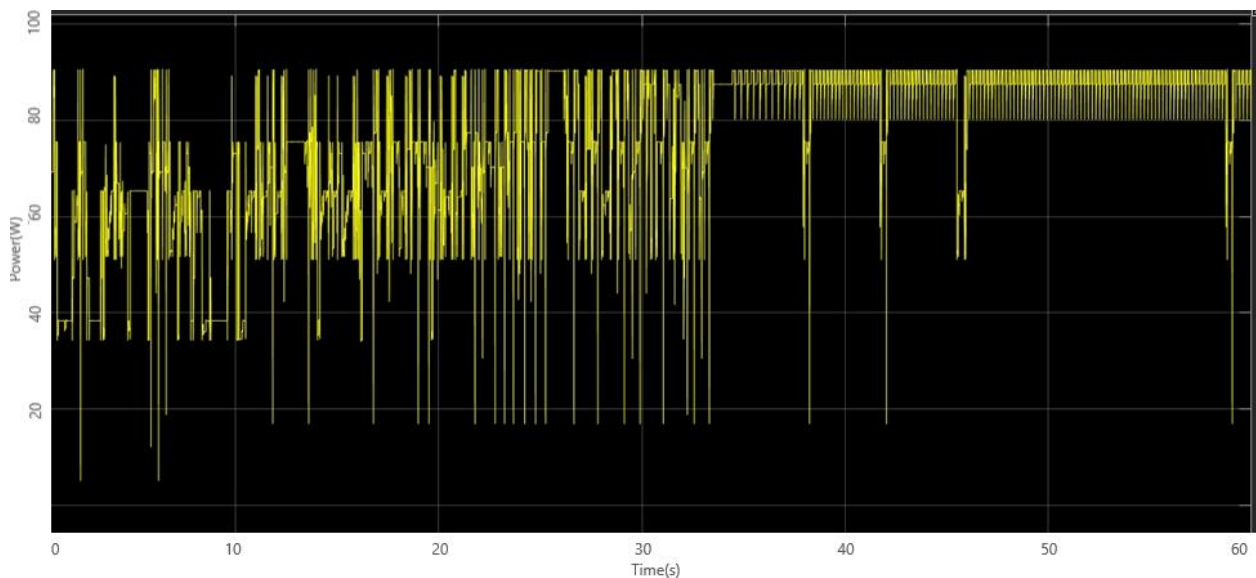
**Fig. 5-84.** PV array output power versus time for shading pattern 9 of the first implementation of the Q-learning-based MPPT method.

It is observed that the agent fails to detect the GMPP. In this shading pattern, the prior knowledge leads the agent to a point far from the MPP. The GMPP of the PV array output power is equal to 90 W and the agent converges to a point that the PV array output power is equal to 65 W.

Figs. 5-85 and 5-86 present the results of the simulations of duty cycle and PV array output power versus time of the second implementation of the Q-learning-based MPPT method for shading pattern 9.



**Fig. 5-85.** Duty cycle versus time for shading pattern 9 of the second implementation of the Q-learning-based MPPT method.

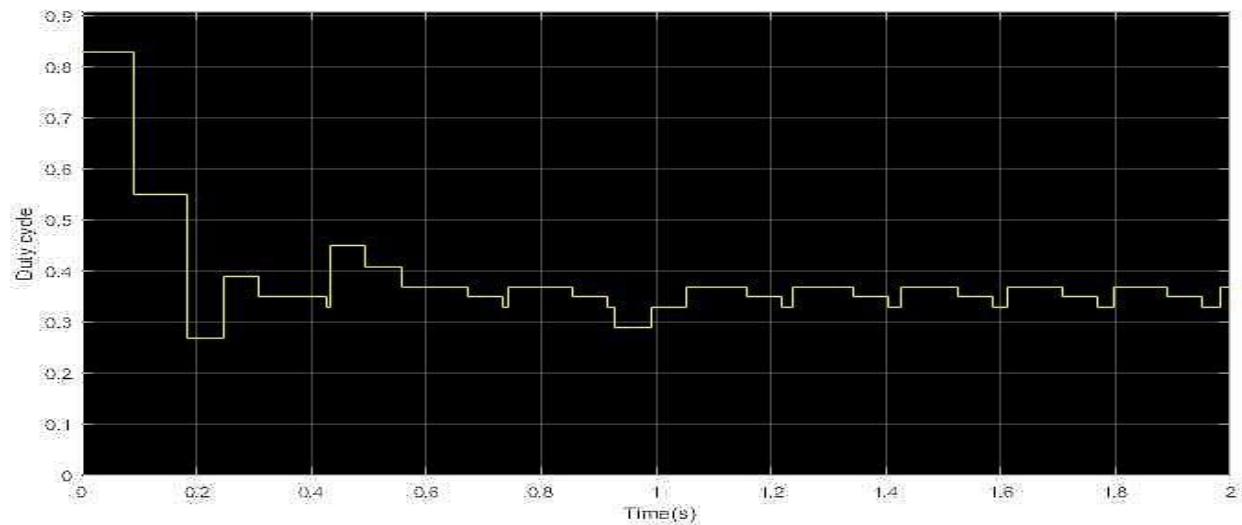


**Fig. 5-86.** PV array output power versus time for shading pattern 9 of the second implementation of the Q-learning-based MPPT method.

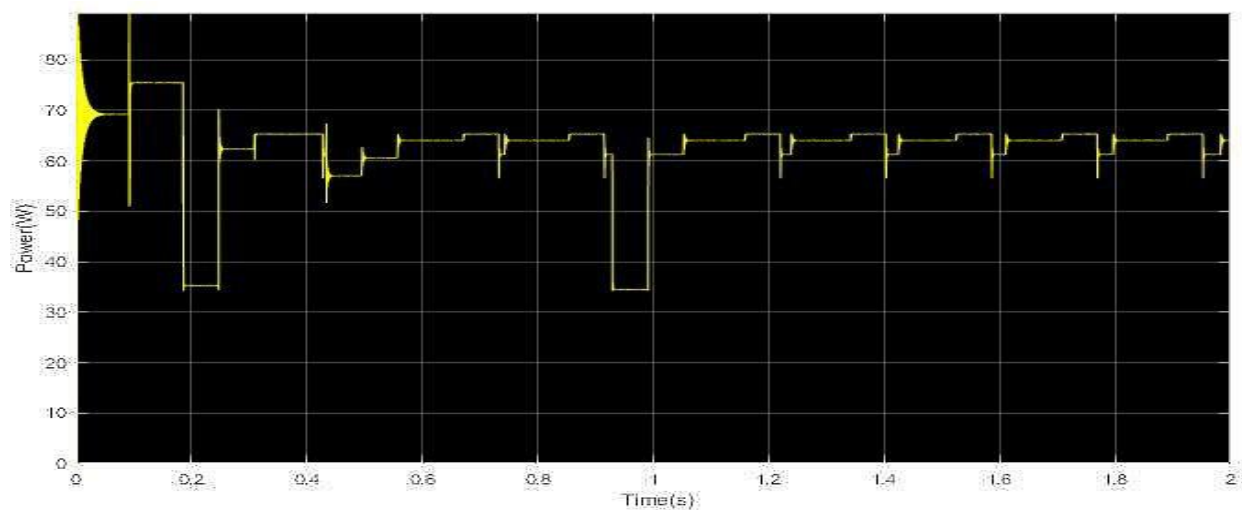
It is observed that the agent succeeds to detect the location of the GMPP in an amount of time, which is equal with the convergence time for shading pattern 1. This is because some states mislead the agent and force it to select an action that would give the agent a positive reward for the specific state of previous shading patterns. However, this action selection gives a negative reward for the specific state of the shading pattern 9. Thus, these states delay the agent during the process of the detection of the location of the GMPP.



Figs. 5-87 and 5-88 present the results of the simulations of duty cycle and PV array output power versus time of the third implementation of the Q-learning-based MPPT method for shading pattern 9.



**Fig. 5-87.** Duty cycle versus time for shading pattern 9 of the third implementation of the Q-learning-based MPPT method.

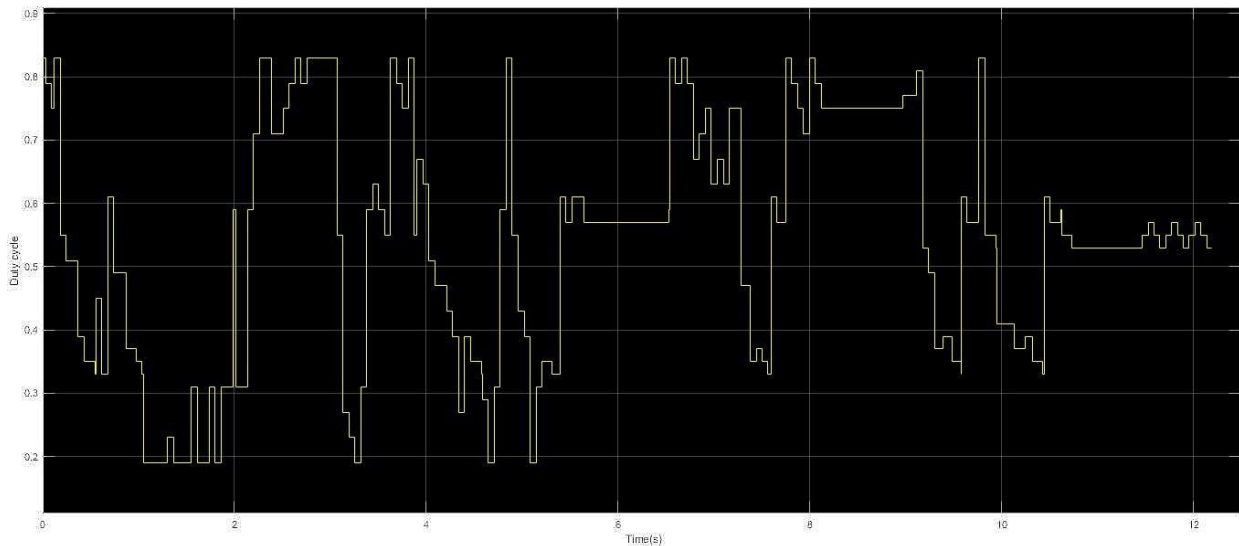


**Fig. 5-88.** PV array output power versus time for shading pattern 9 of the third implementation of the Q-learning-based MPPT method.

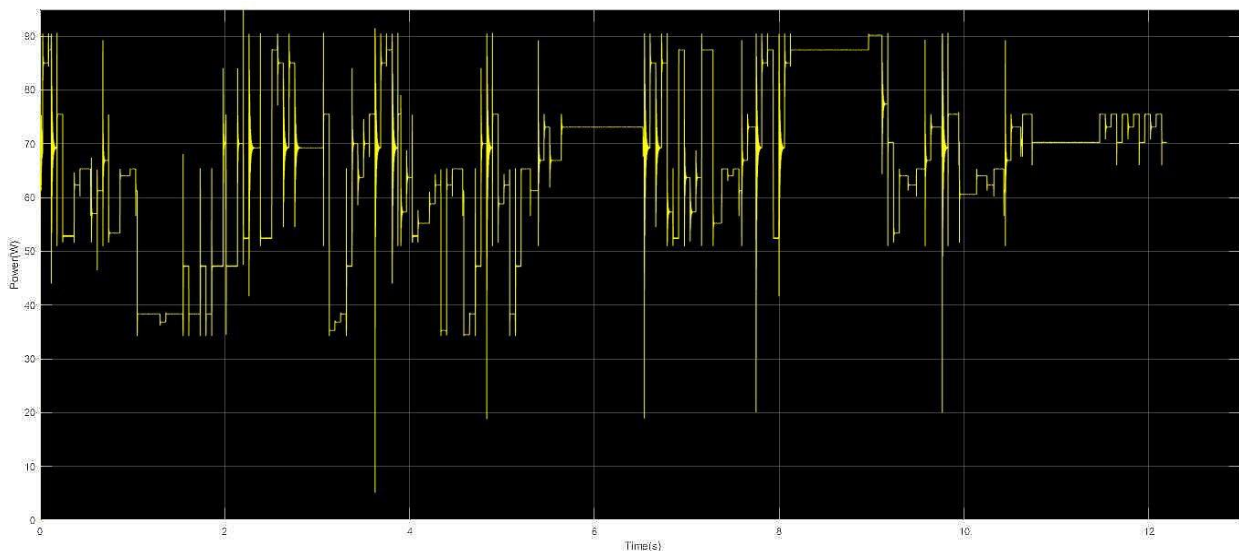
It is observed that the agent fails to detect the MPP. In this shading pattern, the prior knowledge leads the agent to a point far from the MPP. The MPP PV array output power is equal to 90 W and the agent converges to a point that the PV array output power is equal to 65 W.

Figs. 5-89 and 5-90 present the results of the simulations of duty cycle and PV array output power versus time of the fourth implementation of the Q-learning-based MPPT method for shading pattern 9.





**Fig. 5-89.** Duty cycle versus time for shading pattern 9 of the fourth implementation of the Q-learning-based MPPT method.



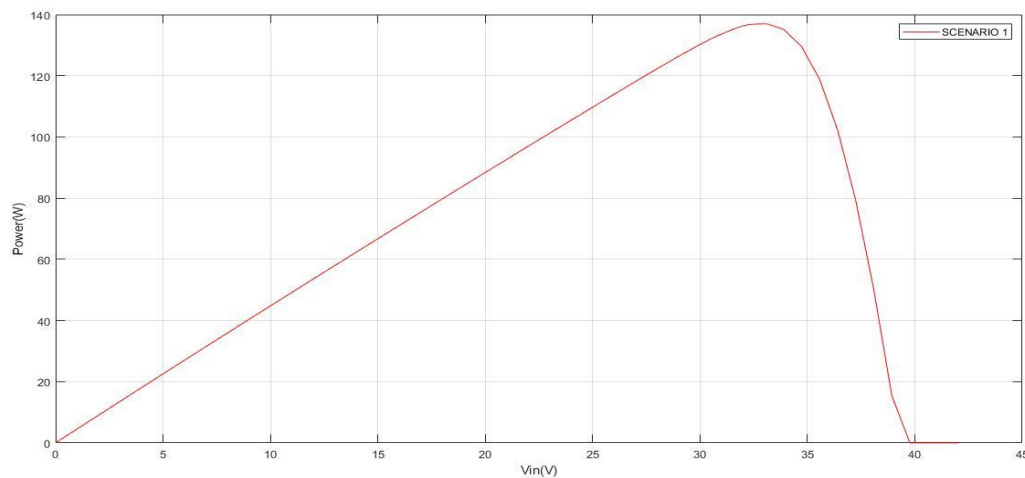
**Fig. 5-90.** PV array output power versus time for shading pattern 9 of the fourth implementation of the Q-learning-based MPPT method.

It is observed that the agent fails to detect the GMPP. In this shading pattern, the prior knowledge leads the agent to a point far from the MPP. The GMPP of the PV array output power is equal to 90 W and the agent converges to a point that the PV array output power is equal to 75 W.

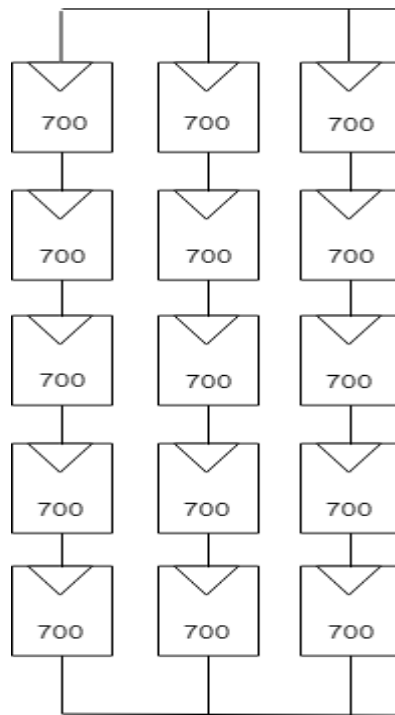
## 5.10 Analysis of the simulation results for shading pattern 10

In Fig. 5-91 the PV array output power-voltage curve is presented and in Fig. 5-92 the distribution of incident solar irradiation over each PV module is presented. The numbers above

each PV module represent the solar irradiation intensity, which is measured in  $W/m^2$ . Shading pattern 10 is a uniform solar irradiance pattern.

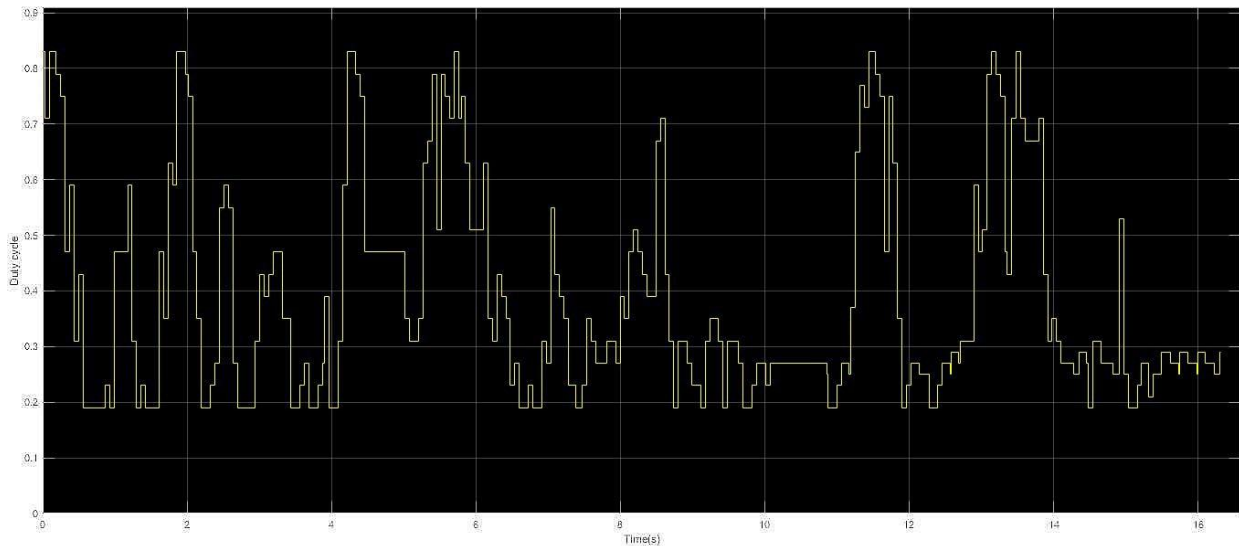


**Fig. 5-91.** The PV array output power-voltage curve for shading pattern 10.

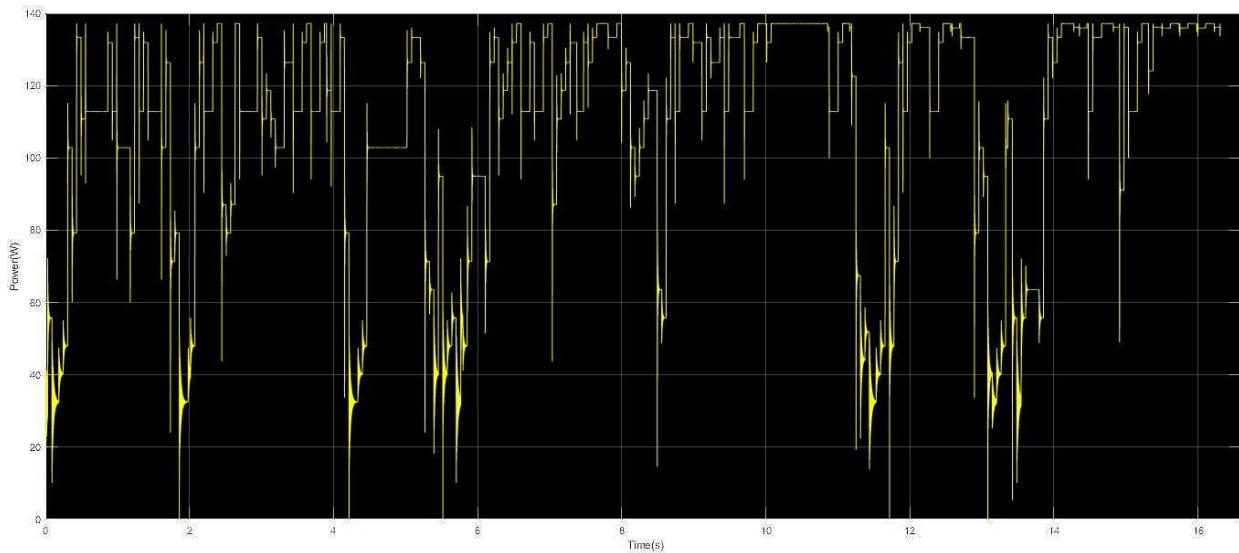


**Fig. 5-92.** The distribution of incident solar irradiation over each PV module for shading pattern 10.

Eight plots were exported, the PV array output power versus time plot and the duty cycle versus time plot in the four aforementioned implementations of the Q-learning-based MPPT method. Figs. 5-93 and 5-94 present the results of the simulations of the duty cycle and the PV array output power versus time for the first implementation of the Q-learning-based MPPT method for shading pattern 10.



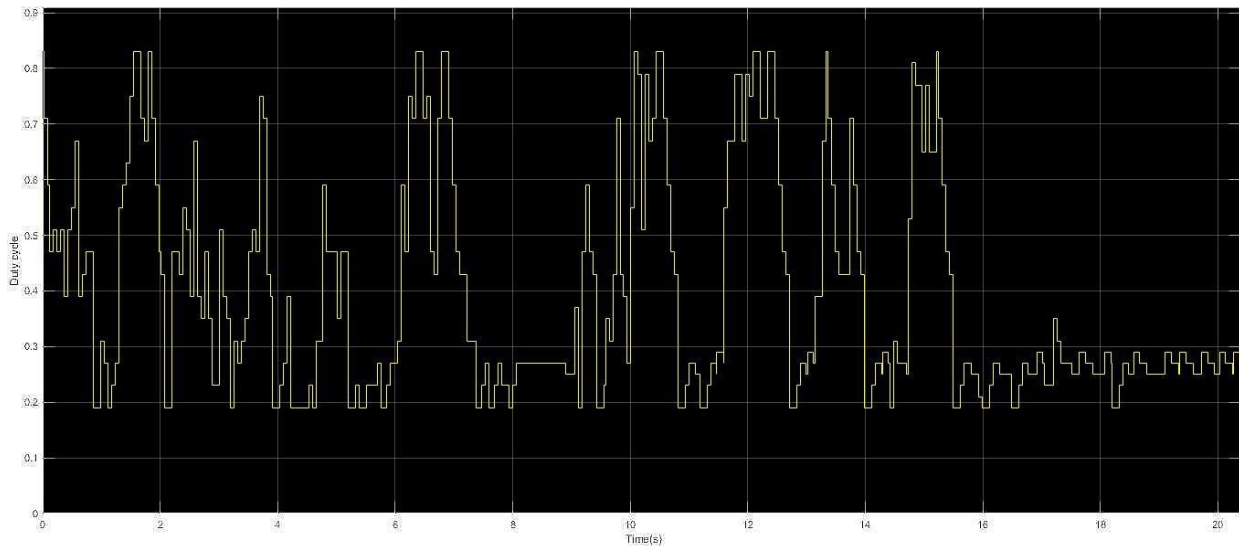
**Fig. 5-93.** Duty cycle versus time for shading pattern 10 of the first implementation of the Q-learning-based MPPT method.



