International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878 (Online), Volume-10 Issue-1, May 2021

Reinforcement Learning with Variable Fractional Order Approach for MPPT Control of PV Systems for the Real Operating Climatic Condition

Ashutosh Yadav, Archana

Abstract: ***The designing of maximum power point tracking (MPPT) controller is an integral part of the PV array system to ensure a continuous supply of energy in dynamic environmental conditions. The most challenging part here is to design a model that can track the maximum point irrespective of variations in environmental conditions and its parametric variations. The model designed in this article combats both the challenges as it is based on reinforcement learning with fractional-order. The application of Deep Q-learning makes the model parametric free and once the model trained can be implanted in a different scenario and run effectively. The amalgamation of fractional-order aids in the process by reducing the tracking time, oscillation around the peak, and total harmonic distortions. The model is well tested on standard conditions and has successfully achieved the desired results. Also, the proposed design is compared against various existing comparative algorithms to showcase its effectiveness in tracking time, THD, and maximum power. The design is also tested on the real data set, from the solcast where the test region is New Delhi, the capital of India. This region is taken as it faces one of extreme climatic condition and also being the second-highest most populated state faces an acute shortage of power throughout the year. The results have demonstrated that the model can produce maximum power even in the least solar irradiance conditions.***

Keywords: ***fractional order factor; Reinforcement learning; Deep Q learning network; Maximum power point tracking (MPPT); Photovoltaic system; Real operating conditions (ROC)***

I. INTRODUCTION

Energy is a vital source of survival and development for humans, it's demand has upshot over the years globally. This has raised concerns about environmental issues and aroused interest in other renewable energy sources. The amalgamation of renewable energy with the different sectors of energy has become an important factor for reducing the environmental effect and dependency on fossil fuels [1]. The main source of energy is the sun and this has been widely used due to its abundance especially in the desert regions [2]. The extraction of maximum energy from a PV generator is a trivial task.

Manuscript received on April 05, 2021.
Revised Manuscript received on April 08, 2021.
Manuscript published on May 30, 2021.
*Correspondence Author
Ashutosh Yadav*, Department of Electrical Engineering, G D Goenka University, Gurugram (Haryana), India.
Archana, Kurukshetra University, Gurugram (Haryana), India.

To achieve this Maximum Power Point Tracking (MPPT) algorithms are implemented to control a DC/DC converter to optimize the power transfer to the load. Also, they must continuously transmit maximum power available to the load irrespective of the conditions, this control problem is coined as Maximum Power Point Tracking (MPPT) [3].

MPPT is to control the change in output voltage as produced due to a change in the PV power. Moreover, PV systems endure a nonlinear relation between output current and voltage, this can result in remarkable losses. There have been many approaches from the most conventional to the traditional MPPT control techniques that have been proposed over time. Few of designed approaches are a modified incremental conduction maximum power point tracking (MPPT) algorithm with fuzzy controller [4], a software and hardware for hill-climbing (HC) modified fuzzy-logic (FL) maximum power point tracking (MPPT) control scheme used in photovoltaic (PV) power systems [5], a hybrid MPPT control that combines a modified P&O and an enhanced PSO [6]. There have been many approaches using heuristic methods like the novel Maximum Power Point Tracking (MPPT) method suitable for any application in which very fast-changing using Artificial Neural Network (ANN) [7], Most of these approaches are model-based and thus they try to modulate the different PV systems. However, obtaining a precision model of the PV systems and its parameters can be a cure for some issues, challenging when PV panels get interconnected in different configurations. This motivates the author to look for a model-free approach.

Reinforcement Learning (RL) is an important branch of artificial intelligence, machine learning, and robotics. This approach has the advantage of having self-learning ability and can interactively resolve much real-time application with trial and error approach. They have been widely applied in many applications like robots, games, industrial control, and other fields. It has bridged the gap between the traditional optimal controls, various nature-inspired algorithms, and the adaptive control theory [8]. RL learning is through experience where an agent continuously interacts with the surroundings through a set of actions. These actions are followed with rewards (positive feedback for right action) or punishment (negative feedback for the wrong action). Hence, iteratively RL algorithm learns how to behave in an optimal way which ensures the growth and survival of an RL agent [9]. The most important advantage of RL is that it does not require any model of the system or its dynamics to mimic it. This makes it insensitive to the model parameters.

The usage of RL to solve the problem of MPPT has gained immense popularity in recent years. A model based on RL was designed by [10] to bring variability in the wind speed energy conversion system. To solve the MPPT, especially for PV was carried out in [11].

Here, a simple RL-based MPPT method was conceived where the learning agent observed the environment conditions and then determined the perturbation to the PV array's operating voltage. This was translated in terms of actions and rewards. The rewards are encouraged to select a pair of (state, action) that would fetch maximum reward (positive value) and minimum (negative) if a wrong pair chosen. Through an iterative process of getting a reward, learning is done. Once the agent of RL MPPT has learned the strategy, it automatically will be able to adjust any change in PV voltage and obtain maximum power for that PV array design. Any RL system is governed through its learning process, which depends on the reward function. The reward function is normally classified into two categories that is either it is a dense function or a sparse. Sparse functions are easy to implement, but they provide less learning efficiency while the dense are difficult to design as the size of the state is too large in real domain applications.

In recent years, with the required robustness and applicability the demand for non-conventional MPP tracking methods has attracted the attention of researchers. The fractional-order approach is one of them. Fractional derivatives are commonly applied to enhance the performance of tracking in varying environmental conditions. This fractional-order control (FOC) method applies the principle of fractional calculus. One of the challenges faced in applying FOC is to find the best order of the system and its gain. Some of the recent work investigated the fractional-order control of MPPT. Fractional order based on Incremental Conductance (FOINC)-MPPT is designed in [12].Radial Movement Optimization (RMO) is used by the author to optimize the various parameters of the fractional integrator. The next is an extremum seeking control (ESC) controller design where a high pass filter is replaced with an integrator and low pass filter [13]. On a similar technique, the other design proposed is incremental conductance algorithm (VFOINC) combined with extenics variable step size (EVSS) control into the MPPT scheme for photovoltaic (PV) [14].There are also a few fractional-order models designed on sliding mode principal for real-time applications like induction generator (DFIG)-based wind energy conversion system [15].There is another design on real atmospheric conditions it is a perturbation observer-based fractional-order sliding-mode controller (POFOSMC) [16]. There have been many Artificial Neural Network-based designs that have been tested on the real data to showcase the effectiveness of the model [17, 18].