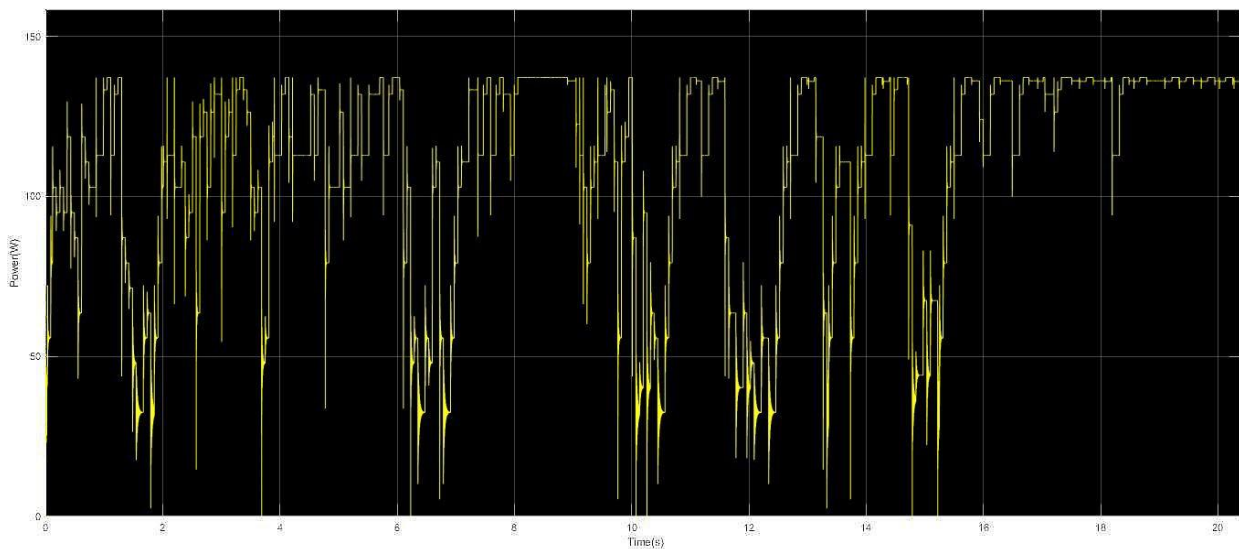
**Fig. 5-94.** PV array output power versus time for shading pattern 10 of the first implementation of the Q-learning-based MPPT method.

It is observed that the agent in the first implementation of the Q-learning-based MPPT method needs approximately 15s to detect GMPP under non-uniform incident solar irradiation conditions, without any prior knowledge of the MPP location.

Figs. 5-95 and 5-96 present the results of the simulations of duty cycle and PV array output power versus time of the second implementation of the Q-learning-based MPPT method for shading pattern 10.



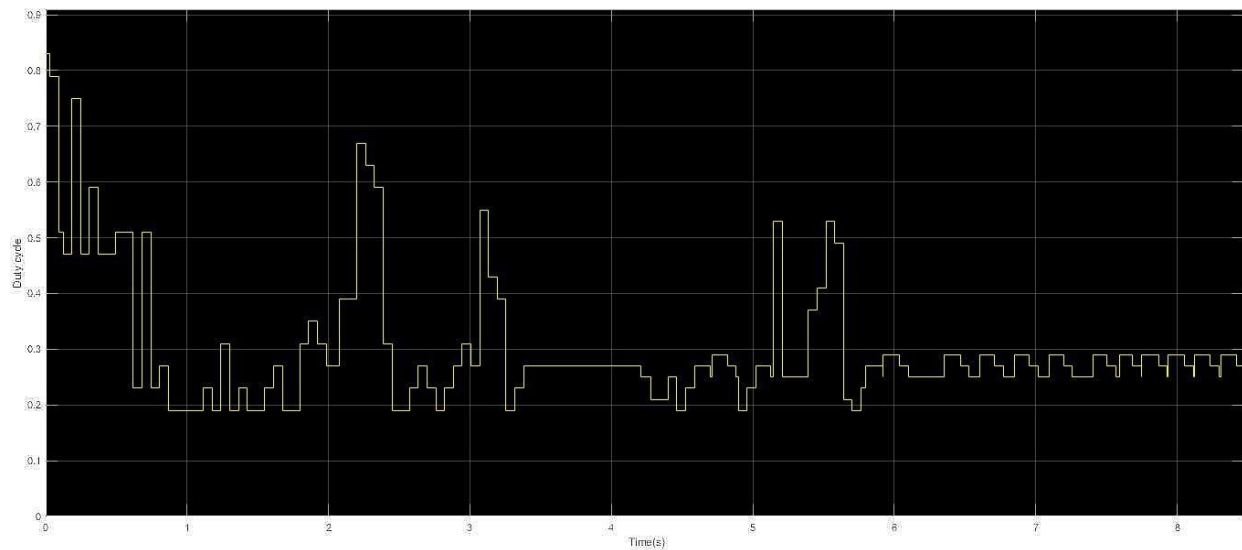
**Fig. 5-95.** Duty cycle versus time for shading pattern 10 of the second implementation of the Q-learning-based MPPT method.



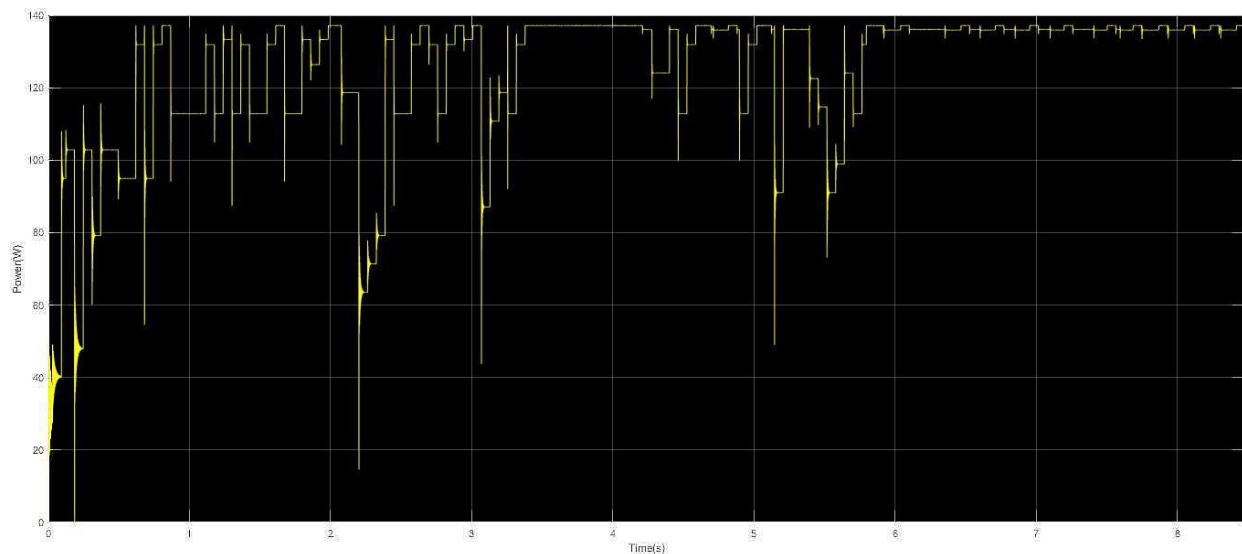
**Fig. 5-96.** PV array output power versus time for shading pattern 10 of the second implementation of the Q-learning-based MPPT method.

It is observed that the agent in the second implementation of the Q-learning-based MPPT method needs approximately 16.5s in order to detect GMPP under non-uniform incident solar irradiation conditions, without any prior knowledge of the MPP position. The longer duration of the MPPT process is due to the fact that the PV array output power is quantized in 30 equal steps, while the PV array output power of the first implementation of the Q-learning-based MPPT method is quantized in 15 equal steps.

Figs. 5-97 and 5-98 present the results of the simulations of duty cycle and PV array output power versus time of the third implementation of the Q-learning-based MPPT method for shading pattern 10.



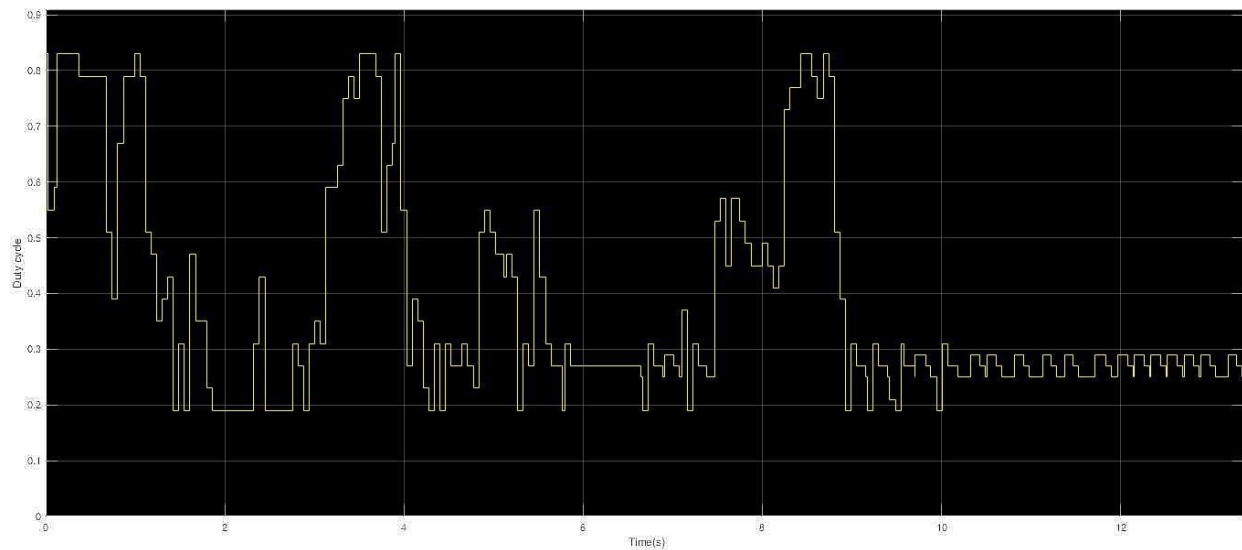
**Fig. 5-97.** Duty cycle versus time for shading pattern 10 of the third implementation of the Q-learning-based MPPT method.



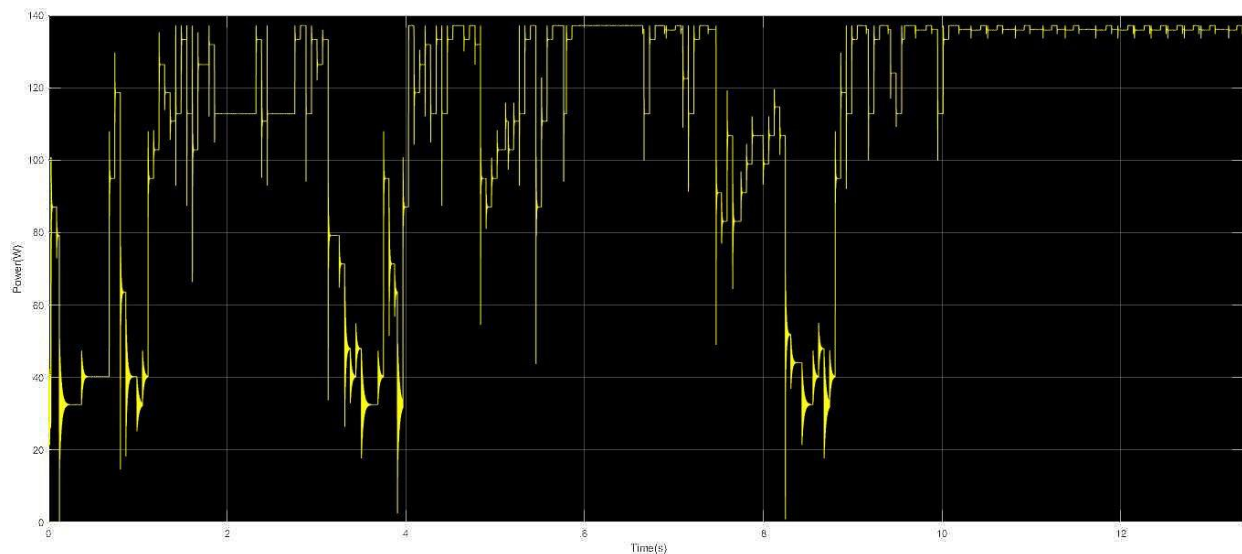
**Fig. 5-98.** PV array output power versus time for shading pattern 10 of the third implementation of the Q-learning-based MPPT method.

It is observed that the agent in the third implementation of the Q-learning-based MPPT method needs approximately 6s to detect GMPP under non-uniform incident solar irradiation conditions, without any prior knowledge of the MPP location.

Figs. 5-99 and 5-100 present the results of the simulations of duty cycle and PV array output power versus time of the fourth implementation of the Q-learning-based MPPT method for shading pattern 10.



**Fig. 5-99.** Duty cycle versus time for shading pattern 10 of the fourth implementation of the Q-learning-based MPPT method.



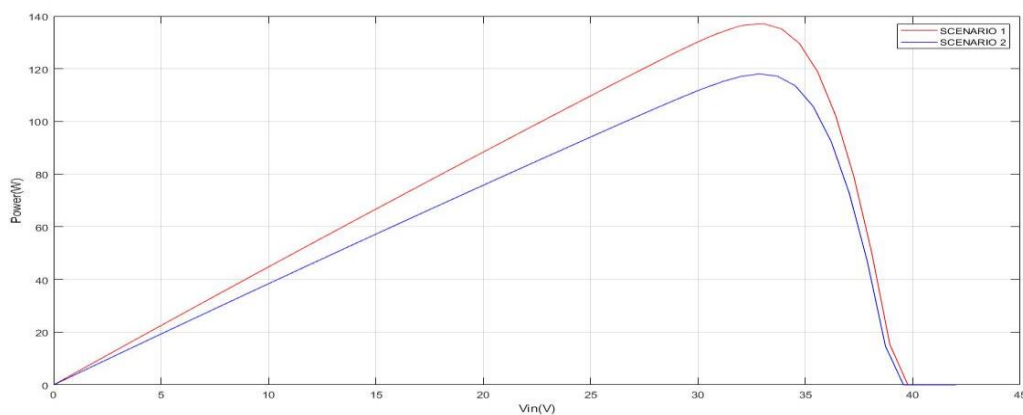
**Fig. 5-100.** PV array output power versus time for shading pattern 10 of the fourth implementation of the Q-learning-based MPPT method.

It is observed that the agent in the fourth implementation of the Q-learning-based MPPT method needs approximately 10s to detect the GMPP under non-uniform incident solar irradiation conditions, without any prior knowledge of the MPP position. The longer duration of the MPPT process is due to the fact that the PV array output power is quantized in 30 equal steps, while the

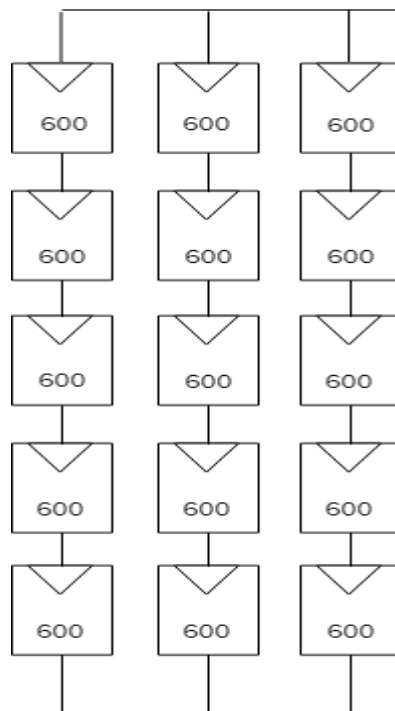
PV array output power of the third implementation of the Q-learning-based MPPT method is quantized in 15 equal steps.

## 5.11 Analysis of the simulation results for shading pattern 11

In Fig. 5-101 the PV array output power-voltage curve is presented and in Fig. 5-102 the distribution of incident solar irradiation over each PV module is presented. The numbers above each PV module represent the solar irradiation intensity, which is measured in  $W/m^2$ . Shading pattern 11 is a uniform solar irradiance pattern.

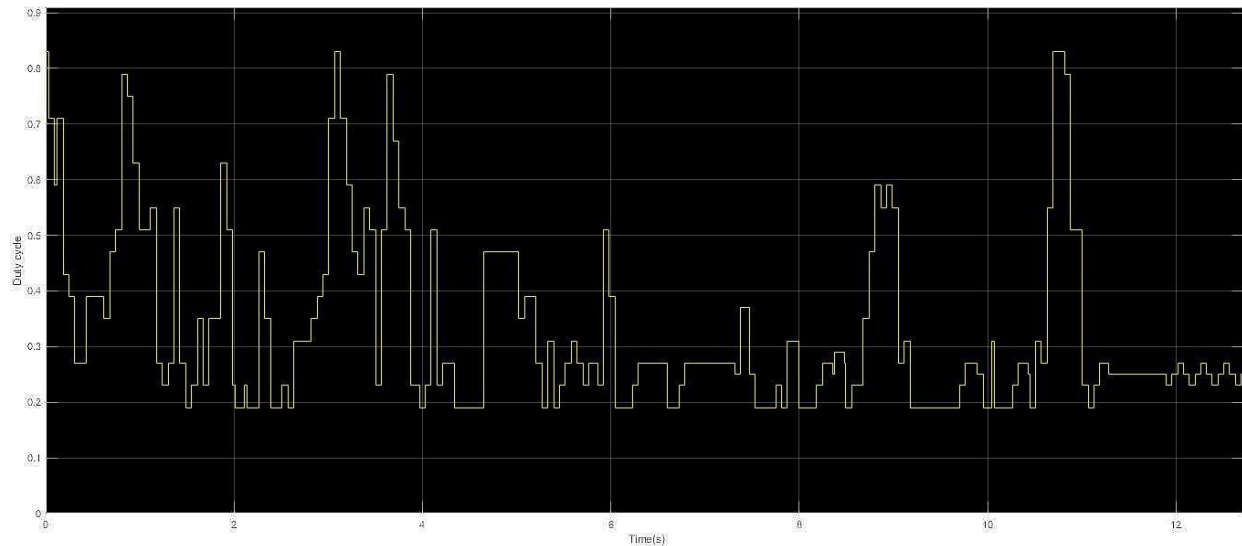


**Fig. 5-101.** The PV array output power-voltage curve for shading pattern 11.

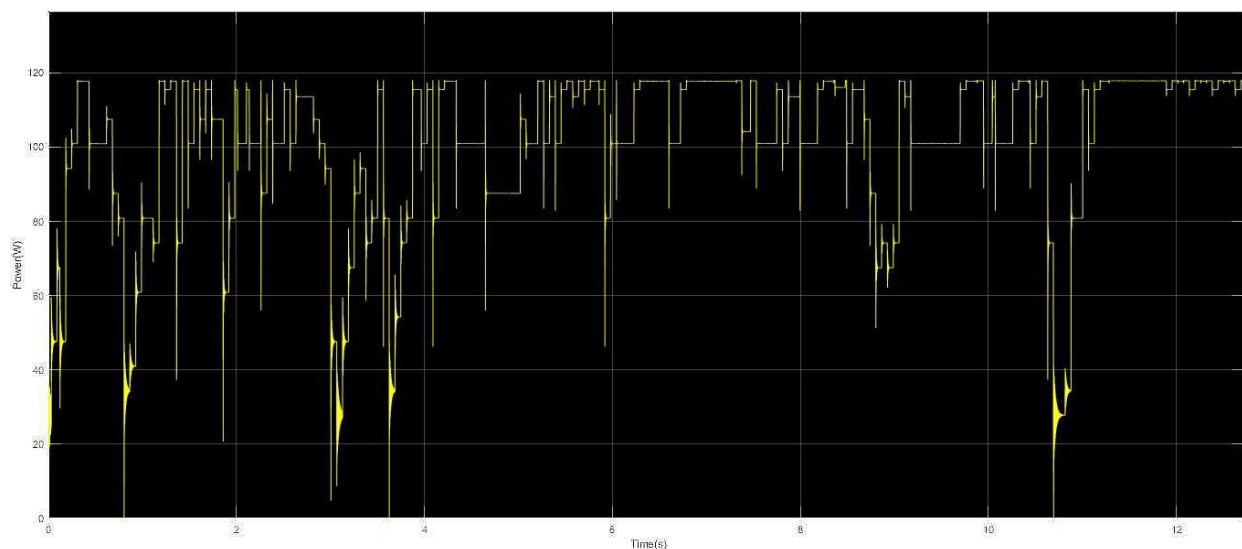


**Fig. 5-102.** The distribution of incident solar irradiation over each PV module for shading pattern 11.

Eight plots were exported, the PV array output power versus time plot and the duty cycle versus time plot in the four aforementioned implementations of the Q-learning-based MPPT method. Figs. 5-103 and 5-104 present the results of the simulations of the duty cycle and the PV array output power versus time for the first implementation of the Q-learning-based MPPT method for shading pattern 11.



**Fig. 5-103.** Duty cycle versus time for shading pattern 11 of the first implementation of the Q-learning-based MPPT method.



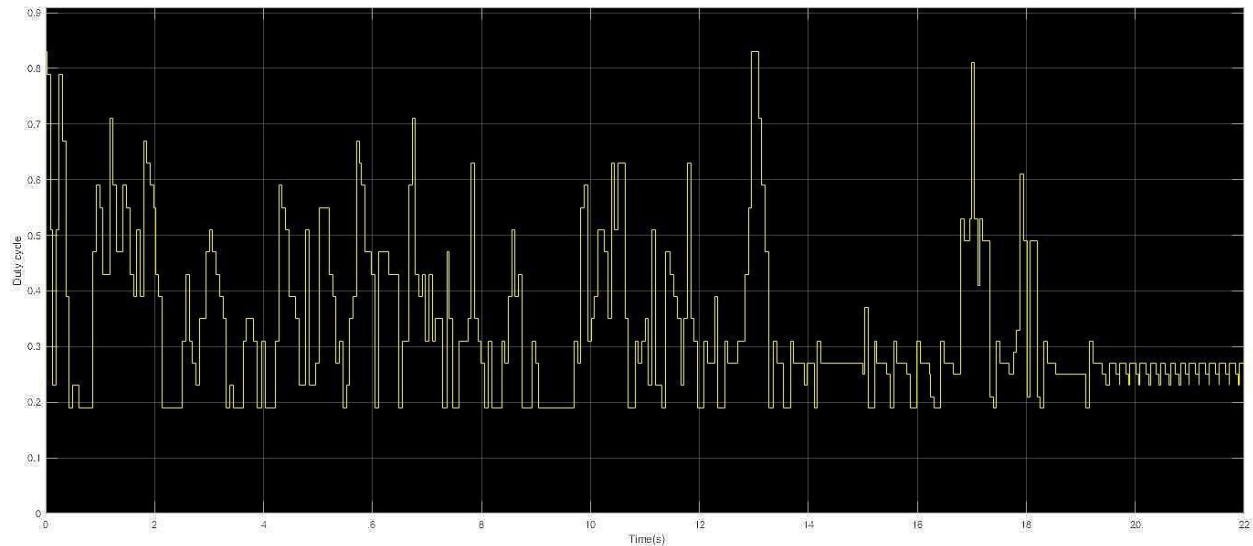
**Fig. 5-104.** PV array output power versus time for shading pattern 11 of the first implementation of the Q-learning-based MPPT method.

It is observed that the agent in the first implementation of the Q-learning-based MPPT method needs approximately 11s in order to detect the GMPP under non-uniform incident solar

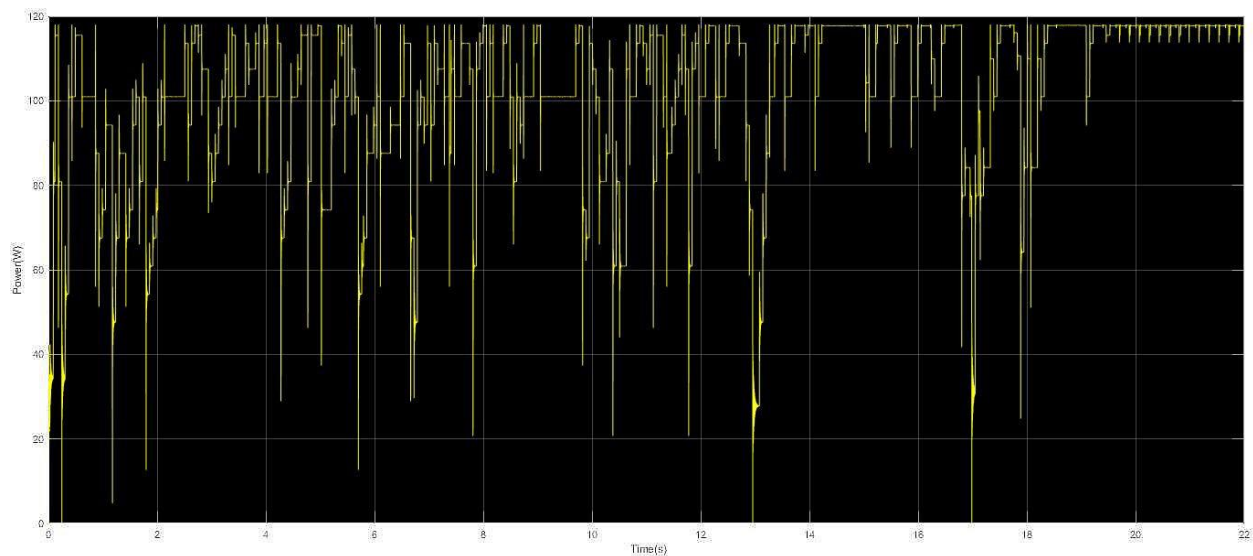


irradiation conditions. The convergence time was reduced, because of the prior knowledge that was acquired by the MPPT process for shading pattern 10.

Figs. 5-105 and 5-106 present the results of the simulations of duty cycle and PV array output power versus time of the second implementation of the Q-learning-based MPPT method for shading pattern 11.



**Fig. 5-105.** Duty cycle versus time for shading pattern 11 of the second implementation of the Q-learning-based MPPT method.

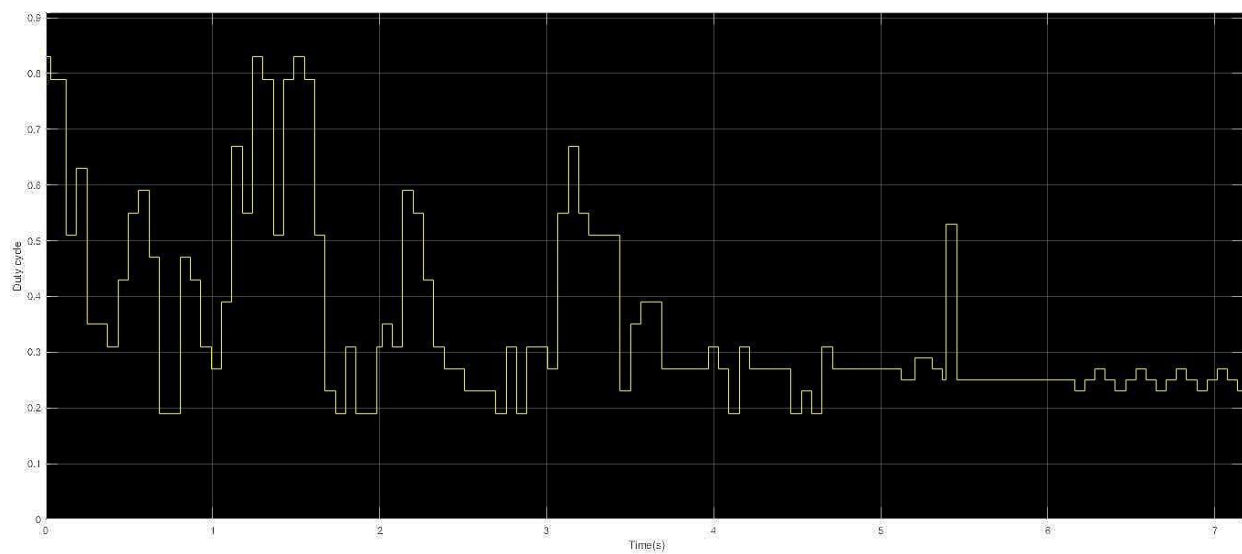


**Fig. 5-106.** PV array output power versus time for shading pattern 11 of the second implementation of the Q-learning-based MPPT method.

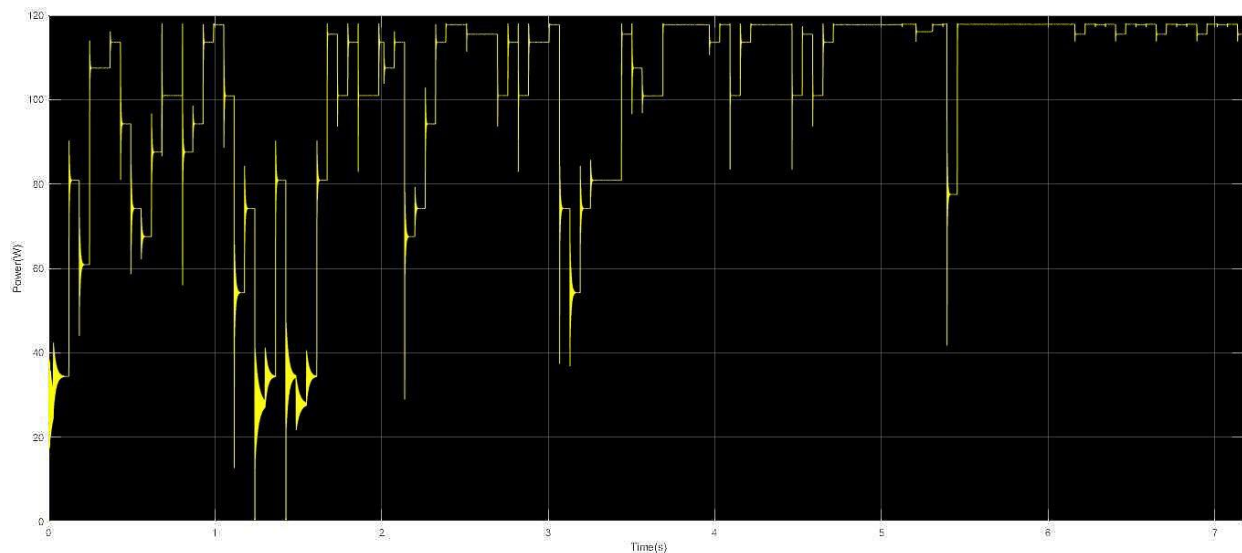
The convergence time of the second implementation of the Q-learning-based MPPT method for shading pattern 11 is approximately equal to the convergence time of the method for shading pattern 10, because the high number of equal quantized steps of the PV array output power, leads

to the conclusion that the agent needs training in more shading patterns in order to reduce the required time for the GMPP detection, compared with the first implementation of the Q-learning-based MPPT method.

Figs. 5-107 and 5-108 present the results of the simulations of duty cycle and PV array output power versus time of the third implementation of the Q-learning-based MPPT method for shading pattern 11.



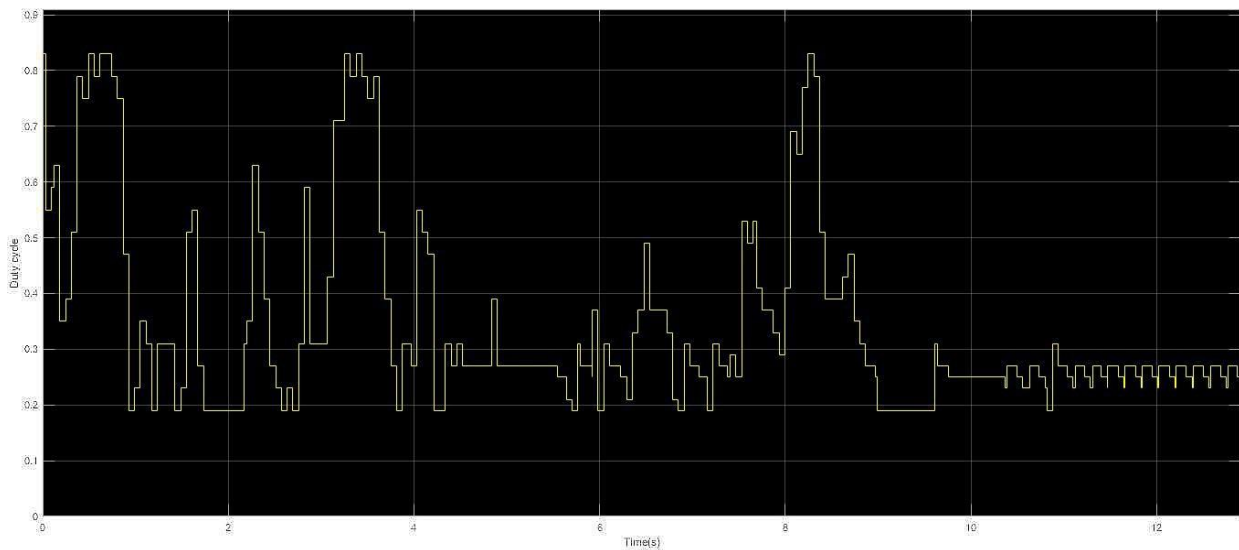
**Fig. 5-107.** Duty cycle versus time for shading pattern 11 of the third implementation of the Q-learning-based MPPT method.



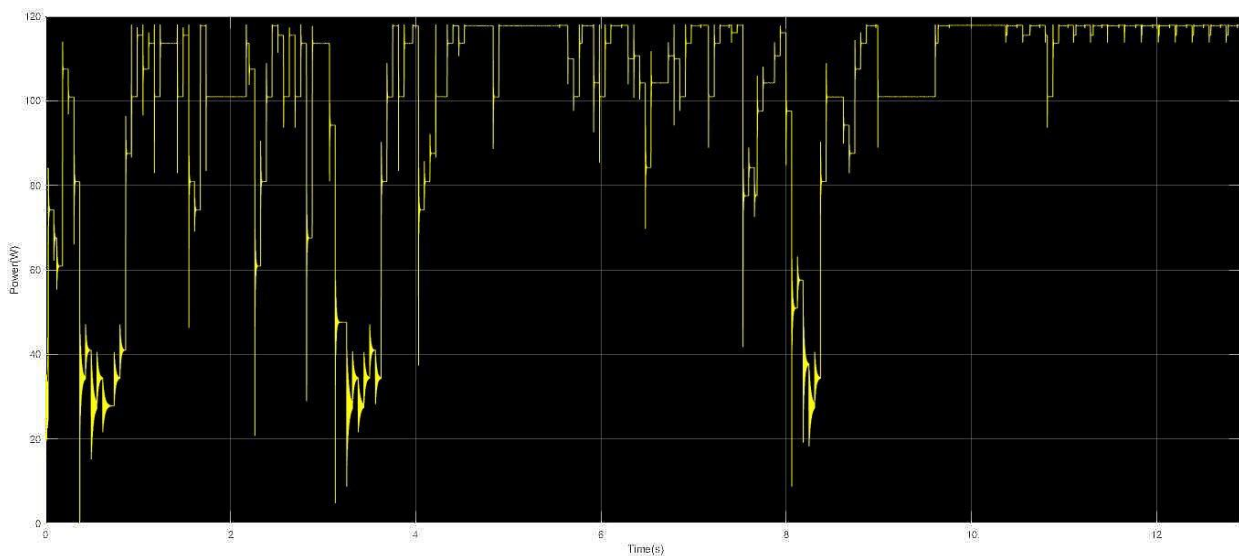
**Fig. 5-108.** PV array output power versus time for shading pattern 11 of the third implementation of the Q-learning-based MPPT method.

It is observed that the convergence time of the third implementation of the Q-learning-based MPPT method for shading pattern 11 is approximately equal to the convergence time for shading pattern 10. It seems that the lack of knowledge of the duty cycle in the previous time-step leads to no reduction of the convergence time, compared with the convergence time for shading pattern 10.

Figs. 5-109 and 5-110 present the results of the simulations of duty cycle and PV array output power versus time of the fourth implementation of the Q-learning-based MPPT method for shading pattern 11.



**Fig. 5-109.** Duty cycle versus time for shading pattern 11 of the fourth implementation of the Q-learning-based MPPT method.

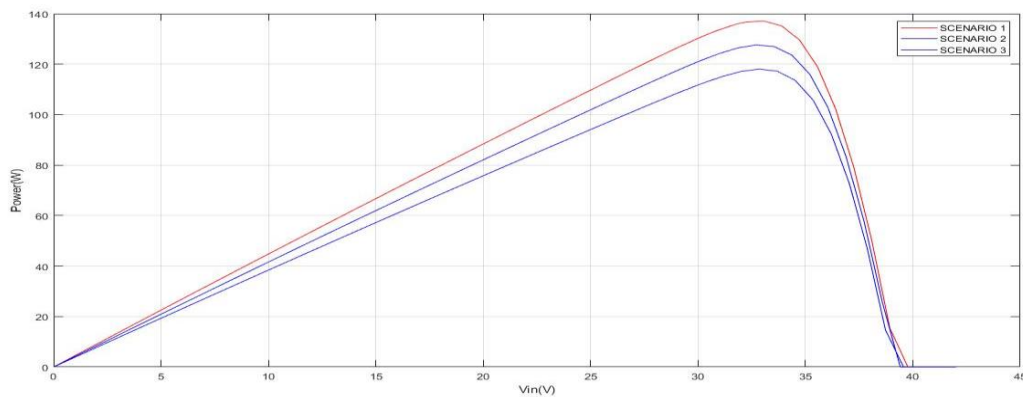


**Fig. 5-110.** PV array output power versus time for shading pattern 11 of the fourth implementation of the Q-learning-based MPPT method.

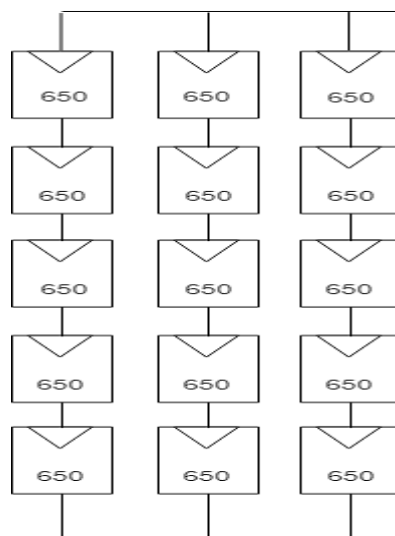
The convergence time of the fourth implementation of the Q-learning-based MPPT method for shading pattern 11 is approximately equal to the convergence time of the method for shading pattern 10, because the high number of equal quantized steps of the PV array output power, leads to the conclusion that the agent needs training in more shading patterns in order to reduce the required time for the MPP detection, compared with the third implementation of the Q-learning-based MPPT method.

## 5.12 Analysis of the simulation results for shading pattern 12

In Fig. 5-111 the PV array output power-voltage curve is presented and in Fig. 5-112 the distribution of incident solar irradiation over each PV module is presented. The numbers above each PV module represent the solar irradiation intensity, which is measured in  $W/m^2$ . Shading pattern 12 is a uniform solar irradiance pattern.

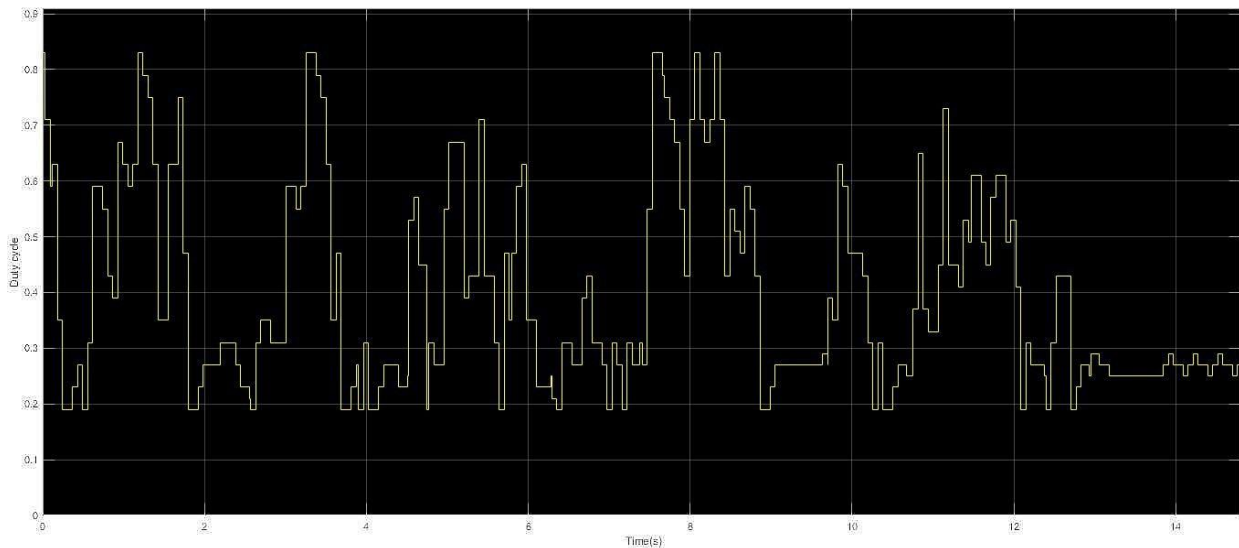


**Fig. 5-111.** The PV array output power-voltage curve for shading pattern 12.

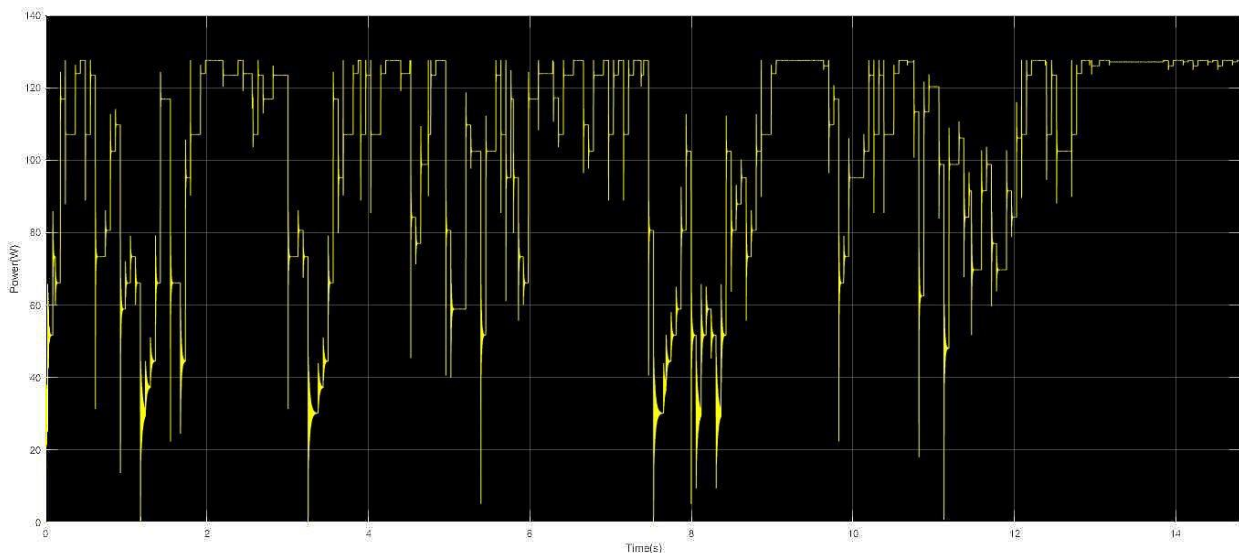


**Fig. 5-112.** The distribution of incident solar irradiation over each PV module for shading pattern 12.

Eight plots were exported, the PV array output power versus time plot and the duty cycle versus time plot in the four aforementioned implementations of the Q-learning-based MPPT method. Figs. 5-113 and 5-114 present the results of the simulations of the duty cycle and the PV array output power versus time for the first implementation of the Q-learning-based MPPT method for shading pattern 12.



**Fig. 5-113.** Duty cycle versus time for shading pattern 12 of the first implementation of the Q-learning-based MPPT method.