This leads the author to jot down few research gaps that are answered using this article:

1. The application of reinforcement learning in MPPT makes the problem simpler as it is a model-free approach and the sensitivity of model parameters on the model design is eliminated. Once the model is designed, it can be effectively used under different environmental conditions with remarkable accuracy without the need for remodeling and retuning the parameters.
2. Q-learning is a model-free approach in the RL algorithm. Here, the agent is guided for a further course of action based on experience. This algorithm estimates the state-action value function for a designed target policy that would select an action with the highest value. This situation is under control till the size of the search space is limited.
3. In the scenario where the search space is of high dimensionality then there is the need for storage of millions of records in a table in the program memory. This shortcoming is eliminated by using the concept of Deep Q-Learning where a neural network selects the action.
4. Fractional order control methods are normally non-integer based which get preferred over the integer due to the precision and more discrete space. Also, they provide a more flexible design and have better results in terms of robustness. Also, in the conventional methods of tracking MPP, the fixed amplitude perturbation is incorporated but due to a mismatch of change in amplitude and tracking speed step size there are oscillations produced, and that results in reduced efficiency. This drawback is effectively answered with the help of the fractional approach.
5. Normally, it has been observed that the convectional design fails to track the MPP in real rapidly changing environmental conditions. To test our design we have tested it on real data. This data is taken from the Solcast API toolkit. The region for analysis is New Delhi, the capital of India. The reason for taking this data is as this region is the capital of our country is one of the densely populated regions and faces an acute shortage of power supply around the year.

The research gap and above points motivate the author to design a novel Fractional-order MPPT control algorithm with DQN in reinforcement learning for a system design consisting of PV array under variable load for partial shading conditions. The tracking of MPP is governed by the agent of RL that is fed with seven different observation viz PV power, voltage, PV generated power diversion with PV's desired capability, integral of power diversion, coupling voltage, and divergence per unit time with reference coupling DC are contained for the learning process. Then, to make the process more efficient, the tracking speed of the agent is indirectly aided by the output of the fractional-order with error and derivative error terms. Also, it is tested on the real data for the region of New Delhi, the capital of India. The subsequent parts of this paper are arranged as follows: Section 2 Fractional-order for MPPT. Section 3 introduces the designed system. Then, Section 4 discusses the results achieved which are subdivided into two parts a) Comparison of the proposed algorithm with the other benchmark algorithms and b) Testing of the model on the real data set for the region under consideration, followed by Section 5 describes the prospects of the designed method along with the concluding remarks.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication

II. FRACTIONAL ORDER FOR MPPT

The functionality of PV array is generally is carried out in the changing atmospheric conditions, that get influenced by the parameters like irradiance and temperature. They constitute an overall impact on the PV array. Therefore, it is mandatory to consider them for modeling a PV module [19].

There are methods for parameter estimation out of which MPP is one of the methods. Here, $(V_m, I_m)$are the corresponding voltage and current of the maximum point of any PV curve. Taking the derivative of PV power P_{PV}concerning PV voltageV_{PV} when equated to zero one can estimate all the parameters for the variation in

MPP for a certain environmental condition.

$$\frac{dP_{PV}}{dV_{PV}}\Big|(V_m, I_m) = I_m + \frac{dI_{PV}}{dV_{PV}}\Big|(V_m, I_m) \times V_m \quad (1)$$

The solar radiation, temperature, and electrical conduction in the PV system are identical to the phenomena like diffusion, etc. This can be explained using the fractional-order approach [20].

Fractional-Order Differentiator

The designing of fractional order (FO) is commonly based on the definitions of fraction derivative of the Riemann–Liouville,and Grunwald–Letnikov(GL)[21,22]. All kinds of variations such as illumination, temperature, electric flow can express by FO using derivative-based formulation in the range of $0 < \alpha \leq 1$.

Fractional-Order Control MPPT Algorithm

Movement of operating point towards the peak level in the P-V curve is done to extract the maximum power output from the source. To reach the peak at a faster rate a fractional factor α (alpha) is applied this enables a variable domain for expansion and contraction for the designed model. In the case when the input variable is expanded then the model designed has a large tracking step size to reach MPP quickly. The Equation.1 when modified in the presence of α

$$\frac{d^\alpha P}{dV^\alpha} = \lim_{\Delta V \to 0} \frac{1}{\Delta V^\alpha} \times \sum_{t=0}^{\infty} \frac{-1^t\, \Gamma(\alpha+1)}{\Gamma(t+1)\Gamma(\alpha-t+1)}\; P(V - t\Delta V) \quad (2)$$

Here α > 0, $P(V - t\Delta V)$ is the output power of the PV array at the instant of time t. The changes in power and voltage get approximated using the fractional-order calculus. The approximated Equaion.2 is given as

$$\frac{d^\alpha P}{dV^\alpha} = \lim_{\Delta V \to 0} \frac{P - \alpha P_0}{(V - V_0)^\alpha} \quad (3)$$

The P and P_0 are the power measured at instant t and $t-1$ respectively. Similarly, is the voltage V and V_O. Theincremental power as approximated using the fractional-order is $d^\alpha P \cong P - \alpha P_0$ and the incremental change in the voltage is $dV^\alpha \cong (V - V_0)^\alpha$. The change in voltage and power in unit time is taken as an error $E_r\,(t)$. The change in power is $\delta\, P(t)$. These values are calculated in Equation.4 and Equation.5 respectively.

$$E_r\,(t) = \frac{P - P_O}{V - V_O} \quad (4)$$

$$\delta\, P(t) = P - P_O \quad (5)$$

The fractional factor α falls in the range of $0 < \alpha \leq 1$, this changes the input domain of the reward function designed for the reinforcement learning agent. When, α=1, the input the reward is directed by the other inputs while for the other values there will be an efficient change in the input pattern of the reward. The smaller value of α will expand the $E_r\,(t)$ that eventually leads less time taken to achieve MPP by increasing the step size of tracking. Also, in the case when MPP is nearby the $E_r\,(t)$is contracted thus step size is reduced and oscillations around the optimal point are prevented. The fast-tracking of MPP and a stable output around MPP will always minimize energy losses and improve the efficiency of the system.

III. PROPOSED IMPLEMENTATION OF FRACTIONAL DQN WITH RL

In this work, the author proposes a fractional order in the RL to cater to the problem of MPPT in PV arrays. RL effectively solves this problem without any parametric information about the dynamic parameters of the model. The algorithm aims to depict the system analogy based on the DQN (Deep Q-Network) model. The complete model designed by the author is shown in Figure.1 and the fractional block internal design is shown in Figure.5.