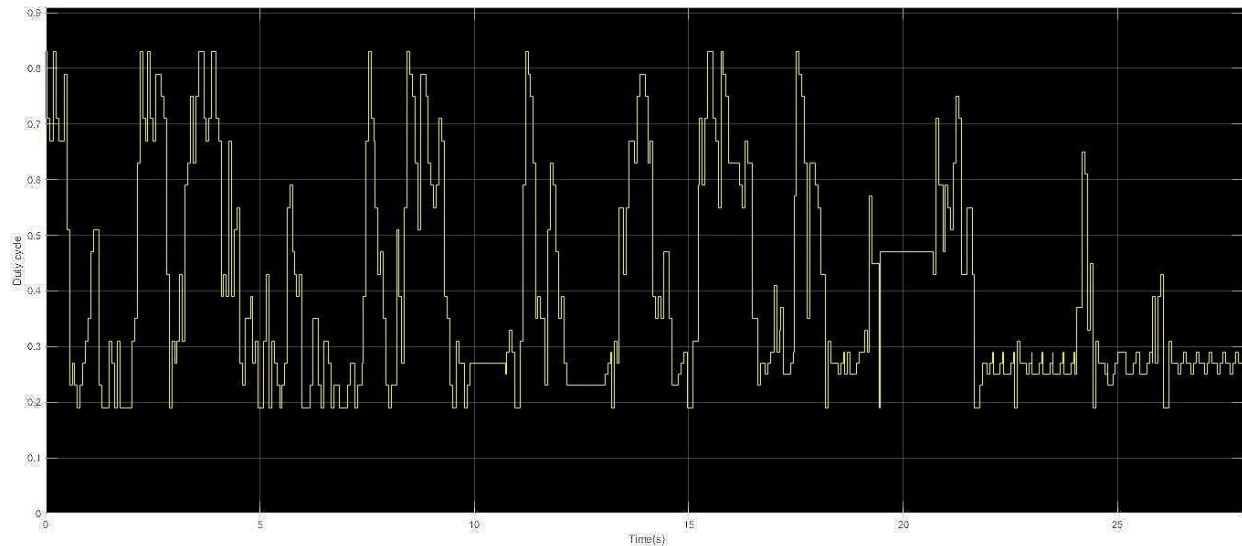


**Fig. 5-114.** PV array output power versus time for shading pattern 12 of the first implementation of the Q-learning-based MPPT method.

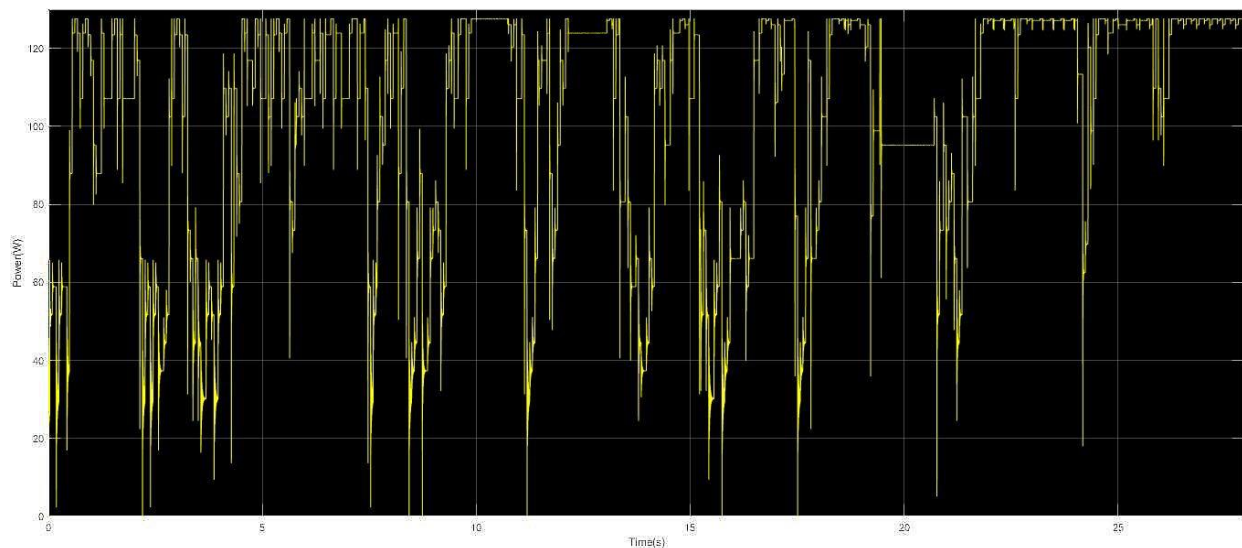
The convergence time of the first implementation of the Q-learning-based MPPT method for shading pattern 12 is approximately equal to the convergence time of the method for shading

pattern 11. This is because the agent needs to be trained in more shading patterns in order to reduce the time required for the GMPP detection.

Figs. 5-115 and 5-116 present the results of the simulations of duty cycle and PV array output power versus time of the second implementation of the Q-learning-based MPPT method for shading pattern 12.



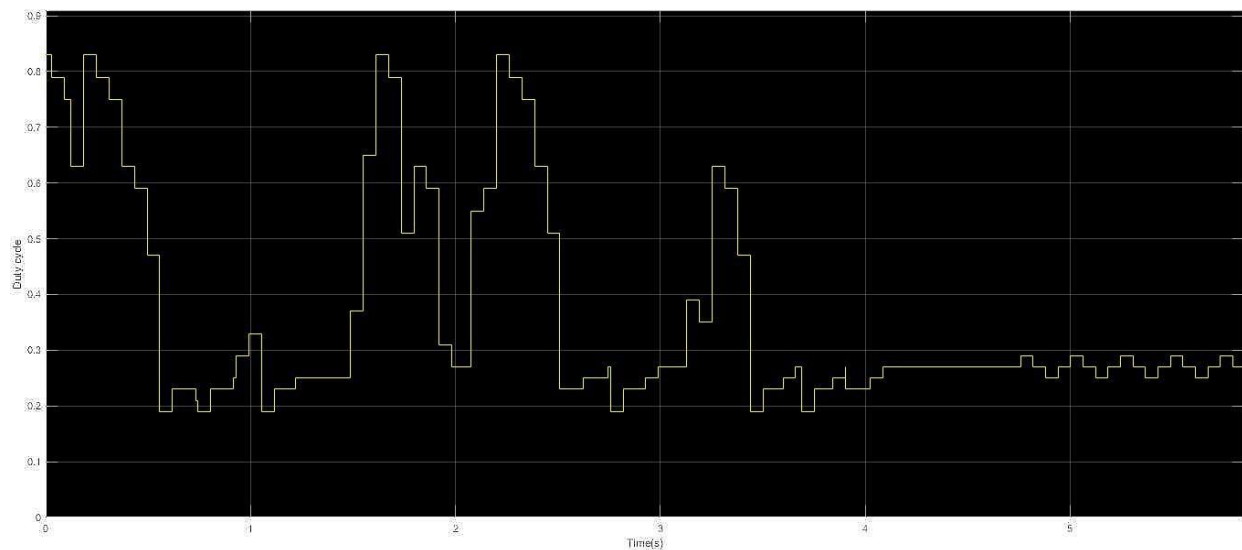
**Fig. 5-115.** Duty cycle versus time for shading pattern 12 of the second implementation of the Q-learning-based MPPT method.



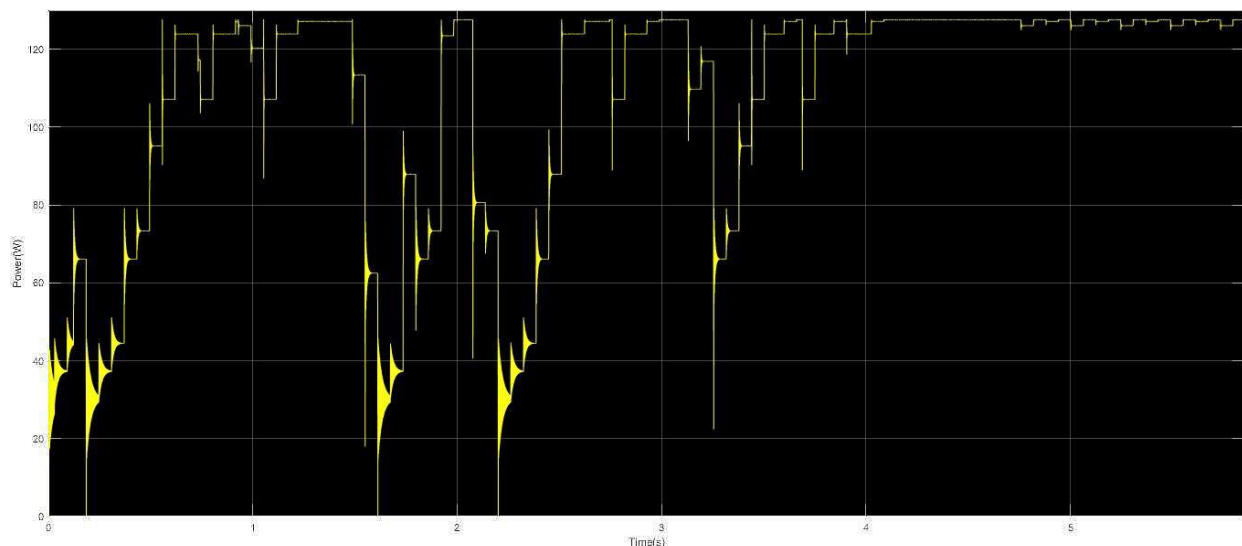
**Fig. 5-116.** PV array output power versus time for shading pattern 12 of the second implementation of the Q-learning-based MPPT method.

The convergence time of the second implementation of the Q-learning-based MPPT method for shading pattern 12 is approximately equal to the convergence time of the method for shading pattern 11. This is because the agent needs to be trained in more shading patterns in order to reduce the time required for the GMPP detection.

Figs. 5-117 and 5-118 present the results of the simulations of duty cycle and PV array output power versus time of the third implementation of the Q-learning-based MPPT method for shading pattern 12.



**Fig. 5-117.** Duty cycle versus time for shading pattern 12 of the third implementation of the Q-learning-based MPPT method.

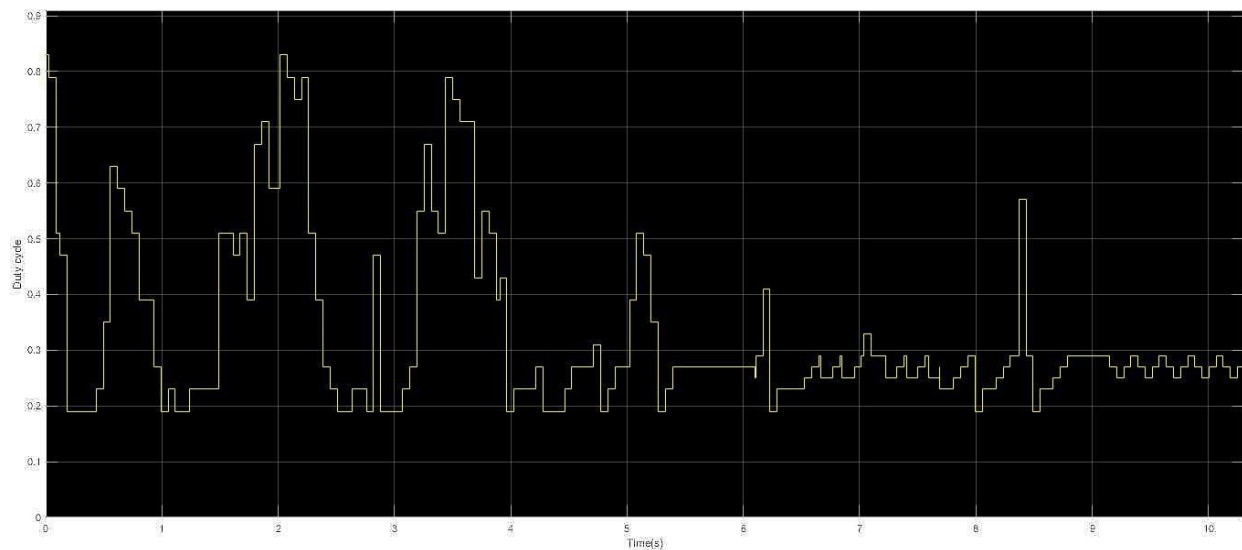


**Fig. 5-118.** PV array output power versus time for shading pattern 12 of the third implementation of the Q-learning-based MPPT method.

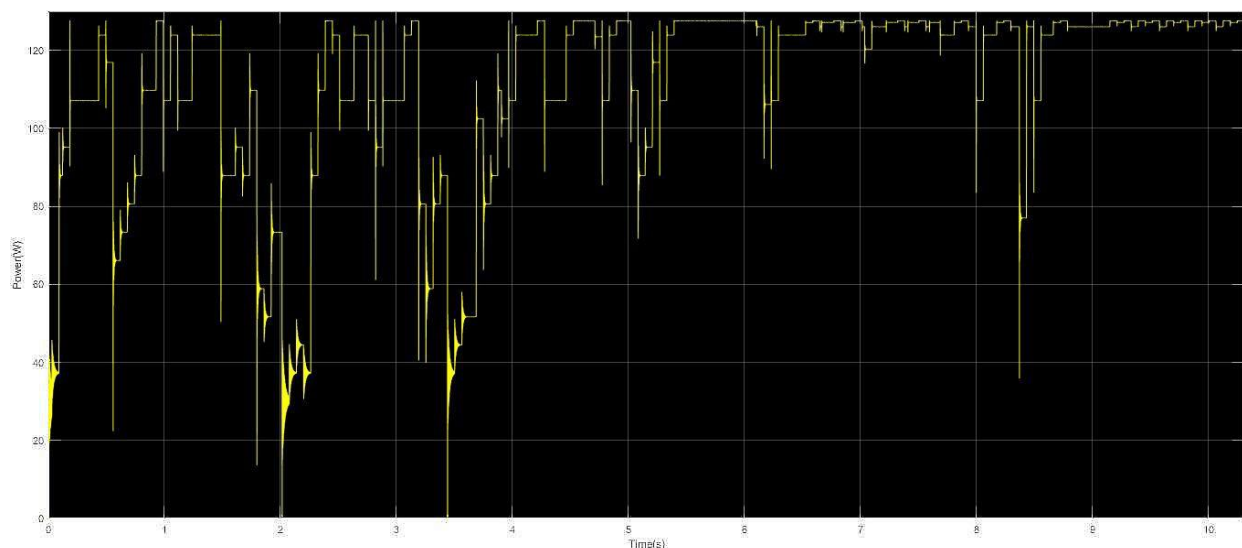


The convergence time of the third implementation of the Q-learning-based MPPT method for shading pattern 12 is approximately equal to the convergence time of the method for shading pattern 11. This is because the agent needs to be trained in more shading patterns in order to reduce the time required for the GMPP detection.

Figs. 5-119 and 5-120 present the results of the simulations of duty cycle and PV array output power versus time of the fourth implementation of the Q-learning-based MPPT method for shading pattern 12.



**Fig. 5-119.** Duty cycle versus time for shading pattern 12 of the fourth implementation of the Q-learning-based MPPT method.



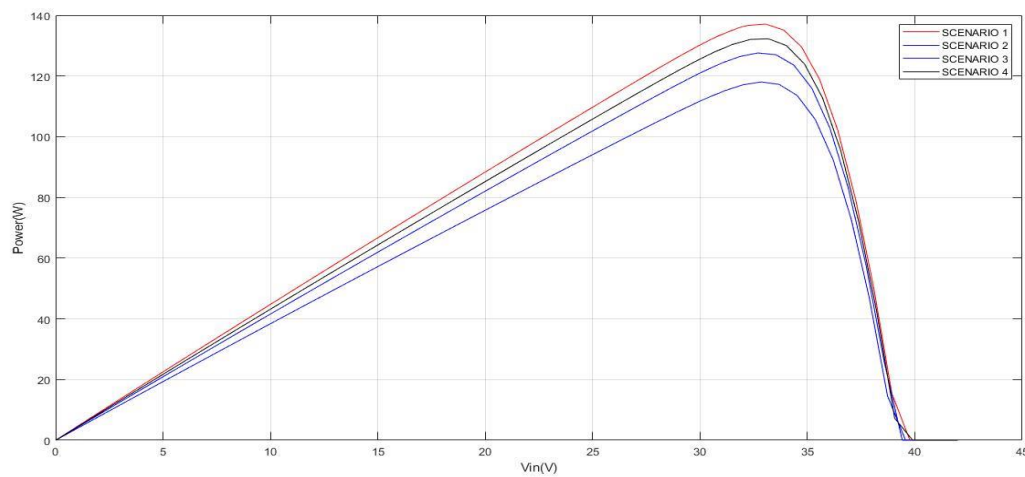
**Fig. 5-120.** PV array output power versus time for shading pattern 12 of the fourth implementation of the Q-learning-based MPPT method.



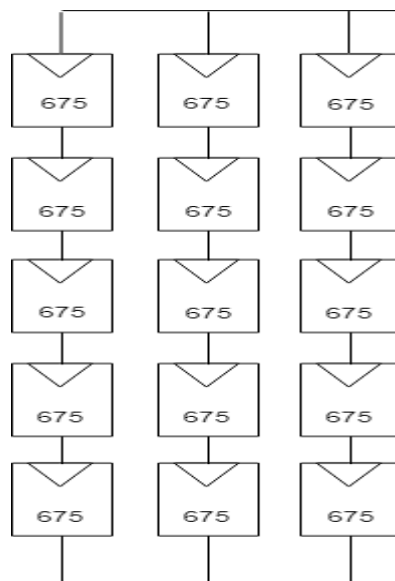
The convergence time of the fourth implementation of the Q-learning-based MPPT method for shading pattern 12 is approximately equal to the convergence time of the method for shading pattern 11. This is because the agent needs to be trained in more shading patterns in order to reduce the time required for the GMPP detection.

### 5.13 Analysis of the simulation results for shading pattern 13

In Fig. 5-121 the PV array output power-voltage curve is presented and in Fig. 5-122 the distribution of incident solar irradiation over each PV module is presented. The numbers above each PV module represent the solar irradiation intensity, which is measured in  $W/m^2$ . Shading pattern 13 is a uniform solar irradiance pattern.

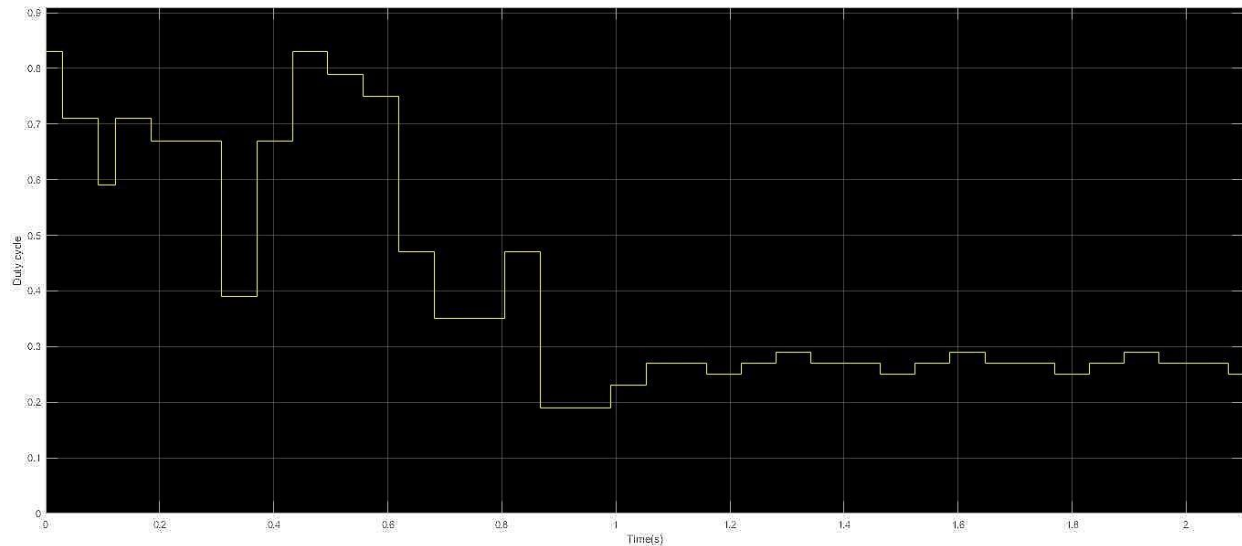


**Fig. 5-121.** The PV array output power-voltage curve for shading pattern 13.

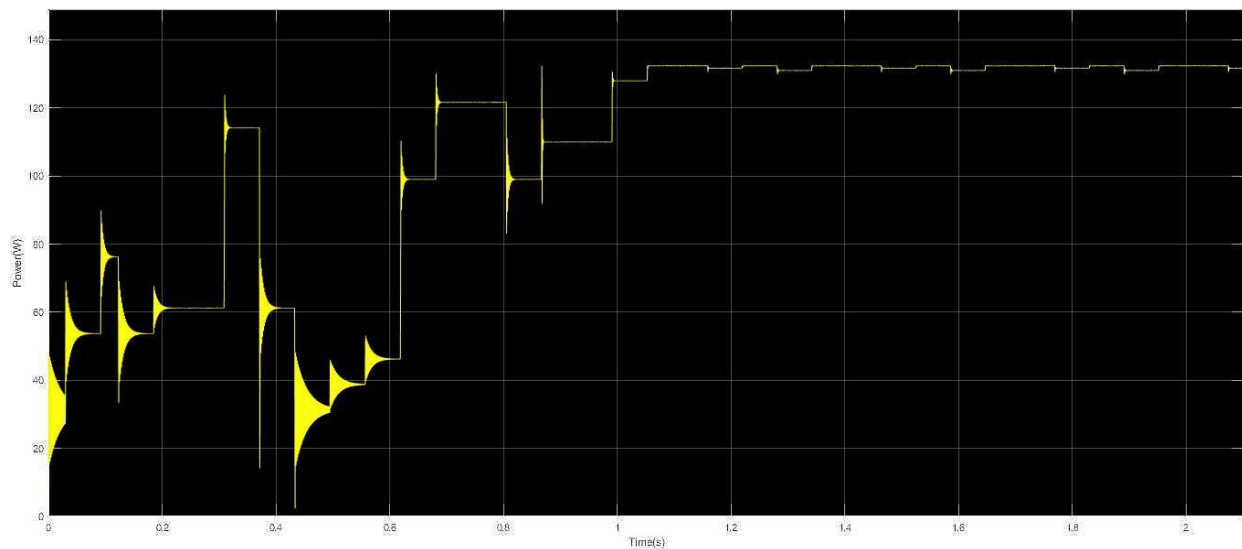


**Fig. 5-122.** The distribution of incident solar irradiation over each PV module for shading pattern 13.

Eight plots were exported, the PV array output power versus time plot and the duty cycle versus time plot in the four aforementioned implementations of the Q-learning-based MPPT method. Figs. 5-123 and 5-124 present the results of the simulations of the duty cycle and the PV array output power versus time for the first implementation of the Q-learning-based MPPT method for shading pattern 13.