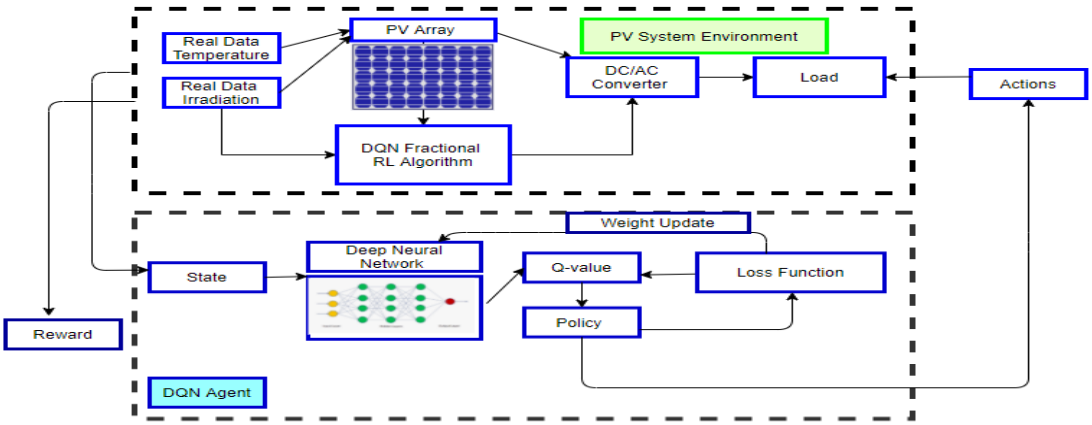

Fig.1. The DQN with the fractional design of MPPT control for the PV arrays model.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication

3.1 RL formulation for MPPT Control

Reinforcement learning is comprising of four main components as already discussed in Section.2. These are state-space$X$, reward r, transition probability p, and the action space U.In the MPPT control problem, any action taken in a speck of time corresponds to any manipulation in the variable V_{PV}.

Any interaction of the agent with the environment leads to the learning process and takes an action $u_t \in U$ through this system evolves from the state $x_t \in X$ to the next state action x_{t+1} then agent achieves feedback known as a reward that quantifies the quality of action or the step chosen by the agent. Therefore, the reward is acting as a "clue" to target the achievable goal or optimal solution. The RL method intends to search for an optimal policy π that satisfies.

$$J^* = \max_\pi J_\pi = \max_\pi E_\pi \{r_t \,|x_t = x\}$$

Here J_π corresponds to the total expected reward against the policy π. Under the assumption that forgiven policy π, the expected cumulative reward $V^\pi(x)$ or the value function for a certain time interval is the function of x^π and is defined as $x^\pi = \{x_t\}_{t=1}^{t=n}$ are the state values $k^\pi = \{k_t\}_{t=1}^{t=n}$ are the sequences of the agent taken based on the actions.

i. State Space

The various states in any MPPT control are designed based on the knowledge of the movement of MPP on the PV curve under varying environmental conditions. The methodology of RL is governed by the current, power, and voltage. The state-space includes $[V_{PV}, I_{PV}, P_{PV}, \Delta P_{PV}, \int \Delta P_{PV}, \alpha e(t)]$ and the coupling point ΔV_{DC}. Here, $\int \Delta V_{DC}$ controls the selection of duty cycle in the interval of $[0,1]$. ΔP_{PV}defines the divergence of PV power from the desired to the generation capacity and $\Delta V_{DC}$is the difference between the reference coupling voltage and measured V_{DC}.

ii. Action Space

Action space normally applied to the MPPT problem is discrete. This guarantees a high precision level and acts as a powerful learning technique that makes it a computationally efficient method. In the action of the RLMPPT agent, there is a defined duty cycle. Through a series of actions, the duty cycleD_{int} is selected in the range $D_{int} = (0,1]$. This results in a matrix with 100 possible actions.

iii. Reward

The reward space has been carefully crafted using four values $[\Delta V_{DC}, \Delta P_{PV}, e(t), exb, \Delta e(t)]$. The threshold is imposed on each input except the exb which denotes the excess bound conditions. The violations in the input values cause the generation of negative reward while the permissible value of input fetches a positive reward value. The internal design of getting $e(t)$ and $\Delta e(t)$ is shown in Figure.2 below.

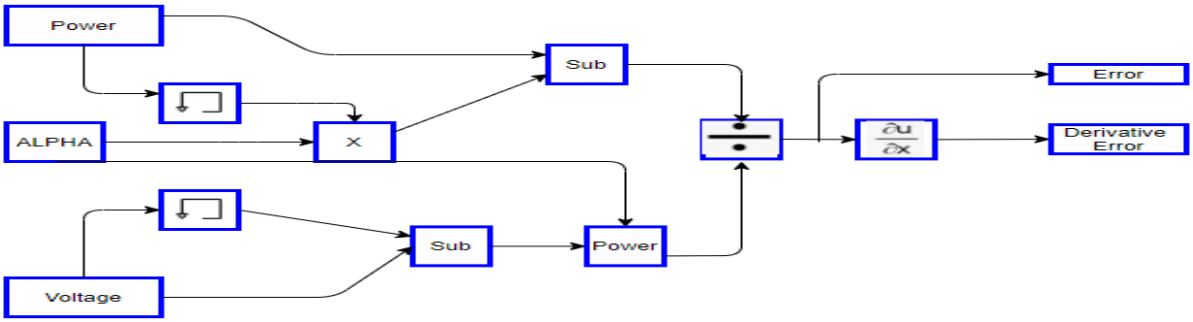

Fig.2. Schematic design of the Fractional Block.

3.2 Deep RL tracking Problem

Designing an RL agent is an especially important and crucial step to attain the required accuracy and robustness. The biggest challenge faced here is how to deal with the continuous space. The right balance between heavy discretion and insufficient discretion is solved. This limits the functionality of RL to a narrow zone that hampers its ability for dynamic environments. This leads to the method is called Deep Q-learning that has the characteristic of being off-policy and model-free. The layered structure of DQN is shown in the Figure. 3. It consists of the stacking of various layers like a fully connected layer and Relu layers. The other necessary parameters taken for designing the network are tabulated in Table. 1 below.

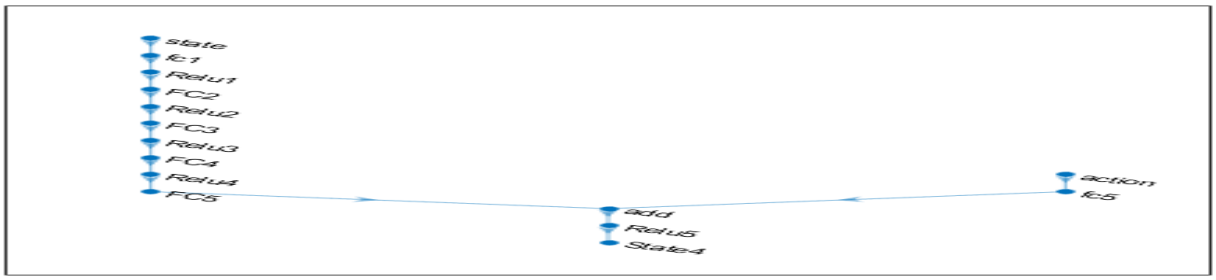

Fig.3. The internal structure of DQN employed for fractional MPPT control.

Retrieval Number: 100.1/ijrte.A56310510121
DOI: 10.35940/ijrte.A5631.0510121

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication

Table 1. The parameter setting for the fractional DQN network.

Parameters	Values
Various Inputs	7
Relu Layers	5
Fully connected Layers	7
Number of Hidden neurons	100
Designed Discount Factor (γ)	0.99
Targeted Learning rate (α)	0.01
Maximum number of Episodes	300
Maximum step size	10

In this part, the proposed algorithm for Deep RL control is discussed in Algorithm.1. RL agent is DQN based design. Thus, the agent consists of neural networks that optimize the policy π and the network Q. The algorithm starts by initializing the parameters for generating power through a connection of PV arrays. Here, a factor α is used as fractional to pass the error and the derivative error terms as designed in the above section. Then, these are given in the state, action, and reward space for the network as discussed in the methodology section. The designed DQN as depicted above is loaded. Then the training of the network is carried out using Q-learning. The network gets loaded. The main loop starts from Line.14. Loop is made to run for (M)several episodes. The initial value of the duty cycle $D_{int}$is provided, followed by all the initial states x_0, observations, and the course of actions u_t. Line 17 leads to the inner loop which is executed for several time steps (T) in each episode. Here, the agent picks an action from the pool of existing actions hanging on the environment (line number 18). As a result, the new state is generated $x_{t+1,}$and reward is provided r based on the action taken by the agent (Line number 20). Finally, all three things are kept aside in a buffer (Line 21). In the next, it shows that if enough state transitions are stored then training of the agent is started. A small number (N) of random transitions are extracted from the buffer. These values update the network uses Q-learning. The total number of training states are finished for target networks and updated. The algorithm ends at line number 19 where both the networks are trained and returned in the buffer. The results achieved through this whole methodology are discussed in the section below.

Algorithm 1: Fractional Order DQN MPPT

1. Connect solar PV array SPR-305E-WHT-D (330 sun power)
2. Set the short-circuited current $I_{SC} = 5.96\,A$ and Open circuited Voltage $V_{oc} = 64.2\,V$
3. Estimate the peak power for $N_S = 1$ (PV's in series) and $N_p = 12$ (PV's in parallel) using $P_{npp} = (N_S \times V_{mpp}) \times (N_P \times I_{mpp})$
4. Select the DC-link voltage
5. Initialize the state, action, and reward for the RL agent.
6. Initialize the duty cycle $D_{int} = 0.5$, $exb = -100$
7. Initialize the $0 < \alpha \leq 1$ fractional factor
8. Calculate the $e(t)$ and $\Delta e(t)$ based on the fractional value of α
9. State-space X=$[V_{PV}, I_{PV}, P_{PV}, \Delta P_{PV}, \int \Delta P_{PV}, \alpha e(t), \Delta V_{DC}]$+ Action space U= (0,1]
10. Reward functions calculate using $[\Delta V_{DC}, \Delta P_{PV}, \alpha e(t), exb, \alpha \Delta e(t)]$
11. Provides these values to the network DQN
12. Initialize / Load Q and replay buffer R. β learning rate and γ discount factor.
13. **for** $j = 1$ *to* M **do**
14. Get initial state x_0
15. **for** $t = 1$ *to* T **do**
16. Select action u_t from the set defined
17. **Execute** the action u_t
18. Get a new state x_{t+1} and reward r
19. Store the transition $(x_t, u_{t,} u_{t+1,} u_{t+1})$ in R
20. IF $|R| > N$
21. Update the network using Q-learning:
22. $Q^{new}(x_t, u_t) \leftarrow Q(x_t, u_t) + \beta(r_t + \gamma \max_a Q(x_{t+1}, a)) - Q(x_t, u_t)$
23. Update the network
24. **end if**
25. Set $x_t = x_{t+1}$
26. **end for**
27. **end for**
28. $Q(.,.|w), r$

IV. RESULTS AND DISCUSSION

The proposed fractional MPPT scheme is tested when powered from the 330-Watt sun module. The module is tested under the standard conditions also tested for the real data set. The parameters employed in PV array simulation are jotted down in Table.2. The performance of the PV cell taken for Simulink modelling is showcased in Figure.4. This curve is plotted for the standard temperature of 25 ℃ under different irradiance conditions.The efficiency of the cell is high as the solar irradiance being as minimum as 200 w/m^2 produces a power of 0.2 kw and it improves with the increase in sun radiations. The design of MPPT is done on MATLAB 2020a with processor Intel ® Core TM i3, 1.98 GHz. The initial duty cycle is $D_{int} = 0.5$. A Deep Q algorithm is used to train the RL agent. The limits are set as(0,1]. The action space for the RL agent is designed in interval over the step size of $(0{:}\,0.01{:}\,1]$ resulting in a matrix size of 100×1. The learning rate is designed as 0.01with regularization L2 having values 0.0001 and a mini-batch size of $N = 64$. The model is trained for $M = 300$ episodes with a step size of $\frac{T}{T_s}$, T is the total simulation time incurred and T_s is the sample time. The deigned model takes the value of α (the fractional factor) over a range of [0,1].

Table 2. Parameter for the PV system utilized in the model.

Sun Power SPR-305E-WHT-D	
No. of PV array series connected	12
No. of PV array parallel connected	1
I_{SC} per PV array	5.96 A
V_{OC} per PV array	64.2 V
Max power P_{PV} at standard operating conditions	305 W
Reference DC coupling voltage $V_{DC_{ref}}$	400 V
Temperature Range $Temp$	0-500%/deg.C
Irradiation Range I_r	680-1000
The voltage at the maximum powerpointV_{PV}	54.7 V
Current at the maximum power pointI_{PV}	5.58 A