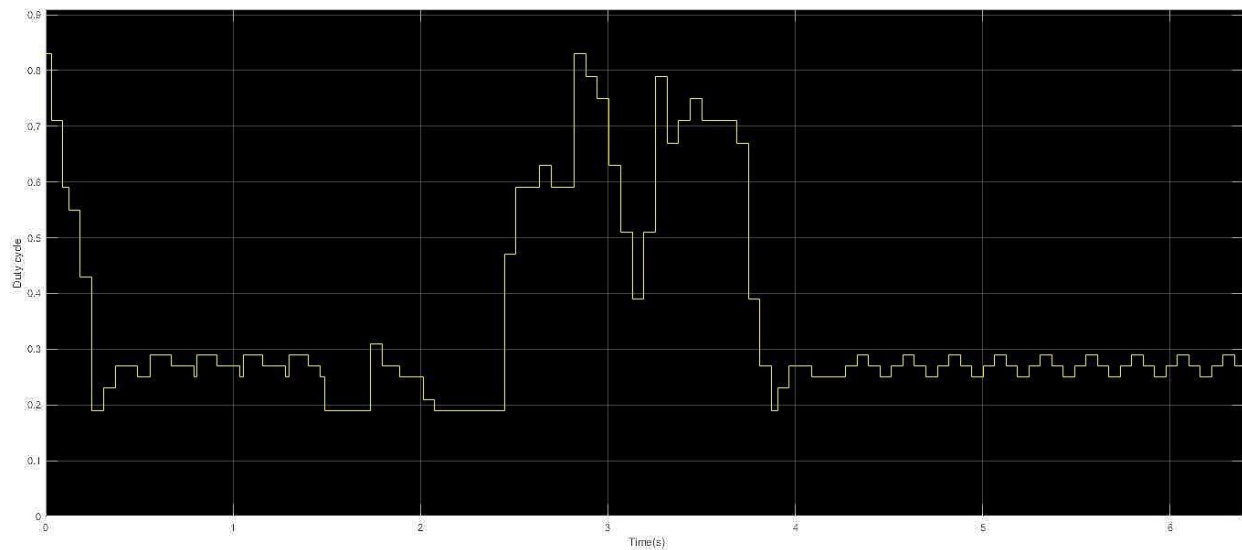
**Fig. 5-123.** Duty cycle versus time for shading pattern 13 of the first implementation of the Q-learning-based MPPT method.



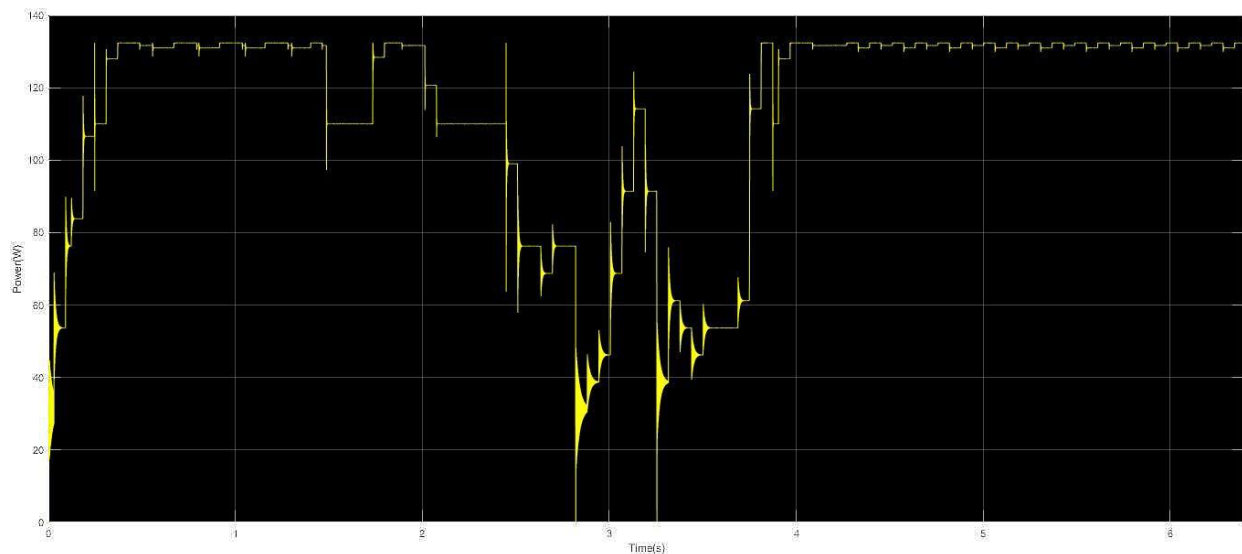
**Fig. 5-124.** PV array output power versus time for shading pattern 13 of the first implementation of the Q-learning-based MPPT method.

It is observed that the convergence time has been reduced. This is due to the fact that the agent has acquired enough knowledge from shading patterns 10, 11 and 12 in order to detect the location the MPP in a few time-steps for shading pattern 13.

Figs. 5-125 and 5-126 present the results of the simulations of duty cycle and PV array output power versus time of the second implementation of the Q-learning-based MPPT method for shading pattern 13.



**Fig. 5-125.** Duty cycle versus time for shading pattern 13 of the second implementation of the Q-learning-based MPPT method.

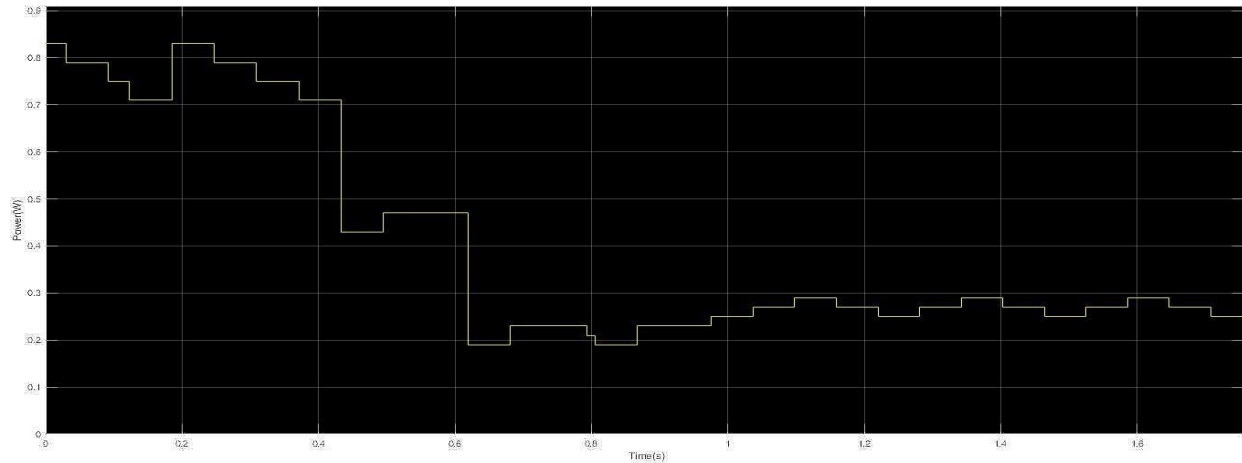


**Fig. 5-126.** PV array output power versus time for shading pattern 13 of the second implementation of the Q-learning-based MPPT method.

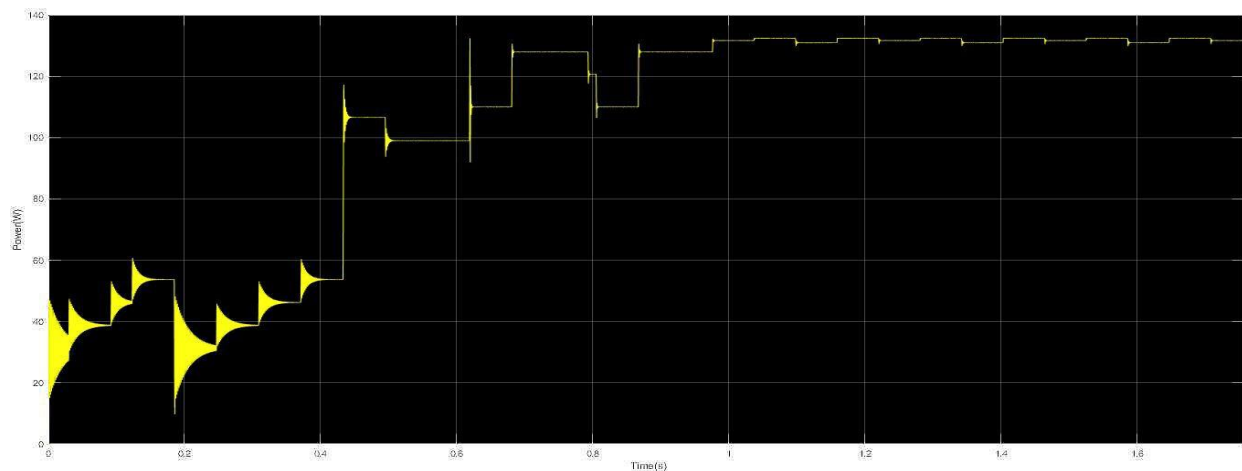
It is observed that the acquired knowledge has helped that agent to reduce the convergence time for shading pattern 13, compared with the convergence time for shading patterns 10, 11 and 12 of the second implementation of the Q-learning-based method. Nevertheless, the agent needs to learn how to detect the location of the GMPP for more shading patterns in order to reduce further the convergence time. This is due to the fact that in the second implementation of the Q-learning

based MPPT method, the PV array output power is quantized in 30 equal steps, while in the first implementation the PV array output power is quantized in 15 equal steps.

Figs. 5-127 and 5-128 present the results of the simulations of duty cycle and PV array output power versus time of the third implementation of the Q-learning-based MPPT method for shading pattern 13.



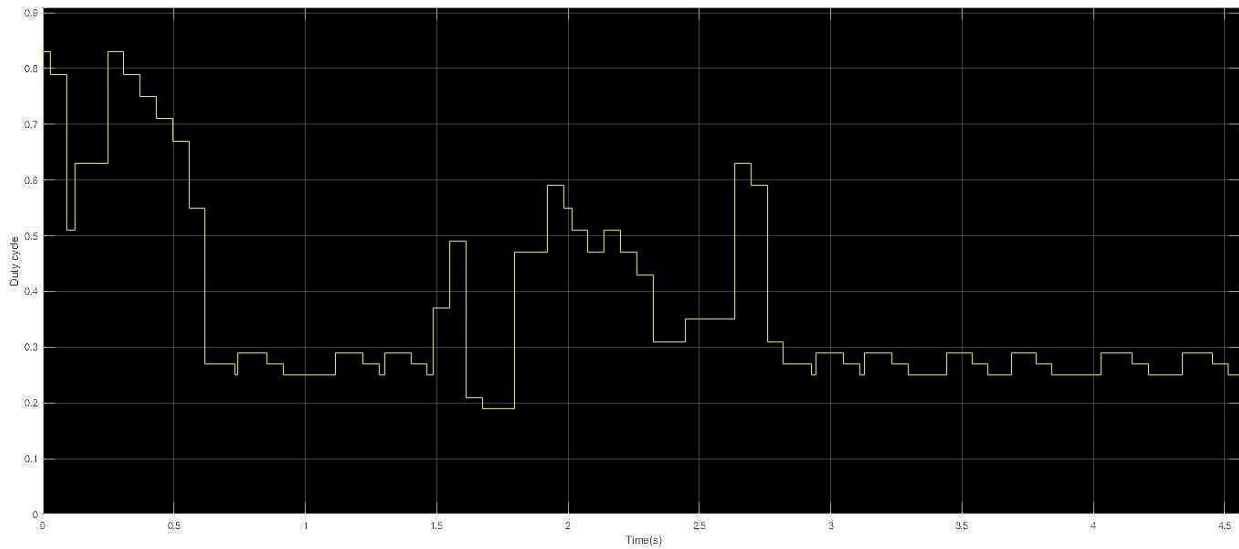
**Fig. 5-127.** Duty cycle versus time for shading pattern 13 of the third implementation of the Q-learning-based MPPT method.



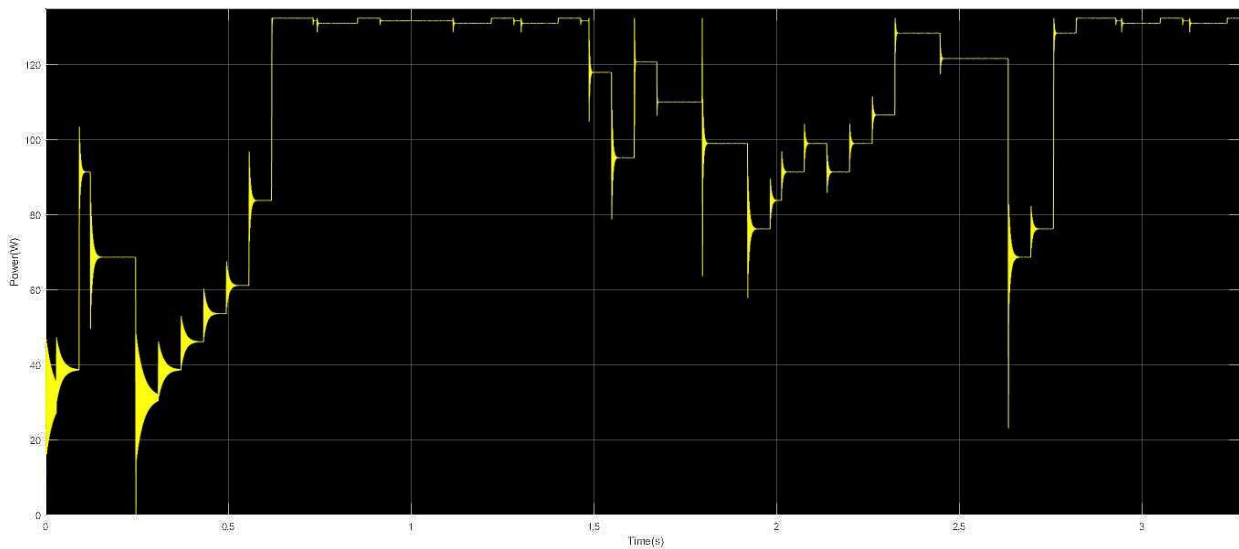
**Fig. 5-128.** PV array output power versus time for shading pattern 13 of the third implementation of the Q-learning-based MPPT method.

It is observed that the convergence time has been reduced. This is due to the fact that the agent has acquired enough knowledge from shading patterns 10, 11 and 12 in order to detect the location the GMPP in a few time-steps for shading pattern 13.

Figs. 5-129 and 5-130 present the results of the simulations of duty cycle and PV array output power versus time of the fourth implementation of the Q-learning-based MPPT method for shading pattern 13.



**Fig. 5-129.** Duty cycle versus time for shading pattern 13 of the fourth implementation of the Q-learning-based MPPT method.

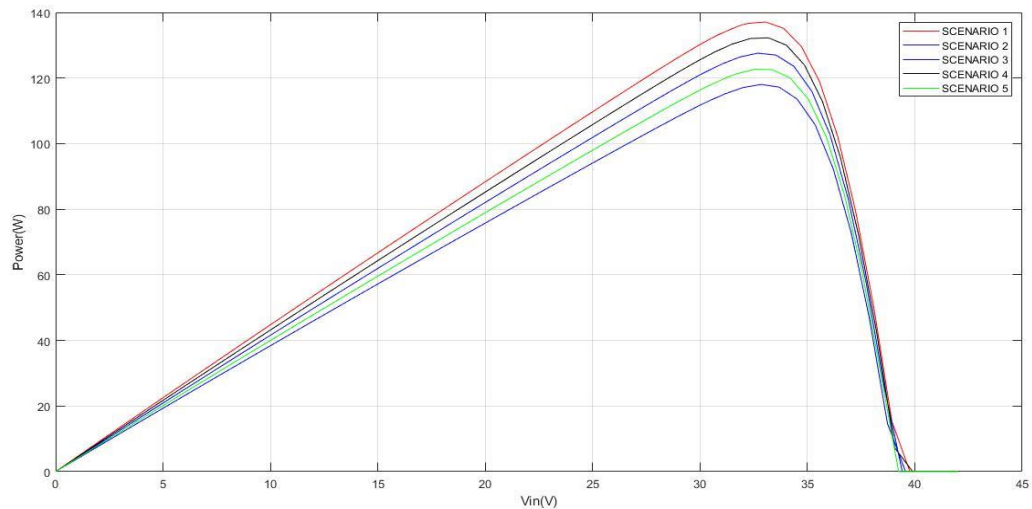


**Fig. 5-130.** PV array output power versus time for shading pattern 13 of the fourth implementation of the Q-learning-based MPPT method.

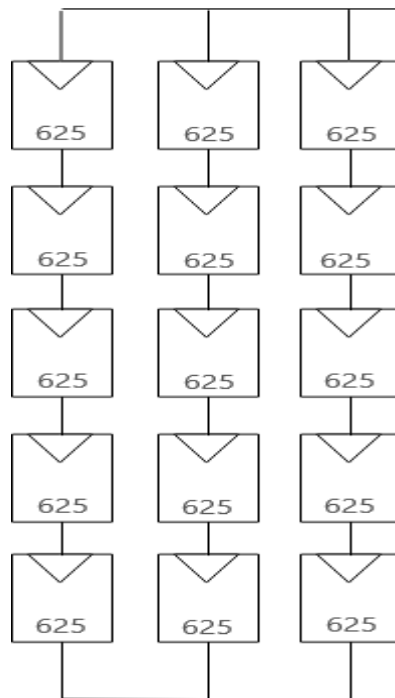
It is observed that the convergence time has been reduced. This is due to the fact that the agent has acquired enough knowledge from shading patterns 10, 11 and 12 in order to detect the location the MPP in a few time-steps for shading pattern 13.

## 5.14 Analysis of the simulation results for shading pattern 14

In Fig. 5-131 the PV array output power-voltage curve is presented and in Fig. 5-132 the distribution of incident solar irradiation over each PV module is presented. The numbers above each PV module represent the solar irradiation intensity, which is measured in  $W/m^2$ . Shading pattern 14 is a uniform solar irradiance pattern.

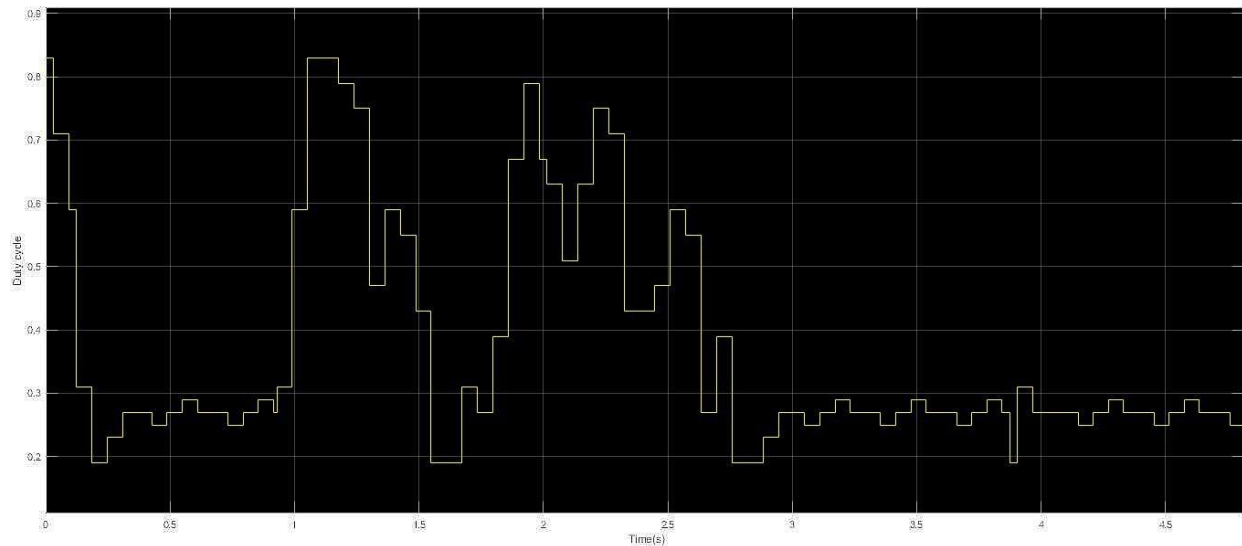


**Fig. 5-131.** The PV array output power-voltage curve for shading pattern 14.

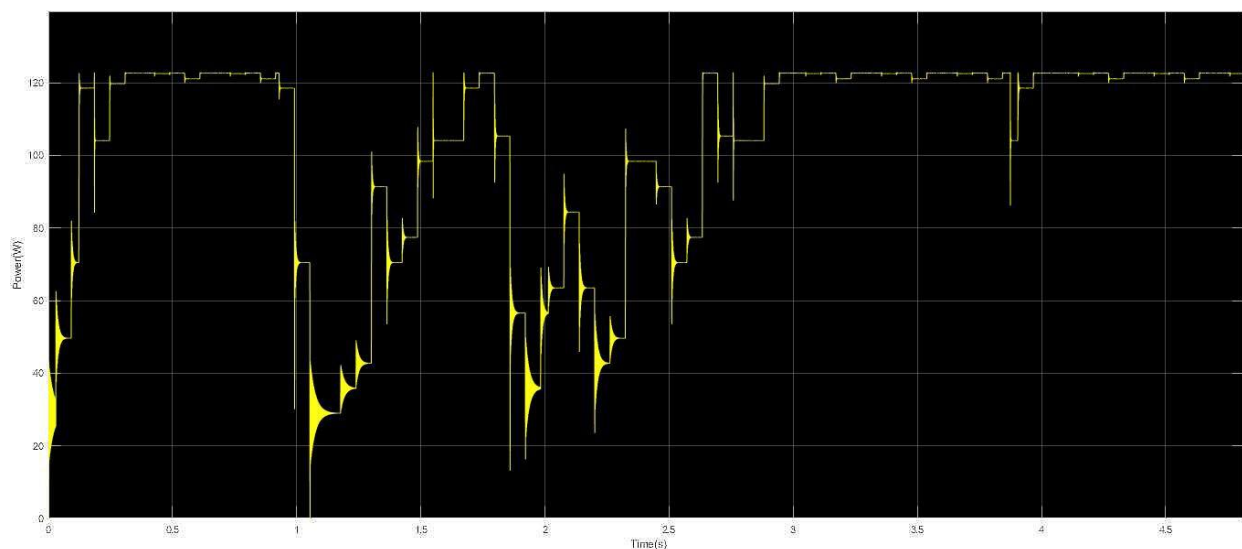


**Fig. 5-132.** The distribution of incident solar irradiation over each PV module for shading pattern 14.

Eight plots were exported, the PV array output power versus time plot and the duty cycle versus time plot in the four aforementioned implementations of the Q-learning-based MPPT method. Figs. 5-133 and 5-134 present the results of the simulations of the duty cycle and the PV array output power versus time for the first implementation of the Q-learning-based MPPT method for shading pattern 14.