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication

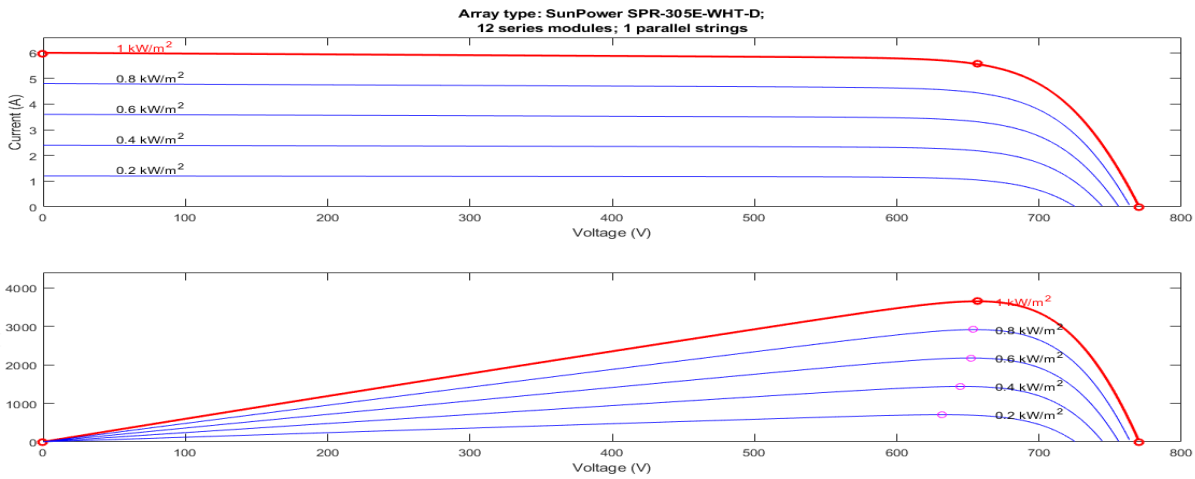

Fig.4. Standard current and voltage of the PV system under different Irradiance at 25 ℃.

The designed model is RL based, the Figure.5 below showcases the training of the DQN learning for the RL model over an episode of 300. This training is done using the standard conditions of temperature and solar irradiance. The model already gets converged around 600 iterations. This illustrates the ability of the designed model to get quickly adapt to the changes in the environmental conditions.

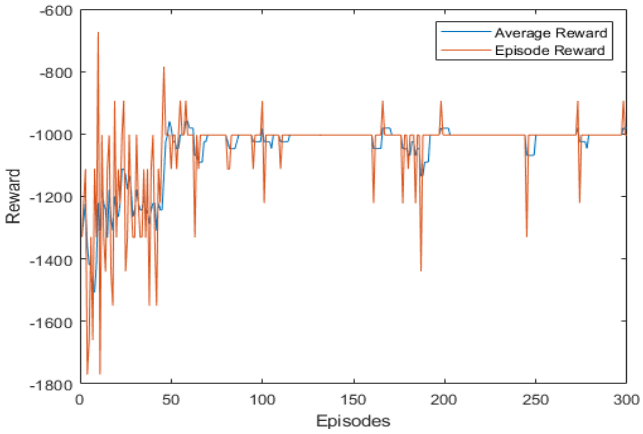

Fig.5. Training Curve of proposed Fractional Order DQN MPPT Control Model.

This completes the brief overview of the parameters of the model designed and the platform. Now, the results section is bifurcated into two parts: The deigned model is compared against the various benchmark algorithm to showcases it is robustness and efficiency and in the last section, the model is validated on the real dataset to prove that the model once trained is efficient to operate different environmental conditions.

Comparative analysis with other Benchmark Approaches

The proposed fractional-order method has been juxtaposed with other comparative benchmark work especially trying to control the problem of MPPT with RL. The methods compared with are perturbation and Observation (P&O) [23], a fuzzy logic controller (FLC) [24], and fractional order fuzzy logic controller (FOFLC) [25]. Here, with help of a controlling parameter fraction of change in the power is reverted into the circuit. This factor α is varied over the range as stated in the algorithm. P&O is one of the most widely used models for MPPT control. The change in duty cycle and power along with voltage get tested with every cycle of perturbation. In the occurrence of violations condition, a reinitialization is done. The comparison in terms of maximum power and voltage are shown in Figure. 6.

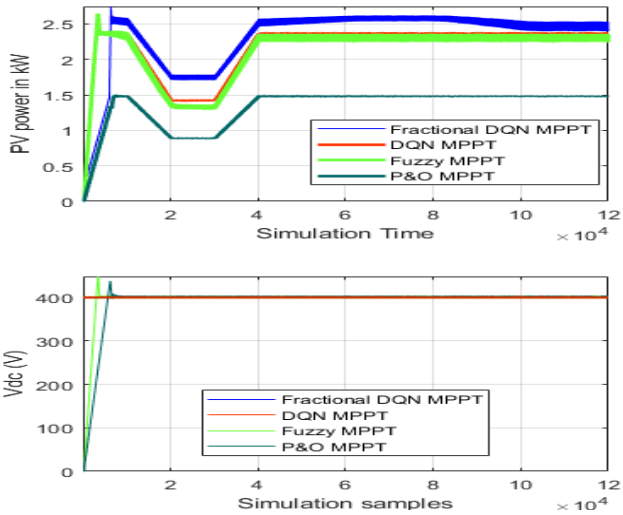

Fig.6. Comparative Analysis of different algorithms for PV power and V_{DC}.

Total harmonic distortion (or THD) is the amount of harmonic distortion present in any signal. It is defined as the ratio of the sum of all the harmonics present to the fundamental frequency. The sun rays incident are fluctuating in nature that eventually generates harmonics in the power that are undesirable for system performance. It is one of the serious concerns in the PV systems that are integrated into grid systems. The designed model with fractional control has been able to reduce THD component under the standard conditions. A comparison of the various methods based on THD has been shown in Table.3. This term is calculated here for all the comparative algorithms based on fundamental frequency and the first 5 harmonics encountered when a modified periodogram (Kaiser Window) is used. It can be observed that the proposed FODQN has the least THD component and also settles in very few seconds.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication

Table 3 Comparative results of various methods against THD and Settling Time.

Parameters	Proposed FODQN	DQNFR	FOFLC	FLC	P&O
Power (kw/m^2)	2.7	2.8	2.26	2.28	1.5
THD	-15.48	-26.1506	-6.6345	-5.250	2.23
Settling Time (s)	0.121	0.231	0.1847	0.1847	0.3225
Temperature : 0-50 ℃ Irradiance ($\frac{W}{m^2}$) = 680-1000					

Also, the proposed algorithm is compared with other two designed algorithms in terms of various output characteristics at MPPT under various solar irradiance conditions and keeping the temperature at a standard condition of 25 ℃. This data has been taken for comparison purposes from [26]. To have uniformity in comparison the values have been normalized using the standard characteristic of the PV array utilized by the designs. Through this analysis, the author could conclude that the design outperforms in terms of maximum power in presence of distinct environmental conditions.

Table 4 Comparative results of various output characteristics at maximum power point tracking (MPPT) for different solar irradiation at standard temperature conditions.

MPPT	Irradiations (w/m^2)				
	200	400	600	800	1000
SST [27]					
I_{PV}	0.305	0.600	0.896	0.924	0.9854
V_{PV}	0.640	0.705	0.715	0.747	0.783
P_{PV}	0.195	0.423	0.641	0.840	0.941
FGSMC [28]					
I_{PV}	0.475	0.9388	0.9425	0.9548	0.9845
V_{PV}	0.6288	0.7006	0.6791	0.7662	0.8067
P_{PV}	0.1599	0.3518	0.5431	0.7743	0.8659
Proposed FODQN					
I_{PV}	0.2249	0.4342	0.6502	0.8647	1.0776
V_{PV}	0.3034	0.4596	0.5244	0.6998	0.8318
P_{PV}	0.2127	0.5188	0.7309	0.8464	0.8483

In this article, the author has also compared the proposed work with other fuzzy methods on basis of the rise time of Voltage V_{dc}, robustness, and deviation. The method is compared to FOFLC, adaptive FLC, Fractional Order (FO), and IncCond Methods (ICM). The proposed method has a smooth transition, with a good tracking speed and steady-state stability. Also, a comparative settling time, a model-free approach have begged good results under varying temperatures and irritation condition. These all make the proposed method a robust design. The deviation defines here the difference that is between V_{dc} and the V_{ref}. There by making it more vulnerable as compared to all the techniques existing in the literature as tabulated in Table.5. These values are taken for comparative analysis. These are on the simulations done by the respective authors in the state-of-the-art schemes. They have adopted the different PV models, so the linguistic approach is used to compare their work with our results

Table 5 Performance Comparison of FODQN MPPT Techniques with other fractional-order designs.

Work Proposed	Rise Time	Robustness	Deviation
Fixed FLC [29]	0.2	Good	Good
Adaptive FLC [30]	0.15	Good	Good
ICM [31]	0.16	Bad	Bad
FO [32]	0.2	Bad	Bad
FOFLC [33]	0.03	Excellent	Excellent
Proposed FODQN	0.221	Excellent	Good
Note: Bad, Good and Excellent are linguistic to compare the methods			

The tracking time is also an important parameter to effectively track the MPP. This has led the author to compare the tracking time of the proposed method with FFGSMC in Table.6 shows the tracking time of the maximum power point for different irradiation.

Table 6 Tracking time of the maximum power point for various solar irradiations.

Irradiations (w/m^2)	200	400	600	800	1000
Proposed FODQN	0.0081	0.0075	0.0062	0.0042	0.0039
FFGSMC [23]	0.0093	0.0078	0.0064	0.0051	0.0046

Testing on Real Data Set of New Delhi

The validation of the model is carried out on the real data. This data is collected from the website Solcast. It is an API toolkit that produces real-time, historical, and forecast estimates of the available solar radiation resources around the globe. Also, it provides actual and forecasts solar irradiance and power data, globally, using satellites and surface measurements. In this article, the region for study chosen is Delhi, officially known as the National Capital Territory of Delhi (NCT), which is a city and a union territory of India. Below, in Figure. 7 the map of India along with Delhi is shown for reference.

The latitude of Delhi is 28.644800, and the longitude is 77.216721. It is in India with the Global Position System (GPS) coordinates of 28° 38' 41.2800" N and 77° 13' 0.1956" E. The NCT covers an area of 1,484 square kilo meters (573 sq mi). According to the 2011 census, Delhi's city proper population was over 11 million, the second-highest in India. The climatic conditions of these regions vary over a range from 2 to 47 °C (35.6 to 116.6 °F), with the lowest and highest temperatures ever recorded being −2.2 and 48.4 °C (28.0 and 119.1 °F), respectively. Normally, it has a dry winter with humid subtropical climatic conditions. The annual mean temperature is 25 °C (77 °F); monthly mean temperatures range from 13 to 32 °C (55 to 90 °F). The different seasons round the year in Delhi are tabulated in Table.7 below.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication

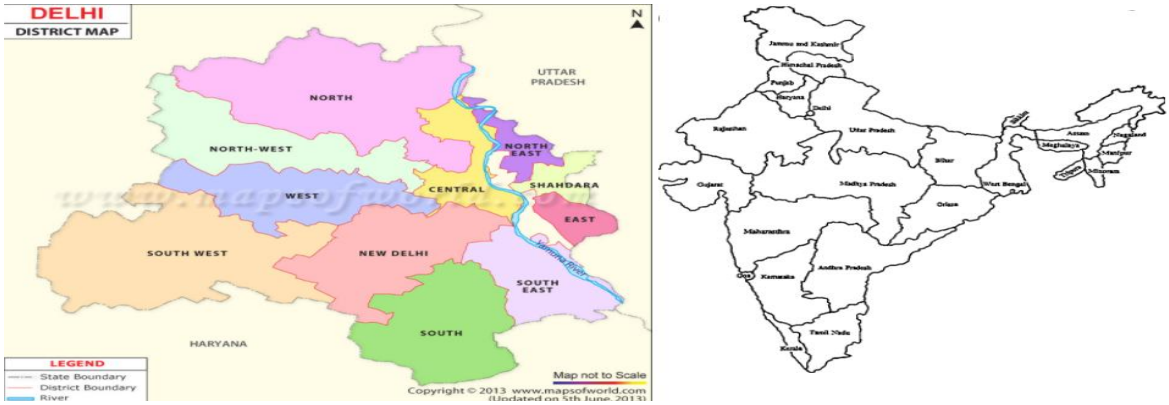

Fig.7. Outline of Map of India and New Delhi.

Table 7 Different Seasons witnessed in New Delhi round the year.

	Month	Climate
Spring	February to March	Sunny and pleasant.
Summer	April to June	Hot
Monsoon	July to Mid-September	Wet, hot, and humid
Autumn	September end to November	Pleasant

In this study, the parameter taken for analysis is Air Temperature (TEMP, °C): The air temperature (2 meters above ground level) and Global Horizontal Irradiance (GHI, W/m2): The total irradiance received on a horizontal surface. It is the sum of the horizontal components of direct (beam) and diffuse irradiance. The mean of temperature and irradiance for the summer season is showcased in Figure. 8 and Figure.9 respectively.

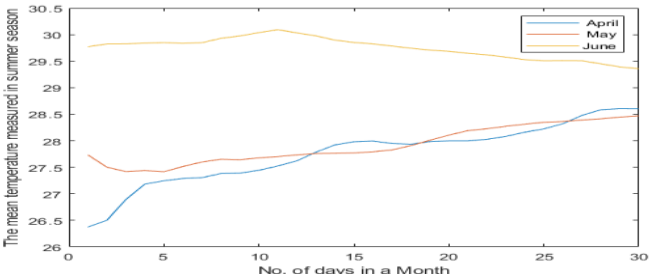

Fig.8. The mean temperature was recorded for New Delhi in the year 2019 during the summer season.