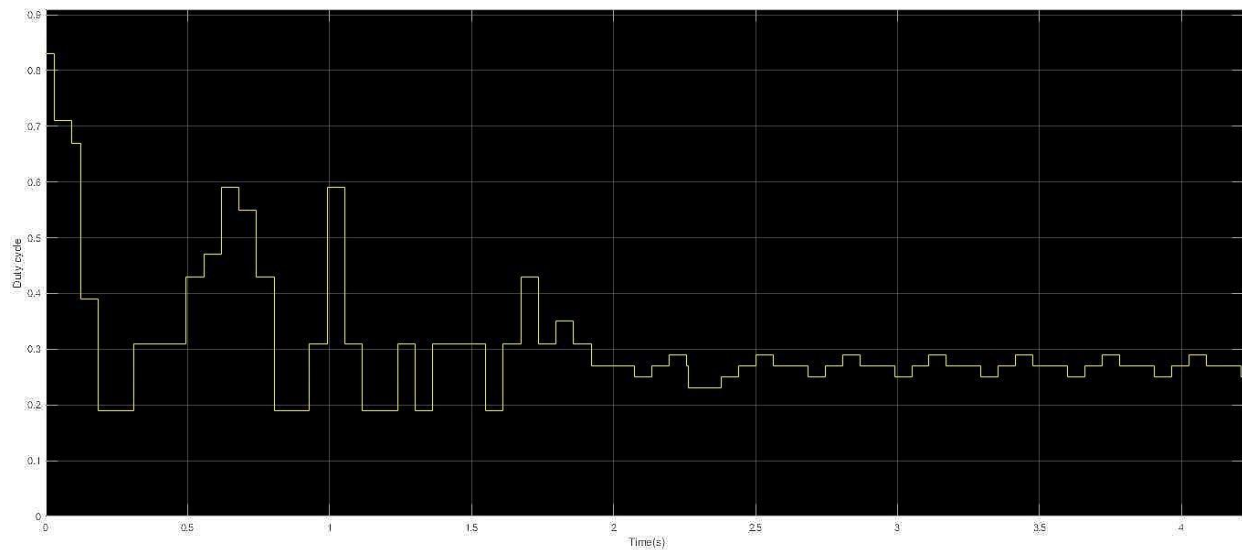
**Fig. 5-133.** Duty cycle versus time for shading pattern 14 of the first implementation of the Q-learning-based MPPT method.



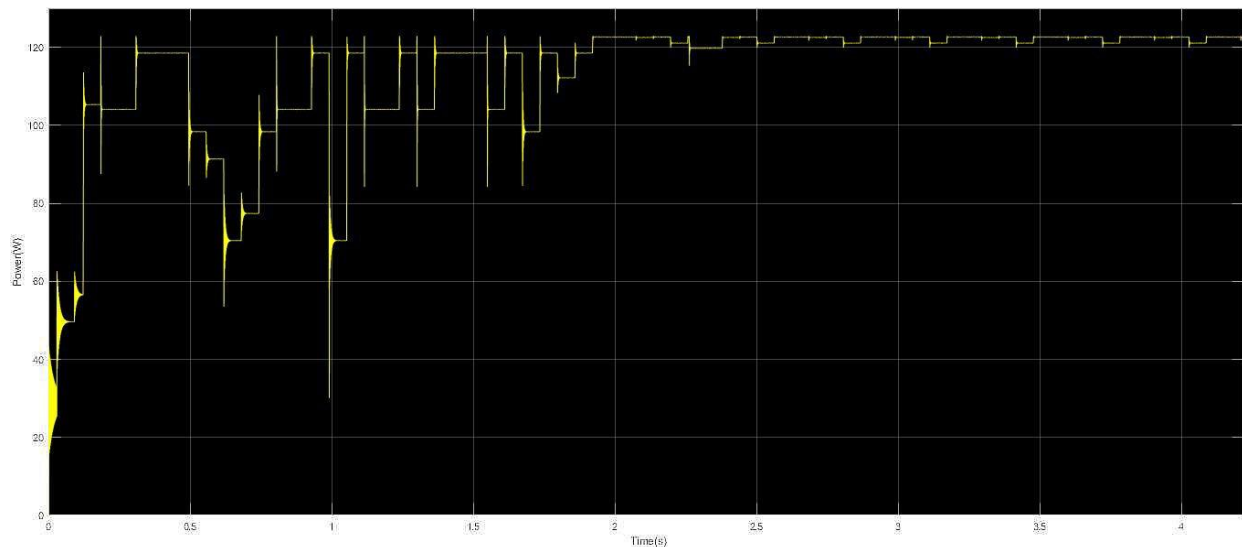
**Fig. 5-134.** PV array output power versus time for shading pattern 14 of the first implementation of the Q-learning-based MPPT method.

It is observed that the agent is able to detect the location GMPP in a few time-steps. The agent will need approximately the same convergence time for the MPPT process in unknown intermediate conditions of solar irradiation and shading pattern.

Figs. 5-135 and 5-136 present the results of the simulations of duty cycle and PV array output power versus time of the second implementation of the Q-learning-based MPPT method for shading pattern 14.



**Fig. 5-135.** Duty cycle versus time for shading pattern 14 of the second implementation of the Q-learning-based MPPT method.

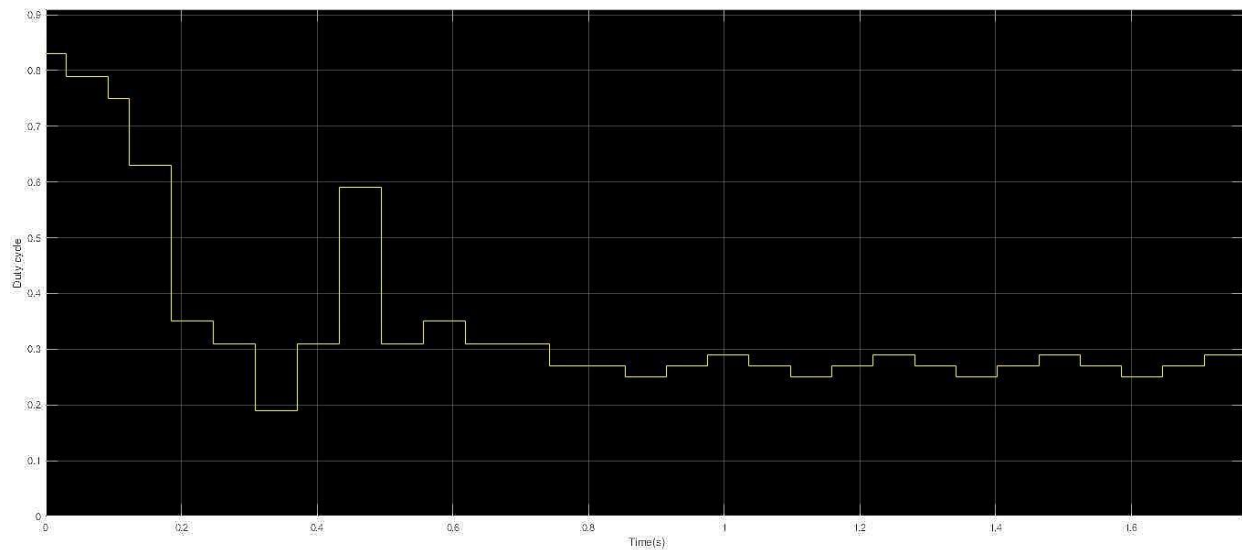


**Fig. 5-136.** PV array output power versus time for shading pattern 14 of the second implementation of the Q-learning-based MPPT method.

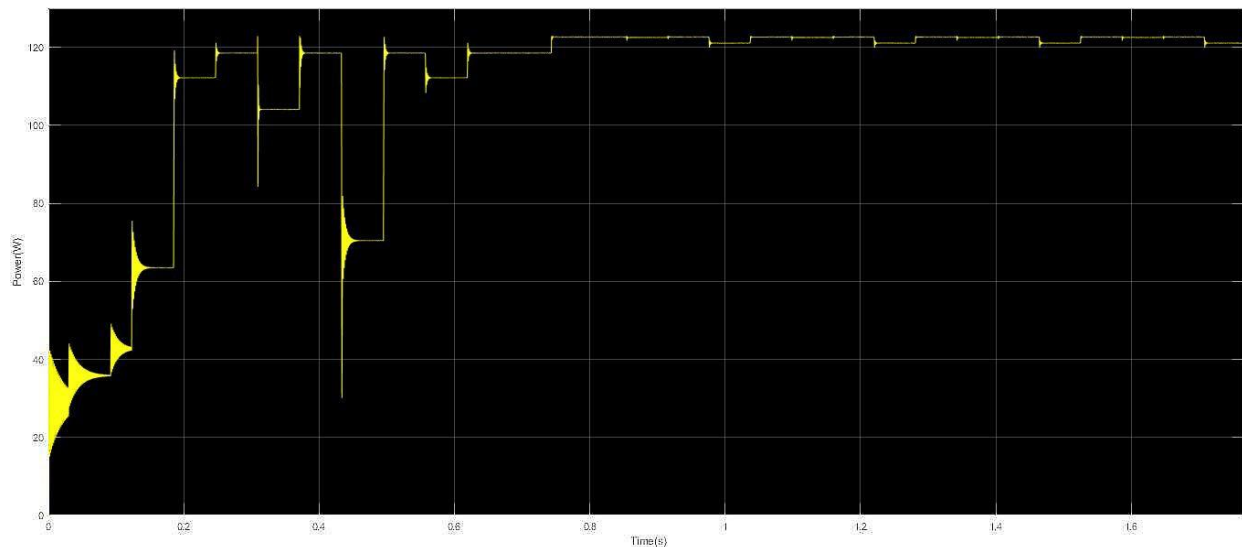
It is observed that the agent is able to detect the location GMPP in a few time-steps. The agent will need approximately the same convergence time for the MPPT process in unknown intermediate conditions of solar irradiation and shading pattern.



Figs. 5-137 and 5-138 present the results of the simulations of duty cycle and PV array output power versus time of the third implementation of the Q-learning-based MPPT method for shading pattern 14.



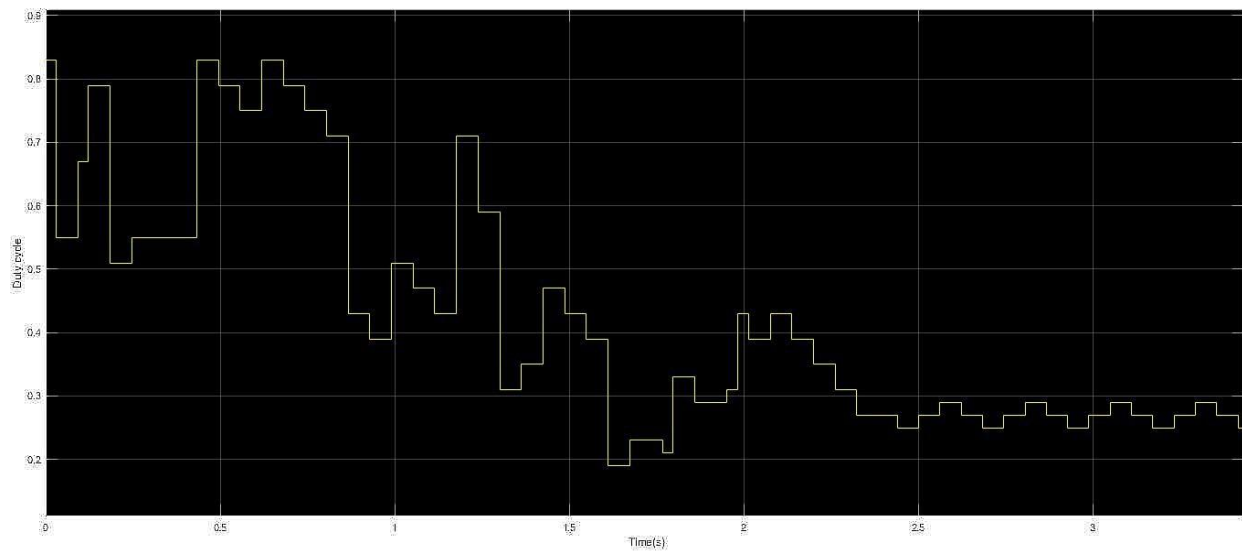
**Fig. 5-137.** Duty cycle versus time for shading pattern 14 of the third implementation of the Q-learning-based MPPT method.



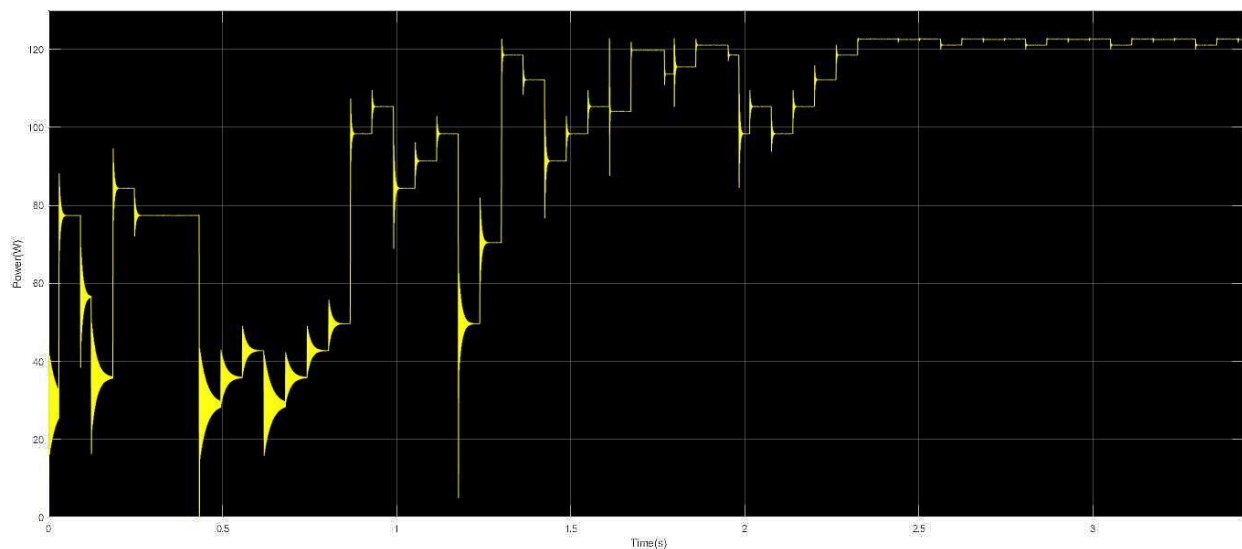
**Fig. 5-138.** Duty cycle versus time for shading pattern 14 of the third implementation of the Q-learning-based MPPT method.

It is observed that the agent is able to detect the location GMPP in a few time-steps. The agent will need approximately the same convergence time for the MPPT process in unknown intermediate conditions of solar irradiation and shading pattern.

Figs. 5-139 and 5-140 present the results of the simulations of duty cycle and PV array output power versus time of the fourth implementation of the Q-learning-based MPPT method for shading pattern 14.



**Fig. 5-139.** Duty cycle versus time for shading pattern 14 of the fourth implementation of the Q-learning-based MPPT method.



**Fig. 5-140.** PV array output power versus time for shading pattern 14 of the fourth implementation of the Q-learning-based MPPT method.

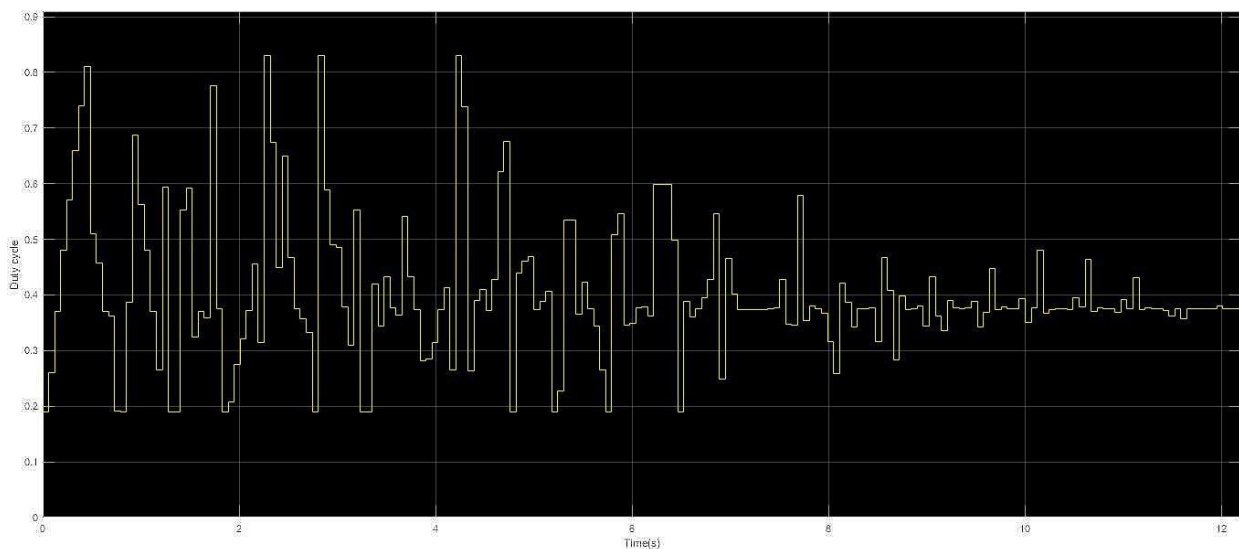
It is observed that the agent is able to detect the location MPP in a few time-steps. The agent will need approximately the same convergence time for the MPPT process in unknown intermediate conditions of solar irradiation and shading pattern.

## 5.15 Analysis of the simulation results of the PSO-based MPPT method

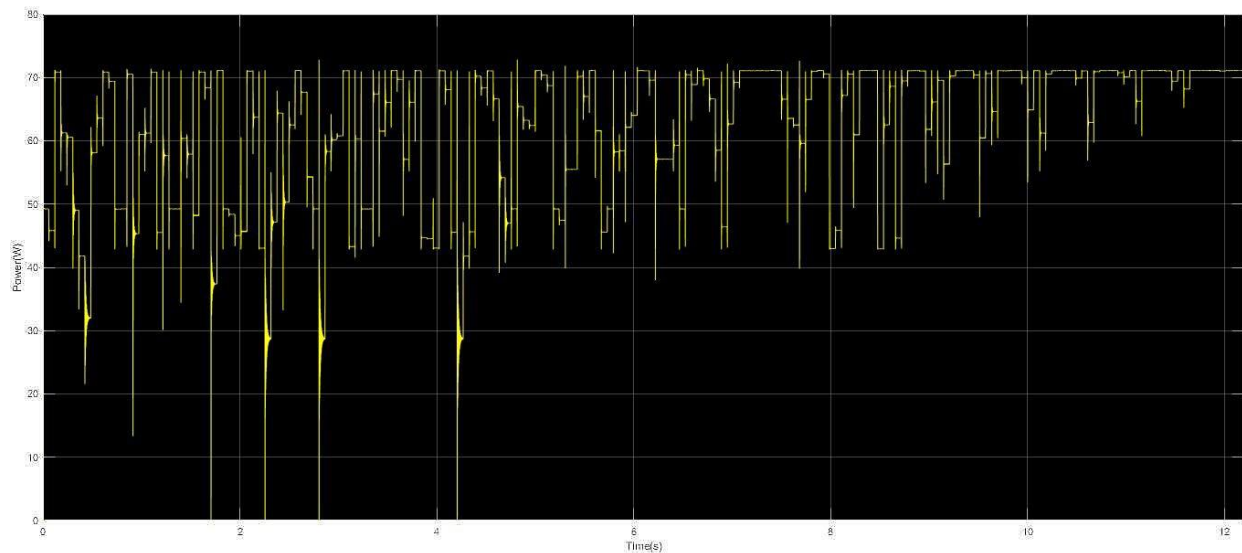
As mentioned at the beginning of Chapter 5, the PSO-based method was simulated in four shading patterns. The shading patterns 5 (non-uniform incident solar irradiation conditions), 6 (non-uniform incident solar irradiation conditions), 8 (non-uniform incident solar irradiation conditions), and 13 (uniform incident solar irradiation conditions) were chosen. Also, for every shading pattern which is related with uniform incident solar irradiation conditions the PSO algorithm needs approximately the same amount of time in order to detect the location of the MPP. Regarding the shading patterns, which are related with non-uniform incident solar irradiation conditions, the unknown intermediate conditions of solar irradiation and shading patterns were selected, in which every implementation of the Q-learning-based methods had successfully detected the location of the GMPP. The parameter settings of the PSO-based method are presented in Table 5.2.

<b>Table 5-2. Parameter settings of the PSO-based method.</b>			
Number of particles	8	w	0.6
Maximum duty cycle	0.83	c <sub>1</sub>	1.6
Minimum duty cycle	0.19	c <sub>2</sub>	1.5

Figs. 5-141 and 5-142 present the results of the simulations of duty cycle and PV array output power versus time of the PSO-based MPPT method for shading pattern 5.