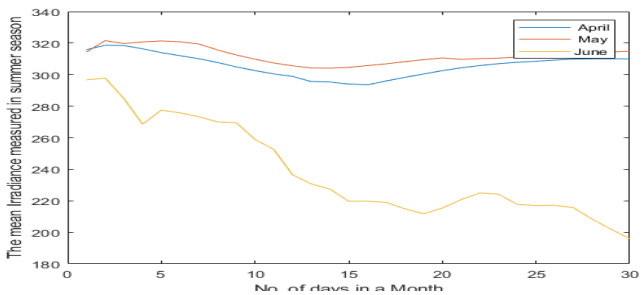

Fig.9. The mean irradiance was recorded for New Delhi in the year 2019 during the summer season.

The model was tested using this data. The data was pre-processed to suit the input conditions of the model. The data for the year 2019 was taken and the output PV power and the other standard values were measured for a day. The data taken for analysis is shown in Figure.10. The model after being trained and tested on the standard conditions. Then, the model is further tested on this real data condition of different temperatures and irradiance. The testing purposed only includes the applying of the pre-processed data and observing the output. To, validate the model author has tested the model for both the extreme season conditions prevailing in New Delhi. As the winter and summer seasons witness extreme conditions these are chosen for testing purposes.

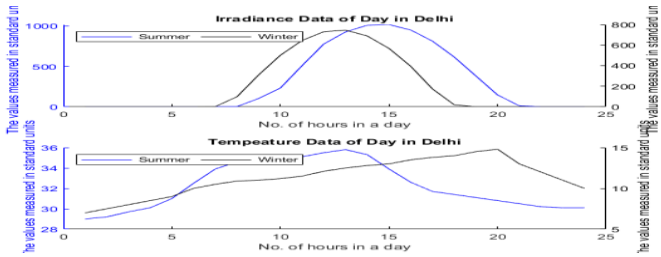

Fig.10. Data taken for testing the model designed (A day in the year 2019 for summer and winter).

The results showcase the utility of the designed model under various environmental conditions. Also, the model simulated in the ideal condition can be effectively utilized in the field conditions without any change in the effective model design. One more advantage of this design is that once the model is designed it can be utilized without retraining it.

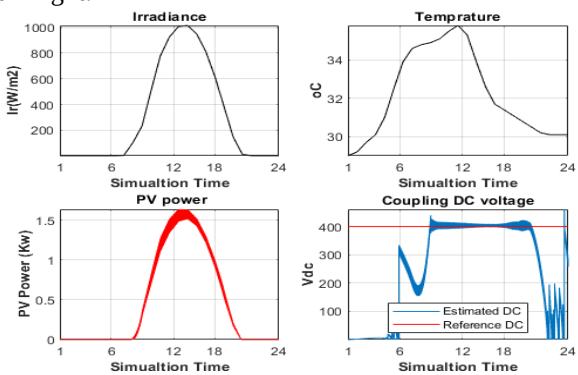

Fig.11.The PV power and DC coupling voltage for testing under summer data.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication

The simulation results for summer and winter testing under different environmental condition is shown in Figure. 11 and Figure. 12 respectively.

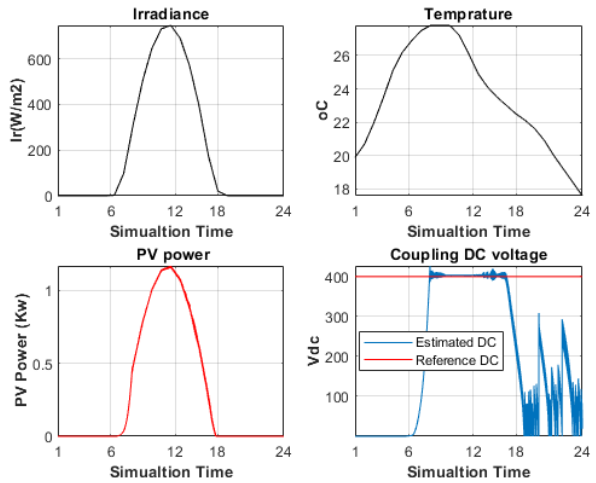

Fig.12.The PV power and DC coupling voltage for testing under winter data.

V. CONCLUSION

In this article, a novel model for tracking maximum power point is designed using reinforcement learning and the fractional-order concept. The introduction of reinforcement learning is done to make the model independent of parametric variations in the design to adapt the environmental effects and also the training is done using a Deep Q-learning algorithm that makes the design a model-free approach. The added advantage of this learning is that once the network trained can withstand any variation still effectively track the MPP. The appending of fractional-order in the proposed design is to reduce the tracking time for reaching the peak level, maintain a steady-state output without oscillation around the MPP, and the reductions THD component in the output due to variation in the solar irradiation that eventually affect the PV array performance. The testing of design is carried out in two phases where it is has been compared with various benchmark existing algorithms and also it has been tested on the real data set to capture the essence of the design under various environmental conditions. The lowest THD component and tracking time with maximum power output at MPP are achieved with FODQN as compared to various algorithms. The comparative analysis confirmed that model is effective in harvesting maximum energy at different solar irradiation conditions. The application of the proposed model on the real data set ensures that this model can be easily implemented in any area without any parametric change in the model and also produces maximum power even under the lowest solar irradiance in extreme winter conditions where the presence of solar energy is lowest recorded.