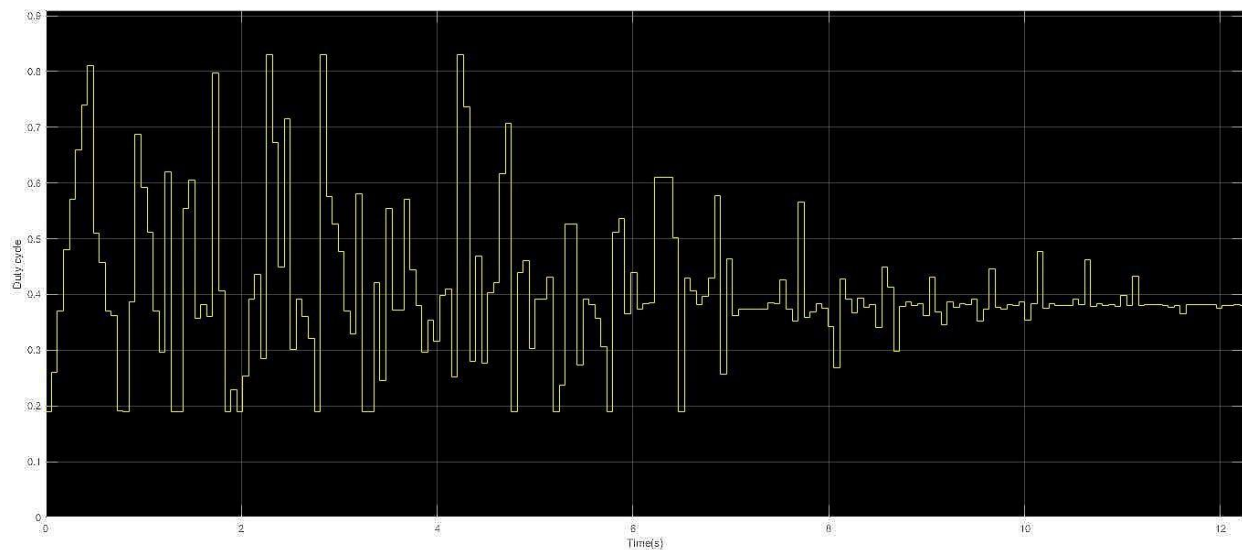
**Fig. 5-141.** Duty cycle versus time for shading pattern 5 of the PSO-based MPPT method.



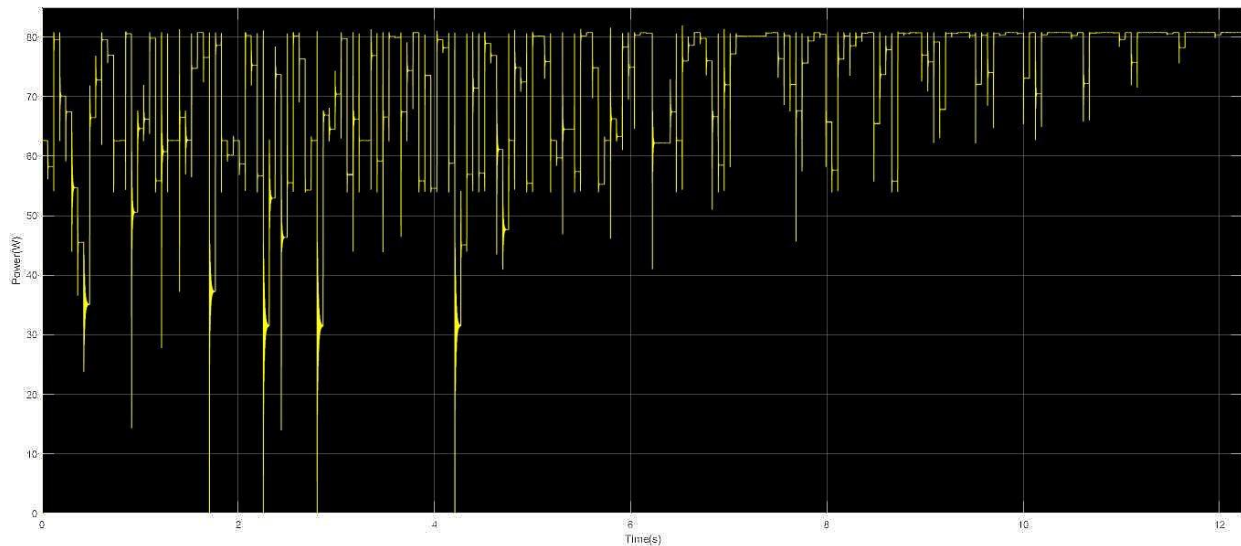
**Fig. 5-142.** PV array output power versus time for shading pattern 5 of the PSO-based MPPT method.

It is observed that the PSO-based method needs approximately 12s in order to detect the location of the GMPP for shading pattern 5.

Figs. 5-143 and 5-144 present the results of the simulations of duty cycle and PV array output power versus time of the PSO-based MPPT method for shading pattern 6.



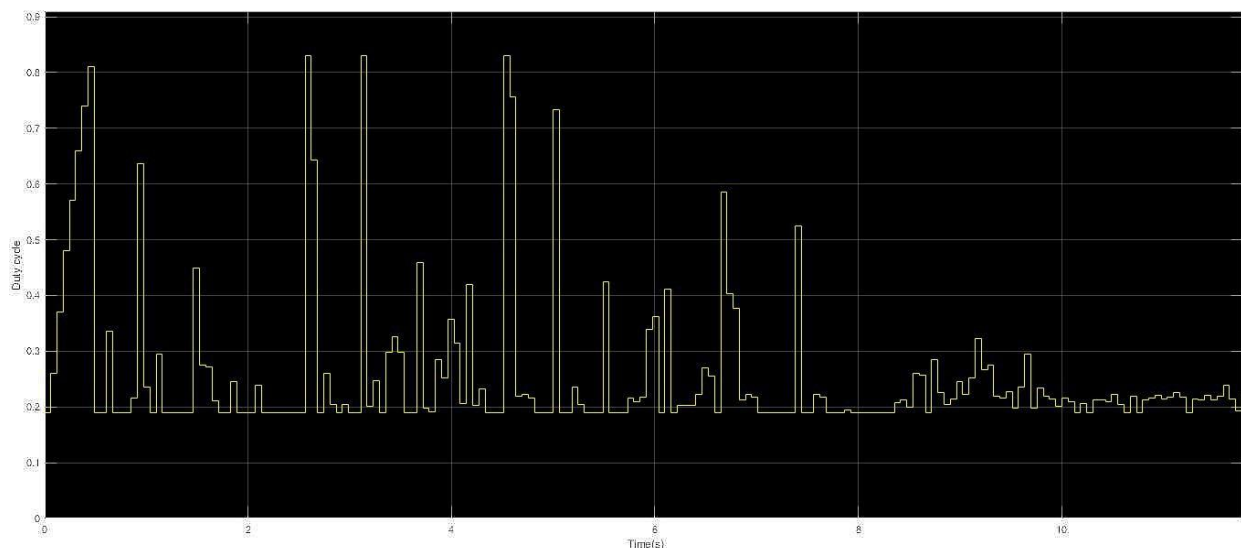
**Fig. 5-143.** Duty cycle versus time for shading pattern 6 of the PSO-based MPPT method.



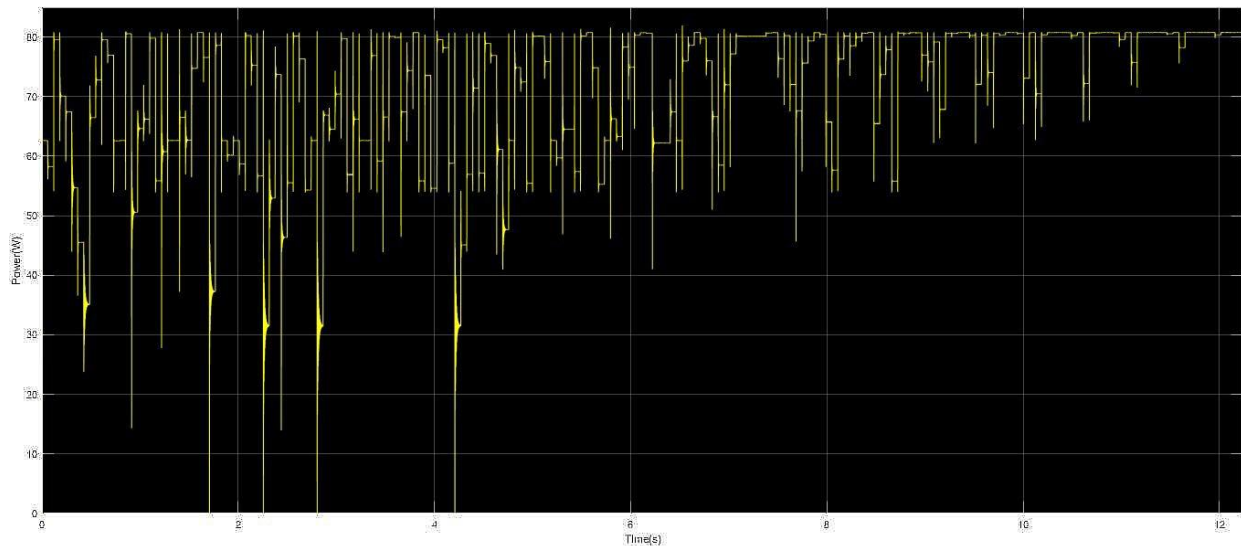
**Fig. 5-144.** PV array output power versus time for shading pattern 6 of the PSO-based MPPT method.

It is observed that the PSO-based method needs approximately 12s in order to detect the location of the MPP for shading pattern 6. This is due to the fact that every time that the shading pattern changes, every parameter of the algorithm is initialized and the knowledge that was acquired about the location of the MPP is lost. Thus, the method starts running from the beginning. This is a major difference compared with the Q-learning-based MPPT methods.

Figs. 5-145 and 5-146 present the results of the simulations of duty cycle and PV array output power versus time of the PSO-based MPPT method for shading pattern 8.



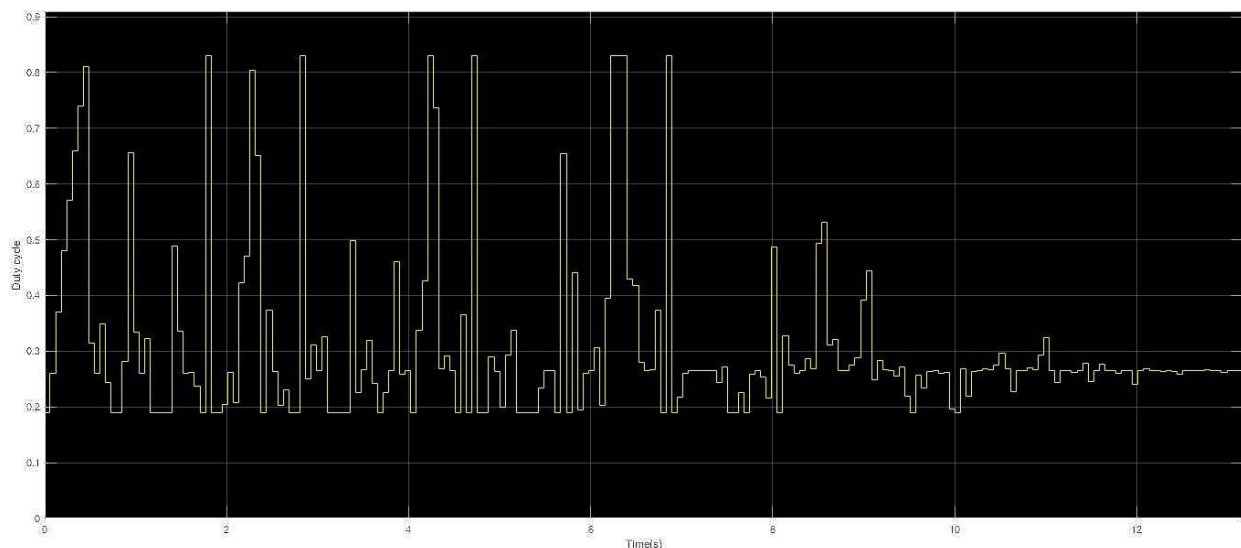
**Fig. 5-145.** Duty cycle versus time for shading pattern 8 of the PSO-based MPPT method.



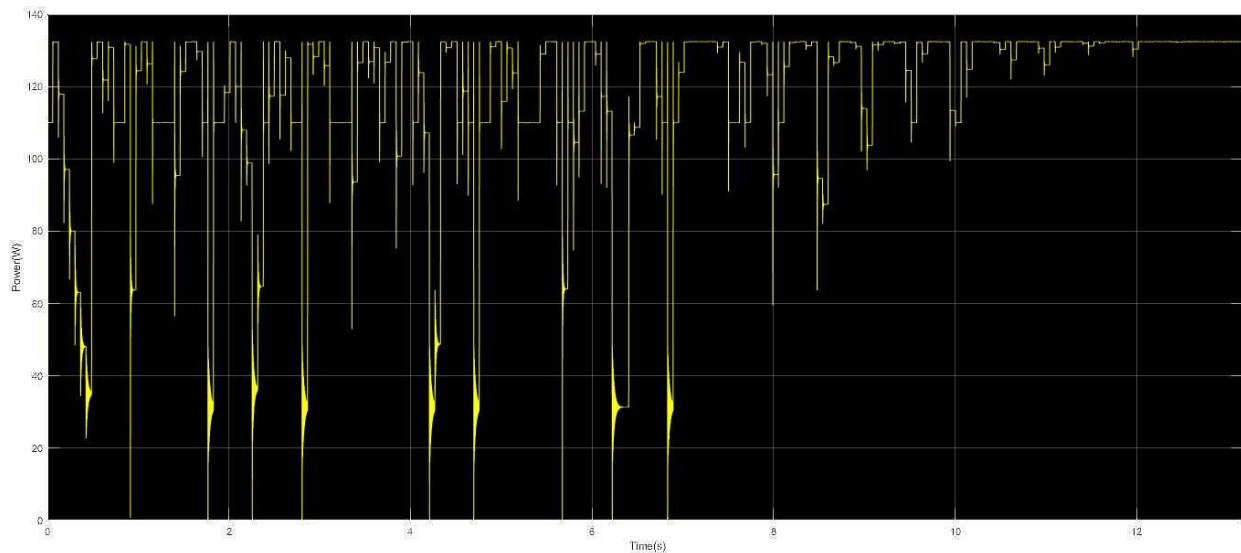
**Fig. 5-146.** PV array output power versus time for shading pattern 8 of the PSO-based MPPT method.

It is observed that the PSO-based method needs approximately 12s in order to detect the location of the MPP for shading pattern 8. The initial assumption is confirmed, that in most cases the PSO-based MPPT method takes a specific amount of time to locate the MPP.

Figs. 5-147 and 5-148 present the results of the simulations of duty cycle and PV array output power versus time of the PSO-based MPPT method for shading pattern 13.



**Fig. 5-147.** Duty cycle versus time for shading pattern 13 of the PSO-based MPPT method.



**Fig. 5-148.** PV array output power versus time for shading pattern 13 of the PSO-based MPPT method.

It is observed that the PSO-based method needs approximately 12s in order to detect the location of the MPP for shading pattern 13. It is clear from the results that the PSO-based MPPT method has one major advantage and one major disadvantage over the Q-learning-based MPPT method. Its advantage is that it provides a very high MPPT efficiency under all environmental conditions. Its disadvantage compared with the Q-learning-based MPPT methods is that the PSO-based MPPT method is unable able to locate the GMPP in a few time-steps.

## 5.16 Evaluation of the simulation results

In this Section, all simulation results of each implementation of the Q-learning-based MPPT method are presented for each shading pattern. Tables 5.3 - 5.4 present the first implementation results. Tables 5.5 – 5.6 present the second implementation results. Tables 5.7 – 5.8 present the third implementation results. Tables 5.9 – 5.10 present the fourth implementation results. Each table contains the convergence time of the MPPT process, the number of duty cycle changes until the beginning of the oscillations, the theoretical GMPP value, the calculated GMPP value and the MPPT efficiency for each shading pattern separately.

**Table 5-3. The results of the first implementation of the Q-learning-based MPPT method under non-uniform solar irradiation.**

	<b>Convergence time (s)</b>	<b>Number of duty cycle changes</b>	<b>Theoretical <math>P_{MPPT}</math> (W)</b>	<b>Calculated <math>P_{MPPT}</math> (W)</b>	<b>MPPT Efficiency</b>
<b>Shading pattern 1</b>	28	395	67.015	66.2	0.98
<b>Shading pattern 2</b>	16	225	83.1516	82	0.98
<b>Shading pattern 3</b>	7	99	75.104	73	0.97
<b>Shading pattern 4</b>	0.3	5	79.09	74	0.936
<b>Shading pattern 5</b>	0.3	4	71.118	69.8	0.98
<b>Shading pattern 6</b>	0.8	11	80.73	79.5	0.98
<b>Shading pattern 7</b>	1.6	20	81.89	76	0.928
<b>Shading pattern 8</b>	11.2	154	78.15	77.5	0.99
<b>Shading pattern 9</b>	6.2	87	90.44	63	0.696

According to the results presented in Table 5-3, it is clear that the knowledge that is acquired by the agent during the learning process of shading pattern 1 is affecting the duration of learning process of the shading patterns 2 and 3, respectively. After the learning process of shading patterns 1, 2 and 3, the agent knows how to react when the shading patterns 4, 5 and 6 are



applied. The convergence time for shading pattern 8 is higher than the convergence time for the above-mentioned shading patterns, because only one small part of the PV array output power-voltage curve is known by the MPPT process for the aforementioned shading patterns. In addition, it is observed that the agent is unable to detect the GMPP in shading patterns 7 and 9. This is due the fact that the PV array output power is quantized in 15 equal steps for each state. Each step includes 10 W ( $150/15=10$  W, 150 W is the maximum PV array output power that could be obtained by the operational characteristics of the PV modules). In addition, it can be observed that the duty cycle value in the MPP for shading pattern 5 is equal to 0.39 and the PV array output power is equal to 69.8 W, while the power value in which the agent converges for shading pattern 9 is equal to 63 W and the duty cycle value is equal to 0.39. It is clear that the agent incorrectly considered that these two states are identical. The reason that this happens is that each step of the PV array output power includes many values. As a result, it confuses different states as identical. This method's benefit is that the agent needs to learn how to detect the location the GMPP for a few shading patterns in order to reduce the convergence time in unknown intermediate conditions of solar irradiation and shading pattern. Nevertheless, it is quite possible to fail to detect the GMPP. Therefore, the acquired knowledge by the previous shading patterns does not always help the agent to detect the location the GMPP.

**Table 5-4. The results of the first implementation of the Q-learning-based MPPT method under uniform solar irradiation.**

	<b>Convergence time (s)</b>	<b>Number of duty cycle changes</b>	<b>Theoretical <math>P_{MPPT}</math> (W)</b>	<b>Calculated <math>P_{MPPT}</math> (W)</b>	<b>MPPT Efficiency</b>
<b>Shading pattern 10</b>	15.6	217	136.98	136.2	0.99
<b>Shading pattern 11</b>	11.7	161	118.04	117.8	0.99
<b>Shading pattern 12</b>	13.5	188	127.53	126.4	0.99
<b>Shading pattern 13</b>	1.2	14	132.26	131.8	0.996
<b>Shading pattern 14</b>	2.8	36	122.8	122.1	0.99

It is observed in Table 5-4 that the knowledge that is acquired by the agent during the learning process of shading patterns 10, 11 and 12 helped the agent to detect the MPP of the intermediate shading patterns, such as 13 and 14 in a few time-steps. The MPPT process under uniform solar irradiation conditions is simple enough and that's the reason that this Q-learning-based MPPT method has high MPPT efficiency.

In Tables 5.5 and 5.6 the simulation results of the second implementation of the Q-learning-based MPPT method are presented, under uniform and non-uniform solar irradiation conditions respectively.

<b>Table 5-5. The results of the second implementation of the Q-learning-based MPPT method under non-uniform solar irradiation.</b>					
	<b>Convergence time (s)</b>	<b>Number of duty cycle changes</b>	<b>Theoretical <math>P_{MPPT}</math> (W)</b>	<b>Calculated <math>P_{MPPT}</math> (W)</b>	<b>MPPT Efficiency</b>
<b>Shading pattern 1</b>	32	442	67.015	66.2	0.98
<b>Shading pattern 2</b>	28	390	83.1516	82	0.98
<b>Shading pattern 3</b>	26	362	75.104	73	0.97
<b>Shading pattern 4</b>	8.1	108	79.09	74	0.936
<b>Shading pattern 5</b>	6.5	86	71.118	69.8	0.98
<b>Shading pattern 6</b>	1.2	13	80.73	79.5	0.98
<b>Shading pattern 7</b>	1.8	22	81.89	78	0.9524
<b>Shading pattern 8</b>	17.8	251	78.15	77.5	0.99
<b>Shading pattern 9</b>	31	380	90.44	88.9	0.9829

<b>Table 5-6. The results of the second implementation of the Q-learning-based MPPT method under uniform solar irradiation.</b>					
	<b>Convergence time (s)</b>	<b>Number of duty cycle changes</b>	<b>Theoretical <math>P_{MPPT}</math> (W)</b>	<b>Calculated <math>P_{MPPT}</math> (W)</b>	<b>MPPT Efficiency</b>
<b>Shading pattern 10</b>	18.1	249	136.98	136.2	0.99
<b>Shading pattern 11</b>	19.5	275	118.04	117.8	0.99
<b>Shading pattern 12</b>	22.6	313	127.53	126.4	0.99
<b>Shading pattern 13</b>	4.1	44	132.26	131.8	0.996
<b>Shading pattern 14</b>	2.4	29	122.8	122.1	0.99

It is noted in Tables 5-6 and 5-7 that there is a better accuracy in the detection of the GMPP location in all investigated cases. However, the learning process time is longer and the agent needs to acquire more knowledge (the second implementation of the Q-learning-based MPPT method needs to be tested for more shading patterns compared with the first) in order to know how to react in unknown intermediate conditions of solar irradiation and shading pattern. For instance, in order to learn how to detect the GMPP in shading pattern 6, the agent needs to learn how to detect the location of the GMPP in shading patterns 1, 2, 3, 4 and 5 previously. Corresponding observations also apply in the case of uniform solar irradiation conditions. It is also observed that a much better accuracy is achieved in shading pattern 7, in a few time-steps. Shading pattern 8 needs more learning time compared to the first implementation of the Q-learning-based MPPT method and shading pattern 9 has a much higher accuracy. This is due to the fact that in the second implementation of the Q-learning-based method, the PV array output power is quantized in 30 equal steps of 5 W ( $150/30=5$  W, 150 W is the maximum PV array output power that could be obtained by the operational characteristics of the PV modules). Thus, it is more difficult for the agent to consider that two states are identical. Regarding the shading

patterns, which are related with the uniform solar irradiation conditions, a longer convergence time is observed.

In Tables 5.7 and 5.8 the simulation results of the third implementation of the Q-learning-based MPPT method are presented, under uniform and non-uniform solar irradiation conditions respectively.

<b>Table 5-7. The results of the third implementation of the Q-learning-based MPPT method under non-uniform solar irradiation.</b>					
	<b>Convergence time (s)</b>	<b>Number of duty cycle changes</b>	<b>Theoretical <math>P_{MPPT}</math> (W)</b>	<b>Calculated <math>P_{MPPT}</math> (W)</b>	<b>MPPT Efficiency</b>
<b>Shading pattern 1</b>	12.1	166	67.015	66.2	0.98
<b>Shading pattern 2</b>	8.3	111	83.1516	82	0.98
<b>Shading pattern 3</b>	0.42	7	75.104	73	0.97
<b>Shading pattern 4</b>	0.3	4	79.09	77.5	0.97
<b>Shading pattern 5</b>	1.5	17	71.118	69.8	0.98
<b>Shading pattern 6</b>	0.62	9	80.73	79.5	0.98
<b>Shading pattern 7</b>	0.4	6	81.89	78	0.9524
<b>Shading pattern 8</b>	6.2	83	78.15	77.5	0.99
<b>Shading pattern 9</b>	1.2	16	90.44	63	0.696

**Table 5-8. The results of the third implementation of the Q-learning-based MPPT method under uniform solar irradiation.**

	<b>Convergence time (s)</b>	<b>Number of duty cycle changes</b>	<b>Theoretical <math>P_{MPPT}</math> (W)</b>	<b>Calculated <math>P_{MPPT}</math> (W)</b>	<b>MPPT Efficiency</b>
<b>Shading pattern 10</b>	6.3	83	136.98	136.2	0.99
<b>Shading pattern 11</b>	5.8	78	118.04	117.8	0.99
<b>Shading pattern 12</b>	4.5	58	127.53	126.4	0.99
<b>Shading pattern 13</b>	1.3	15	132.26	131.8	0.996
<b>Shading pattern 14</b>	0.78	11	122.8	122.1	0.99

It is observed that the MPPT efficiency is similar with the first implementation of the Q-learning-based MPPT method. This is because every state is defined by the same number of equal quantized steps of the duty cycle and equal quantized steps of the PV array output power. The shorter convergence time is due to the fact that the agent does not need any knowledge of duty cycle value in the previous time-step. This greatly reduces the number of the required time-steps until the agent converges to a value. It should be mentioned that in this implementation as well, the agent fails to detect the location the GMPP for shading pattern 9. The reason is the same as that of the first implementation of the Q-learning-based method.

In Tables 5.9 and 5.10 the simulation results of the fourth implementation of the Q-learning-based MPPT method are presented, under uniform and non-uniform solar irradiation conditions respectively.

<b>Table 5-9. The results of the fourth implementation of the Q-learning-based MPPT method under non-uniform solar irradiation.</b>					
	<b>Convergence time (s)</b>	<b>Number of duty cycle changes</b>	<b>Theoretical <math>P_{MPPT}</math> (W)</b>	<b>Calculated <math>P_{MPPT}</math> (W)</b>	<b>MPPT Efficiency</b>
<b>Shading pattern 1</b>	11.5	157	67.015	66.2	0.98
<b>Shading pattern 2</b>	13.2	179	83.1516	82	0.98
<b>Shading pattern 3</b>	8.8	121	75.104	73	0.97
<b>Shading pattern 4</b>	0.8	11	79.09	74	0.97
<b>Shading pattern 5</b>	3.8	49	71.118	69.8	0.98
<b>Shading pattern 6</b>	0.65	9	80.73	79.5	0.98
<b>Shading pattern 7</b>	0.38	6	81.89	78	0.9524
<b>Shading pattern 8</b>	9.1	123	78.15	77.5	0.99
<b>Shading pattern 9</b>	11.8	161	90.44	73	0.81

**Table 5-10. The results of the fourth implementation of the Q-learning-based MPPT method under uniform solar irradiation.**

	<b>Convergence time (s)</b>	<b>Number of duty cycle changes</b>	<b>Theoretical <math>P_{MPPT}</math> (W)</b>	<b>Calculated <math>P_{MPPT}</math> (W)</b>	<b>MPPT Efficiency</b>
<b>Shading pattern 10</b>	10.2	138	136.98	136.2	0.99
<b>Shading pattern 11</b>	10.5	141	118.04	117.8	0.99
<b>Shading pattern 12</b>	9	119	127.53	126.4	0.99
<b>Shading pattern 13</b>	2.8	37	132.26	131.8	0.996
<b>Shading pattern 14</b>	2.4	34	122.8	122.1	0.99

It is observed that the MPPT efficiency is similar with the second implementation of the Q-learning-based MPPT method. This is because every state is defined by the same number of equal quantized steps of the duty cycle and equal quantized steps of the PV array output power. The shorter convergence time is due to the fact that the agent does not need any knowledge of duty cycle value in the previous time-step. This greatly reduces the number of the required time-steps until the agent converges to a value. However, it is observed that in this implementation as well, the agent fails to detect the location the MPP for shading pattern 9. Nevertheless, it presents better MPPT efficiency compared with the first and the third implementation of the Q-learning-based method.

Table 5.11 presents the tracking performance of the PSO-based MPPT method for shading patterns 5, 6, 8 and 13. Tables 5.12 – 5.15 present the tracking performances of the first, second, third and the fourth implementation of the Q-learning-based MPPT methods for shading patterns 5, 6, 8 and 13, respectively. Each table contains the convergence time of the MPPT process for each method, the energy that is produced during the MPPT process, the calculated MPP and the MPPT efficiency. Then, a comparison of the performance of the Q-learning-based MPPT



methods with the PSO-based MPPT method is performed. The energy production,  $E_{total}$  (Wh) was calculated as follows:

$$E_{total} = \sum_{\forall i} \frac{P_i \cdot t_i}{3600} \quad (5.2)$$

where

$P_i$  is the PV array output power (W) and

$t_i$  is the time-step, which is equal to 0.071 sec

The number of time-steps for each shading pattern depends on the convergence time of the corresponding implementation of the Q-learning MPPT method.

In Table 5.11 the simulation results of the PSO-based MPPT method are presented.

<b>Table 5-11. Tracking performance of the PSO-based MPPT method.</b>				
	<b>Convergence time (s)</b>	<b>Energy production (Wh)</b>	<b>Calculated <math>P_{MPPT}</math> (W)</b>	<b>MPPT Efficiency</b>
<b>Shading pattern 5</b>	11.5	0.20256	71.1	0.99
<b>Shading pattern 6</b>	11.6	0.23363	80.5	0.99
<b>Shading pattern 8</b>	10.8	0.2285	78.15	0.99
<b>Shading pattern 13</b>	11	0.333	132.1	0.99

In Table 5.12 the simulation results of the first implementation of the Q-learning-based MPPT method are presented.

<b>Table 5-12. Tracking performance of the first implementation of the Q-learning-based MPPT method.</b>				
	<b>Convergence time (s)</b>	<b>Energy production (Wh)</b>	<b>Calculated <math>P_{MPPT}</math> (W)</b>	<b>MPPT Efficiency</b>
<b>Shading pattern 5</b>	0.3	0.22865	69.8	0.9817
<b>Shading pattern 6</b>	0.8	0.2497	79.5	0.984
<b>Shading pattern 8</b>	11.2	0.2102	77.5	0.9916
<b>Shading pattern 13</b>	1.2	0.3821	131.8	0.997

In Table 5.13 the simulation results of the second implementation of the Q-learning-based MPPT method are presented.

<b>Table 5-13. Tracking performance of the second implementation of the Q-learning-based MPPT method.</b>				
	<b>Convergence time (s)</b>	<b>Energy production (Wh)</b>	<b>Calculated <math>P_{MPPT}</math> (W)</b>	<b>MPPT Efficiency</b>
<b>Shading pattern 5</b>	6.5	0.2069	69.8	0.9817
<b>Shading pattern 6</b>	1.2	0.2483	79	0.9801
<b>Shading pattern 8</b>	17.8	0.1927	77.5	0.9916
<b>Shading pattern 13</b>	4.1	0.3727	131.8	0.997

In Table 5.14 the simulation results of the third implementation of the Q-learning-based MPPT method are presented.

<b>Table 5-14. Tracking performance of the third implementation of the Q-learning-based MPPT method.</b>				
	<b>Convergence time (s)</b>	<b>Energy production (Wh)</b>	<b>Calculated <math>P_{MPPT}</math> (W)</b>	<b>MPPT Efficiency</b>
<b>Shading pattern 5</b>	1.5	0.2276	69.8	0.9817
<b>Shading pattern 6</b>	0.3	0.2530	79	0.9801
<b>Shading pattern 8</b>	6.2	0.2355	77.6	0.9916
<b>Shading pattern 13</b>	1.3	0.3902	131.8	0.997

In Table 5.15 the simulation results of the fourth implementation of the Q-learning-based MPPT method are presented.

<b>Table 5-15. Tracking performance of the fourth implementation of the Q-learning-based MPPT method.</b>				
	<b>Convergence time (s)</b>	<b>Energy production (Wh)</b>	<b>Calculated <math>P_{MPPT}</math> (W)</b>	<b>MPPT Efficiency</b>
<b>Shading pattern 5</b>	3.8	0.2175	69.8	0.9817
<b>Shading pattern 6</b>	0.7	0.2516	79	0.9801
<b>Shading pattern 8</b>	9.1	0.2216	77.6	0.9916
<b>Shading pattern 13</b>	2.8	0.3875	131.8	0.997

Regarding shading pattern 5, the first implementation of the Q-learning-based MPPT method produces the highest amount of energy, followed by the third implementation of the Q-learning-based MPPT method, the fourth implementation, the second implementation and the PSO-based MPPT method. For shading pattern 6, the third implementation of the Q-learning-based MPPT method produces the highest amount of energy, followed by the fourth implementation of the Q-learning-based MPPT method, the second implementation, the first implementation and the PSO-based MPPT method. When shading pattern 8 was applied, the third implementation of the Q-learning-based MPPT method produced the highest amount of energy, followed by the PSO-based MPPT method, the fourth implementation, the first implementation and the third implementation. Finally, regarding shading pattern 13, the third implementation of the Q-learning-based MPPT method produces the highest amount of energy, followed by the fourth implementation, the first implementation, the second implementation and PSO-based MPPT method. From the above it is obvious that the PSO-based MPPT method performs worse than the Q-learning-based methods, provided that an integrated learning process has been applied before.

## 5.17 Overall remarks

### Tests for uniform incident solar irradiation on the PV array:

Due to fact that there is only one peak in the PV array output power-voltage curve, after a short learning period, the agent is able to detect the GMPP in a small number of time-steps with very high MPPT efficiency (approximately 99% for each case).

### In the simulation results presented in Sections 5.1-5.15:

- **First implementation of the Q-learning-based MPPT method** needs approximately 13s for learning how to detect the MPP and 2s for detecting the MPP in unknown intermediate conditions of solar irradiation and shading pattern.
- **Second implementation of the Q-learning-based MPPT method** needs approximately 20s for learning how to detect MPP and 3s for detecting the MPP in unknown intermediate conditions of solar irradiation and shading pattern.
- **Third implementation of the Q-learning-based MPPT method** needs approximately 6s for learning how to detect MPP and 2.5s for detecting the MPP in unknown intermediate conditions of solar irradiation and shading pattern.
- **Fourth implementation of the Q-learning-based MPPT method** needs approximately 10s for learning how to detect MPP and 2.5s for detecting the MPP in unknown intermediate conditions of solar irradiation and shading pattern.
- The **PSO-based MPPT method** needs approximately 11s in order to detect the MPP for every case under uniform solar irradiation conditions.

### Tests for non-uniform incident solar irradiation on the PV array:

Studying MPPT under partial shading is a much more complex problem. This happens because there are many peaks in the PV array output power-voltage curve of the PV array and the position of the GMPP may change during abrupt changes of the environmental conditions. This affects both the learning period and the MPPT efficiency of the MPPT method based on Q-learning.

### In the simulation results presented in Sections 5.1-5.15:

- **First implementation of the Q-learning-based MPPT method** needs approximately 17s for learning how to detect MPP and usually less than 1s for detecting the MPP in unknown intermediate conditions. However, it is observed that prior knowledge does not

always help the agent to locate the MPP. For instance, in shading patterns 7 and 9, some of the intermediate states are recognized as MPP of a previous shading pattern and this leads the algorithm to convergence at a point which is not the GMPP. On the other hand, it is observed that in shading pattern 8, prior knowledge helped the agent to detect the MPP in 11s, which is approximately one third of the learning time for an unknown shading pattern.

- **Second implementation of the Q-learning-based MPPT method** needs approximately 28s for learning how to detect MPP and usually less than 2s for detecting the MPP in unknown intermediate conditions. A key difference with the first case is that the agent needs to learn more shading patterns in order to know how to react in unknown patterns. However, prior knowledge helps the agent to detect every MPP in less time. For instance, patterns 1, 2, and 3 need 28s, while 4 and 5 need 7s. As regards patterns 7 and 9, it is noted clearly a higher MPPT efficiency. The agent is able to detect the location of the MPP in both cases, but in shading pattern 9 it needs a long learning period (30s). Pattern 8 is treated similarly with the first case by the agent, but with a longer learning period (17s). Generally, a much better MPPT efficiency is observed and a longer learning period.
- **Third implementation of the Q-learning-based MPPT method** needs approximately 10s for learning how to detect the MPP and usually less than 1s for detecting the MPP in unknown intermediate conditions. Similar results were derived with the first case, except that shading pattern 3 is learned by the agent with the prior knowledge that was acquired during the learning period of patterns 1 and 2.
- **Fourth implementation of the Q-learning-based MPPT method** needs approximately 11s for learning how to detect the MPP and usually less than 1s for detecting the MPP for unknown intermediate conditions. Similar results were derived with the second case, except that shading pattern 9 has lower MPPT efficiency.
- The **PSO MPPT method** needs approximately 11s in order to detect the MPP for every case of shading pattern under non-uniform solar irradiation conditions.

# 6.

## Conclusions

---

The detection of the location of the GMPP is certainly one major factor in order to maximize the power generated by the PV arrays under uniform and non-uniform solar irradiation conditions. The GMPP is represented by the highest peak of the PV array output power-voltage curve. The process of detecting the duty cycle that will make the PV system to operate at this point (GMPP) is being called MPPT process. However, this process may often not be successful. This can happen during the partial shading of the PV panels, where multiple peaks appear on the PV array output power-voltage curve due to the use of the bypass diodes.

In this thesis, the design of a PV MPPT system was performed, which is based on reinforcement learning. For that purpose, a Q-learning-based MPPT method was applied. This algorithm was developed in Matlab<sup>TM</sup>/Simulink software. For the implementation of the PV system, a DC/DC Boost-type power converter, a PV array, an MPPT control unit and a load were designed in Matlab<sup>TM</sup>/Simulink.

The PV system under study was simulated for multiple alternative shading patterns of the PV array. These shading patterns include uniform solar irradiation patterns (5 shading patterns) and non-uniform solar irradiation patterns (9 shading patterns). The Q-learning-based MPPT method was simulated in 4 different cases. In the first case, the state-space was quantized in 18 equal steps of the duty cycle in the current time-step and in 15 equal steps of the generated PV power, as well as in 9 equal steps of the value of duty cycle in a previous time-step. In the second case, the state-space was quantized in 18 equal steps of the duty cycle in the current time-step and in 30 equal steps of the PV power, as well as in 9 equal steps of the value of duty cycle in a



previous time-step. In the third case, the state-space was quantized in 18 equal steps of the duty cycle in the current time-step and in 15 equal steps of the PV power. In the fourth case, the state-space was quantized in 18 equal steps of the duty cycle in the current time-step and in 30 equal steps of the PV power. For comparison purposes, a PSO-based MPPT method was also developed and applied for 4 representative shading patterns of the PV array.

According to the simulation results, under uniform solar irradiation conditions of the PV system, the third implementation of the Q-learning-based MPPT method would be chosen, because of its short learning period and its high MPPT efficiency. Under non-uniform solar irradiation conditions, the second implementation of the Q-learning-based method would be chosen, because despite the fact that it needs a longer learning period, it presents higher MPPT efficiency than the other three methods (i.e. the second implementation was the only Q-learning-based that located the MPP in shading pattern 9). The main advantage of the Q-learning-based MPPT methods is that they have the ability of learning how to detect the GMPP in specific shading patterns and then apply this knowledge in unknown intermediate conditions of these shading patterns. Thus, they are able to locate the GMPP of the unknown intermediate solar irradiation conditions in a few time-steps. Nevertheless, there is always the probability that the GMPP will not be detected. This is due to the fact that it is impossible to create a complete discrete state-space under all environmental conditions, resulting in the agent confusing different states as identical. Regarding the PSO-based MPPT method, its main advantage is that it always detects the GMPP, but its major disadvantage is that it cannot reduce its convergence time under similar environmental conditions that were reapplied in the past.

Future work includes the experimental evaluation of the Q-learning-based MPPT system that was designed in this thesis, in order to evaluate its performance under real operating conditions.

# 7.

## References

---

- [1] European Commission, “Green Paper - A 2030 framework for climate and energy policies”, *COM(2013) 169 Final*, pp. 1–16, 2013.
- [2] N. S. Lewis, “Toward cost-effective solar energy use”, *Science*, vol. 315, no. 5813, pp. 798-801, 2007.
- [3] L. Lakatos, G. Hevessy, and J. Kovács, “Advantages and disadvantages of solar energy and wind-power utilization”, *World Futur. J. Gen. Evol.*, vol. 67, no. 6, pp. 395–408, 2011.
- [4] J. Mohtasham, “Renewable Energies”, *Energy Procedia*, vol. 74, pp. 1289-1297, 2015.
- [5] E. O. Ogunniyi, and H. Pienaar, “Overview of battery energy storage system advancement for renewable (photovoltaic) energy applications”, *Proceedings of the 25th Conference on the Domestic Use of Energy*, pp. 233–239, 2017.
- [6] T. Tsoutsos, N. Frantzeskaki, and V. Gekas, “Environmental impacts from the solar energy technologies”, *Energy Policy*, vol. 33, no. 3, pp. 289–296, 2005.
- [7] X. Xia and J. Xia, “Evaluation of potential for developing renewable sources of energy to facilitate development in developing countries”, *Asia-Pacific Power and Energy Engineering Conference*, pp. 1–3, 2010.
- [8] V. Devabhaktuni, M. Alam, S. S. S. R. Depuru, R. C. Green, D. Nims, and C. Near, “Solar energy: Trends and enabling technologies”, *Renewable and Sustainable Energy Reviews*, vol. 19, pp. 555–564, 2013.
- [9] Y. Yang, K. A. Kim, F. Blaabjerg, and A. Sangwongwanich, “Advances in Grid-Connected Photovoltaic Power Conversion Systems”, *Woodhead publishing*, 2019.
- [10] N. Kishor, M. G. Villalva, S. R. Mohanty, and E. Ruppert, “Modeling of PV module with

- consideration of environmental factors”, *IEEE PES Innov. Smart Grid Technol. Conf.*, pp. 1–5, 2010.
- [11] R. Faranda, S. Leva, and V. Maugeri, “MPPT techniques for PV systems: energetic and cost comparison”, *IEEE Power Energy Soc. 2008 Gen. Meet.*, pp. 1–6, 2008.
  - [12] M. Raugei and P. Frankl, “Life cycle impacts and costs of photovoltaic systems: current state of the art and future outlooks”, *Energy*, vol. 34, no. 3, pp. 392–399, 2009.
  - [13] T. Ikegami, T. Maezono, F. Nakanishi, Y. Yamagata, and K. Ebihara, “Estimation of equivalent circuit parameters of PV module and its application to optimal operation of PV system”, *Sol. Energy Mater. Sol. Cells*, vol. 67, no. 1–4, pp. 389–395, 2001.
  - [14] S. Manias, “Power electronics”, *Simeon publishing*, 2017.
  - [15] S. Palanidoss, and T.V. Vishnu, “Experimental analysis of conventional buck and boost converter with integrated dual output converter”, *International Conference on Electrical, Electronics, Communication, Computer, and Optimization Techniques (ICEECCOT)* IEEE, pp. 323-329, 2017.
  - [16] G. Cipriani, G., V. Di Dio, D. La Manna, R. Miceli and G. R. Galluzzo, “Technical and economical comparison between different topologies of PV plant under mismatch effect”, *2014 Ninth International Conference on Ecological Vehicles and Renewable Energies (EVER)*, pp. 1-6, 2014.
  - [17] A. Ingle, D. I. Sangotra, R. B. Chadge, and P. Thorat, “Module configurations in photovoltaic system: a review”, *Mater. Today Proc.*, vol. 4, no. 14, pp. 12625–12629, 2017.
  - [18] G M. Jazayeri, S. Uysal, and K. Jazayeri, “A comparative study on different photovoltaic array topologies under partial shading conditions”, *Proc. IEEE Power Eng. Soc. Transm. Distrib. Conf.*, pp. 1–5, 2014.
  - [19] R. Sutton and A. Barto, “Reinforcement learning”, *MIT Press*, 2018.
  - [20] S. Pareek and R. Dahiya. "Output power maximization of partially shaded 4\* 4 PV field by altering its topology", *Energy Procedia*, vol. 54, pp. 116-126, 2014.
  - [21] A. N. A. Nasr, H. Saied, M. Z. Mostafa and T.M. Abdel-Moneim, “A survey of MPPT techniques of PV systems”, *IEEE Energytech*, pp. 1-17, 2012.
  - [22] T. P. Sahu and T. V. Dixit, “Modelling and analysis of perturb and observe and incremental conductance MPPT algorithm for PV array using Ćuk converter”, *IEEE Students’ Conf. Electr. Electron. Comput. Sci. (SCEECS 2014)*, vol. 4, no. 2, pp. 213–224, 2014.
  - [23] A. Safari and S. Mekhilef. "Simulation and hardware implementation of incremental conductance MPPT with direct control method using cuk converter.", *IEEE Transactions on Industrial Electronics*, vol. 58, no. 4, pp. 1154-1161, 2011.

- [24] Y. H. Liu, S. C. Huang, J. W. Huang, and W. C. Liang, "A particle swarm optimization-based maximum power point tracking algorithm for PV systems operating under partially shaded conditions", *IEEE Trans. Energy Convers.*, vol. 27, no. 4, pp. 1027–1035, 2012.
- [25] R. Sutton and A. Barto, "Reinforcement learning", *MIT Press*, 2018.
- [26] C. J. Watkins and P. Dayan, "Q-learning", *Machine learning*, vol. 8, no. 3-4, pp.279-292, 1992.
- [27] V. Mnih, "Human-level control through deep reinforcement learning", *Nature*, vol. 518, no. 7540, pp. 529–533, 2015.
- [28] N. Cesa-Bianchi, C. Gentile, G. Lugosi and G. Neu, "Boltzmann exploration done right", *Advances in Neural Information Processing Systems*, pp. 6284-6293, 2017.
- [29] C. Wei, Z. Zhang, W. Qiao, L. Qu, "Reinforcement-learning-based intelligent maximum power point tracking control for wind energy conversion systems" *IEEE Transactions on Industrial Electronics*, vol. 62, no. 10, pp. 6360-6370, 2015.