REFERENCES

1. Adams, Samuel, and Alex O. Acheampong. "Reducing carbon emissions: The role of renewable energy and democracy." Journal of Cleaner Production 240 (2019): 118245.
2. Bouraiou, Ahmed, et al. "Status of renewable energy potential and utilization in Algeria." Journal of Cleaner Production 246 (2020): 119011.
3. Moharm, State of the art in big data applications in microgrid: A review, Adv.Eng.Informatics.42(2019)100945,http://dx.doi.org/10.1016/j.aei.2019.100945.
4. L. Panwar, S.C. Kaushik, S. Kothari, Role of renewable energy sources in environmental protection: A review, Renew. Sustain. EnergyRev.15(2011)1513–1524, http://dx.doi.org/10.1016/j.rser.2010.11.037
5. Punitha, K., D. Devaraj, and S. Sakthivel. Development and analysis of adaptive fuzzy controllers for photovoltaic system under varying atmospheric and partial shading condition., Applied Soft Computing 13.11 (2013): 4320-4332.
6. Liu, Liqun, et al. "Simulation and hardware implementation of a hill-climbing modified fuzzy-logic for maximum power point tracking with direct control method using a boost converter." Journal of Vibration and Control 21.2 (2015): 335-342.
7. Kermadi, Mostefa, et al. "An effective hybrid maximum power point tracker of photovoltaic arrays for complex partial shading conditions." IEEE Transactions on Industrial Electronics 66.9 (2018): 6990-7000.
8. Rizzo, Santi Agatino, and Giacomo Scelba. "ANN-based MPPT method for rapidly variable shading conditions." Applied Energy 145 (2015): 124-132.
9. Khan, Said G., et al. "Reinforcement learning and optimal adaptive control: An overview and implementation examples." Annual reviews in control 36.1 (2012): 42-59.
10. Zhu, Zhuangdi, Kaixiang Lin, and Jiayu Zhou. "Transfer Learning in Deep Reinforcement Learning: A Survey." arXiv preprint arXiv:2009.07888 (2020).
11. Wei, Chun, et al. "Reinforcement-learning-based intelligent maximum power point tracking control for wind energy conversion systems." IEEE Transactions on Industrial Electronics 62.10 (2015): 6360-6370.
12. Hsu, Roy Chaoming, et al. "A reinforcement learning-based maximum power point tracking method for the photovoltaic array." International Journal of Photoenergy 2015 (2015).
13. Al-Dhaifallah, Mujahed, et al. "Optimal parameter design of fractional order control based INC-MPPT for the PV system." Solar Energy 159 (2018): 650-664.
14. Liu, Jianxin, Tiebiao Zhao, and YangQuan Chen. "Maximum power point tracking with fractional order high pass filter for proton exchange membrane fuel cell." IEEE/CAA Journal of Automatica Sinica 4.1 (2017): 70-79.
15. Yu, Kuo-Nan, Chih-Kang Liao, and Her-Terng Yau. "A new fractional-order based intelligent maximum power point tracking control algorithm for photovoltaic power systems." International Journal of Photoenergy 2015 (2015).
16. Ebrahimkhani, Sadegh. "Robust fractional-order sliding mode control of doubly-fed induction generator (DFIG)-based wind turbines." ISA transactions 63 (2016): 343-354.
17. Yang, Bo, et al. "Perturbation observer based fractional-order sliding-mode controller for MPPT of grid-connected PV inverters: Design and real-time implementation." Control Engineering Practice 79 (2018): 105-125.
18. Bouakkaz, Mohammed Salah, et al. "ANN based MPPT Algorithm Design using Real Operating Climatic Condition." 2020 2nd International Conference on Mathematics and Information Technology (ICMIT). IEEE, 2020.
19. Kumar, Pankaj, Nitai Pal, and Himanshu Sharma. "Performance analysis and evaluation of 10 kWp solar photovoltaic array for remote islands of Andaman and Nicobar." Sustainable Energy Technologies and Assessments 42 (2020): 100889.
20. Shongwe, Samkeliso, and Moin Hanif. "Comparative analysis of different single-diode PV modeling methods." IEEE Journal of photovoltaics 5.3 (2015): 938-946.
21. Tang, Shiqing, et al. "An enhanced MPPT method combining fractional-order and fuzzy logic control." IEEE Journal of Photovoltaics 7.2 (2017): 640-650.
22. Ma, Chengbin, and Yoichi Hori. "Fractional-order control: Theory and applications in motion control [past and present]." IEEE Industrial Electronics Magazine 1.4 (2007): 6-16.
23. Bonilla, B., et al. "Fractional differential equations as alternative models to nonlinear differential equations." Applied Mathematics and Computation 187.1 (2007): 79-88.
24. Ahmed, Jubaer, and Zainal Salam. "An improved perturb and observe (P&O) maximum power point tracking (MPPT) algorithm for higher efficiency." Applied Energy 150 (2015): 97-108.

25. Al-Majidi, Sadeq D., Maysam F. Abbod, and Hamed S. Al-Raweshidy. "A novel maximum power point tracking technique based on fuzzy logic for photovoltaic systems." International Journal of Hydrogen Energy 43.31 (2018): 14158-14171.
26. Kumar, Anupam, and Vijay Kumar. "Performance analysis of optimal hybrid novel interval type-2 fractional-order fuzzy logic controllers for fractional-order systems." Expert Systems with Applications 93 (2018): 435-455.
27. Babaei, S. Milad, Meisam Yahyazadeh, and Hassan Fatehi Marj. "Novel MPPT for Linear-Rotational Sun-Tracking System Using Fractional Fuzzy Grey-Based Sliding Mode Control." Iranian Journal of Science and Technology, Transactions of Electrical Engineering (2020): 1-23.
28. Fathabadi, Hassan. "Comparative study between two novel sensorless and sensor based dual-axis solar trackers." Solar Energy 138 (2016): 67-76.
29. Chu, Chen-Chi, and Chieh-Li Chen. "Robust maximum power point tracking method for photovoltaic cells: A sliding mode control approach." Solar Energy 83.8 (2009): 1370-1378.
30. Eltamaly, Ali M. "Performance of smart maximum power point tracker under partial shading conditions of photovoltaic systems." Journal of Renewable and Sustainable Energy 7.4 (2015): 043141.
31. Guenounou, Ouahib, Boutaib Dahhou, and Ferhat Chabour. "Adaptive fuzzy controller based MPPT for photovoltaic systems." Energy Conversion and Management 78 (2014): 843-850.
32. Lin, Chia-Hung, et al. "Maximum photovoltaic power tracking for the PV array using the fractional-order incremental conductance method." Applied Energy 88.12 (2011): 4840-4847.
33. Yu, Kuo-Nan, Her-Terng Yau, and Chi-Kang Liao. "Development of a fractional-order chaos synchronization dynamic error detector for maximum power point tracking of photovoltaic power systems." applied sciences 5.4 (2015): 1117-1133.
34. Tang, Shiqing, et al. "An enhanced MPPT method combining fractional-order and fuzzy logic control." IEEE Journal of Photovoltaics 7.2 (2017): 640-650.

AUTHORS PROFILE

Ashutosh Yadav, is pursuing Ph.D. from Department of Electrical Engineering, from G D Goenka University, Gurugram. He completed his B.Tech in Electrical Engineering from YMCA University, Faridabad and M Tech from MDU, Rohtak.

Dr. Archana, has completed B.Tech from Kurukshetra university one of the top leading university and also completed masters from PEC university of technology. She obtained Phd. From MRIIRS (Faridabad) Haryana. She also work with different engineering and research institute of Delhi NCR and Pune. She has 13 years of teaching and research experience with more than 15 published papers. The teaching and research experience is more than 10 years.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication