Enhancement of Power Quality of Three-Phase GC Solar Photovoltaics

Sukhbir Singh

sukhbirsinghdtu26051972@gmail.com

Delhi Technological University

J N Rai

Delhi Technological University

Research Article

Keywords: Three-phase grid connected, PV, inter-harmonics, DC offset, EDRL, COA-Fuzzified PLL

Posted Date: August 31st, 2023

DOI: https://doi.org/10.21203/rs.3.rs-3280839/v1

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at Electrical Engineering on March 7th, 2024.
See the published version at https://doi.org/10.1007/s00202-024-02304-z.

Enhancement of Power Quality of Three-Phase GC Solar Photovoltaics

[1]Sukhbir Singh
Department of Electrical Engineering, Delhi Technological University, New Delhi, India
Email: sss.singh149@gmail.com

[2]J N Rai
Department of Electrical Engineering, Delhi Technological University, New Delhi, India
Email: Jnraiphd1968@gmail.com

Abstract:

The proliferation of grid-connected photovoltaic (PV) systems has generated considerable apprehension among power system operators due to worries about electricity quality, leading to the implementation of increasingly strict standards and regulations. Inter-harmonics and DC offset have emerged as prominent power quality issues in grid-connected photovoltaic (PV) systems, constituting significant obstacles. This article provides a thorough examination of the methods used to improve the performance of a three-phase grid-connected photovoltaic (PV) system, with a specific focus on mitigating inter-harmonics and DC offset. The presence of inter-harmonics and DC offset may have a substantial negative impact on the overall performance of a system, resulting in compromised power quality and diminished energy extraction capabilities. In order to address these challenges, a method known as ensembled Deep Reinforcement learning (EDRL) Maximum electricity Point Tracking (MPPT) is used to optimize the extraction of electricity from the photovoltaic (PV) array. Furthermore, the integration of a Coati Optimization Algorithm (COA) with a fuzzified Phase-Locked Loop (PLL) synchronization mechanism is used to ensure precise synchronization with the grid. The EDRL MPPT approach demonstrates a proficient ability to accurately monitor and follow the maximum power point of the photovoltaic (PV) array. This is achieved by using a reward system that is based on the lowest overall harmonic distortion in the grid current. The COA (Centralized Optimization Algorithm) is used to effectively tune the hyperparameters of the fuzzy system. The primary objective of this optimization process is to reduce the DC offset, hence ensuring a steady and precise synchronization between the fuzzy system and the grid. The efficacy of the proposed system is assessed by means of comprehensive simulations and experimental validation. The findings of this study provide evidence supporting the efficacy of the Enhanced Distributed Reactive Load Maximum Power Point Tracking (EDRL MPPT) approach in optimizing power extraction and reducing the impact of inter-harmonics. The COA-fuzzified-PLL synchronization system is designed to provide precise grid synchronization while mitigating the adverse effects of a 2.89% total harmonic distortion (THD) in grid current, particularly the influence of direct current (DC) offset. The integration of many approaches presents notable improvements in terms of power quality, energy extraction efficiency, and system stability.

Keywords: Three-phase grid connected, PV, inter-harmonics, DC offset, EDRL, COA-Fuzzified PLL,

Abbreviations

MPPT	Maximum Power Point Tracking
EDRL	ensembled Deep Reinforcement learning
COA	Coati Optimization Algorithm
PV	photovoltaic
PLL	Phase-Locked Loop
DC	Direct Current
SRF	synchronous reference frame
SOGI	second-order generalized integrator
ROGI	Reduced-Order Generalized Integrator
TOGI	Third Order Generalised Integrator
P & O	perturb and observe
THD	Total Harmonic Distortion
PWM	Pulse width modulation
DQN	Distinct neural network
DDPG	Deep Deterministic Policy Gradient

rlTD3	twin-delayed deep deterministic policy gradient
PPO	Proximal Policy Optimization
LCL	Inductor–Capacitor–Inductor filter

Introduction

The escalating global demand for sustainable and clean energy sources has led to the emergence of solar photovoltaic (PV) systems as a viable solution to address the challenges of environmental pollution and the depletion of fossil fuel reserves. The deployment of solar photovoltaic (PV) systems that are connected to the grid has attracted significant attention due to their ability to efficiently harness abundant solar energy and seamlessly incorporate it into the power grid. The grid-connected solar PV system is a promising form of a nonconventional energy resource that generates electricity without emitting carbon, thereby contributing to a cleaner environment. Nevertheless, the intermittent nature of the solar PV system precludes its direct connection to the utility grid, necessitating asynchronous coupling through the utilization of converter devices.

The utilization of a grid-tied AC/DC converter, coupled with an appropriate control methodology, facilitates the fulfilment of grid connection prerequisites for solar photovoltaic systems. These prerequisites encompass electrical power flow management, harmonic mitigation, enhanced quality of power, stability, and grid harmonization [1]. The employment of photovoltaic (PV) arrays in the design of grid systems is widely adopted and acknowledged as the predominant method. The control problem of transmitting maximal electrical energy accessible to the load regardless of the situation is commonly referred to as MPPT. The primary objective of Maximum Power Point Tracking (MPPT) is to effectively regulate the fluctuations in output voltage that arise due to variations in PV power. Furthermore, photovoltaic systems demonstrate a non-linear correlation between output current along with voltage, resulting in substantial efficiency reductions. [2]

Typically, the interface circuit for the grid is expected to carry out three primary functions, namely voltage sensing, filtering, and A/D conversion. The occurrence of direct current (DC) offset in the measured grid voltage can be ascribed to the non-linear characteristics of voltage sensors, the analog-to-digital (A/D) conversion procedure, and the thermal drift exhibited by analog components. This phenomenon may occur despite the implementation of a well-designed grid interface circuit. [3] The input sine signal's undesirable component on the resultant waveform of the PLL structure is the addressed DC offset. The reference sine signal is commonly employed in the creation of reference currents for photovoltaic or grid-tied converters. Several established standards, including IEEE 1547-200, EN61000-3-2, and IEC 61727, delineate the permissible limit of direct current injection into the grid that may be attributed to photovoltaic systems or other converters that are connected to the grid. Various techniques can be employed to eliminate the induced DC offset in the measured grid voltage. The existence of inter-harmonics within the power system has a detrimental impact on its overall performance. The connection of a significant quantity of non-linear loads to the power system results in a range of complications, including inadequate power factor, excessive heating of transformers and power cables, impaired functioning of protection devices, heightened transmission losses, and substandard voltage regulation [3]. Inter-harmonic components result in significant energy wastage. The escalation of inter-harmonics within the power system can be closely associated with non-linear power electronic loads, such as variable frequency drives, converters, and inverters. Environmental conditions, such as temperature variations and irradiance fluctuations, can affect the performance of PV modules and inverters. These variations can introduce oscillation in the output power and current harmonics, including inter-harmonics.

As a result of the previously stated DC offset in voltage and current, the waveforms exhibit asymmetry along the x-axis. Consequently, the process of interrupting asymmetrical current is considerably more challenging than interrupting symmetrical current, as per reference [3]. Moreover, as a result of the direct current offset, there is a possibility of imbalanced grid voltages. This phenomenon has the potential to result in the degradation of power quality issues. Therefore, it can be concluded that harmonics and DC offset are unnecessary in a functional power system.

1.1 Problem Statement and Contributions

The efficiency of solar PV systems that are associated with the grid is subject to the impact of multiple factors, including fluctuations in solar irradiance, environmental circumstances, and distortions in grid voltage due to non-linear loads and conversion processes. These distortions include inter-harmonics and DC offset [2]. Several causes can account for the fact that some observed systems' output currents have unwelcome direct current components. Non-linearities in switching devices, small errors and offset drifts in voltage and current measurement sensors

used to provide feedback signals for control systems are all examples of sources of imprecision and error in PWM signals [14].

In the context of PV inverters, it is possible that inter-harmonics are attributable to the MPPT control mechanism. Partly shedding conditions lead to irradiance fluctuations that can affect the performance of PV modules and inverters. These variations can introduce fluctuations in the output power and current harmonics, including inter-harmonics. The conventional techniques employed for achieving maximum power point tracking (MPPT) and grid synchronization are constrained by issues related to inter-harmonics. Additionally, traditional synchronization methods, such as Phase Locked Loop (PLL), may struggle to maintain accurate synchronization with the grid in the presence of voltage distortions or frequency variations, results DC offset problems.

These challenges are addressed in this work, and contributions are:

- To address these challenges, this research aims to improve the power quality of grid-tied solar PV systems by proposing the utilization of a novel ensembled Deep Reinforcement Learning (EDRL) MPPT controller, and a Coati Optimization Algorithm tuned Fuzzified-Phase Locked Loop (COA Fuzzified-PLL) based synchronizing system.
- The ensembled DRL MPPT controller leverages the power of deep learning and reinforcement learning techniques to optimize the MPPT process, enabling the mitigation of inter-harmonics and efficient power extraction with a constant reference DC voltage at the DC link.
- The objective of the synchronizing system based on COA Fuzzified-PLL is to achieve reliable synchronization with the electrical grid, even under voltage distortions and frequency variations, while also being able to reject DC-offset.

1.2. Article Structure

The article begins with an introduction that highlights the significance of power quality enhancement in grid-tied PV systems and outlines the existing challenges. A comprehensive literature review is provided, examining the limitations of current control strategies and exploring previous research on adaptive control techniques. The system configuration and proposed controller of the grid-tied PV system with Ensembled DRL are presented in section III. The proposed COA-Fuzzified-PLL control strategy, including the Fuzzy membership function, is then described. The article proceeds to present experimental results to validate the effectiveness of the proposed strategy. A discussion section provides an in-depth evaluation of the approach, its advantages, limitations, and implications for power quality enhancement. Finally, the paper concludes with a summary of the key findings, contributions, and potential impact of the research, along with directions for future work.

Related Work

In current years, considerable research and development has focused on the power quality improvement of grid-connected photovoltaic scheme regarding harmonics and DC offset. Several research studies have examined the creation and impact of inter-harmonics, underscoring the necessity for efficient reduction methods. Moreover, direct current offset in photovoltaic systems results in transformer saturation and elevated losses. The scholarly community has shown interest in identifying and eliminating DC offset. Various methods for mitigating these issues have been suggested to alleviate the issues of harmonics and DC offset, such as grid synchronization, phase-locked loop (PLL) [3], synchronous reference frame (SRF-PLL) [4], and non-PLL techniques such as SOGI [9], Cascaded SOGI [10] within the context of this document. Thus, the attainment of accurate estimation and the elimination of unwanted periodic ripple in SRF-PLL [4] can be accomplished in the presence of a DC offset in the input signal. The paper [4] presents a novel phase-locked loop (PLL) that utilizes the Extended Second-Order Generalized Integrator (ESOGI) and is reconfigured by the Second-Order Generalized Integrator (SOGI) with Active Power Filter (APF). The study also suggests that the ESOGI-PLL exhibits satisfactory performance, even under conditions of elevated dc offset values and a high frequency of low-order harmonics. Furthermore, due to its uncomplicated architecture, the ESOGI-PLL suggested in this study can be readily executed on an inexpensive microcontroller. According to source [5], the second-order generalized integrator, which corresponds to the SOGI-PLL, exhibits a relatively rapid transient response, a high capacity for rejecting disturbances, and a robust performance. One of the previously published works has suggested the utilization of SOGI-FLL [6] to improve the efficacy of SOGI-PLL in the context of frequency fluctuations. This is because the performance of SOGI-PLL is deemed inadequate in scenarios with variations in frequency and DC offset in the grid voltage. The article [1] presents a proposal for a comb filter utilizing a modified sliding Goertzel discrete Fourier transform (SGDFT)-based phase-locked loop (PLL). The design incorporates the three degrees of freedom (DOFs) of second-order

fraction delay, to mitigate the effects of non-integer frequency components. The SGDFT-based PLL [1] operates at a constant sampling frequency and aims to calculate the fundamental frequency, amplitude, and phase angle in order to achieve optimal synchronization. The paper [7], utilizes the Advanced Third Order Generalised Integrator (ATOGI) for controlling a two-stage three-phase photovoltaic system power quality. The ATOGI [7] is utilized for the purpose of extracting essential components from distorted grid voltages and non-linear load current. The proposed method effectively addresses the integrator delay and inter-harmonic challenges inherent in conventional SOGI techniques. Additionally, the DC-Offset Estimator component exhibits robust dc-offset rejection capabilities. The study [8] showcases a noteworthy implementation of two distinct generalized integrators: SOGI and the Reduced-Order Generalized Integrator (ROGI) controller, which are interconnected in a cascade configuration. The outcome of SOGI-ROGI yields several desirable characteristics, including the absence of phase shift, optimal filtering, minimal harmonic distortion, and favourable dynamic response. Furthermore, filtering techniques, such as passive filters and active power filters, have been explored to suppress these issues. Despite progress, challenges remain, including the need for improved detection methods and the integration of emerging technologies.

As discussed in the introduction section, the voltage disturbance of photovoltaic arrays results in power oscillations, particularly during partial shedding operation. These power oscillations contain inter harmonics. Maximum Power Point Tracking (MPPT) control techniques have been suggested to harness maximum output and increase power quality over the years. Several approaches have been proposed in the literature for maximum power point tracking (MPPT) in photovoltaic (PV) power systems. These include a modified incremental conduction MPPT algorithm with a fuzzy controller [11], a hill-climbing (HC) modified fuzzy-logic (FL) MPPT control scheme implemented in both software and hardware [12], and a hybrid MPPT control strategy that integrates a modified perturb and observe (P&O) algorithm with an enhanced particle swarm optimization (PSO) algorithm [13]. Numerous heuristic techniques have been employed in the field, including the innovative Maximum Power Point Tracking (MPPT) method, which is adaptable to applications with rapid fluctuations through the utilization of Artificial Neural Network (ANN) [14]. Most of these techniques are based on models and aim to regulate various photovoltaic (PV) systems but not predict the inter-harmonics characteristics. The acquisition of an accurate model for photovoltaic (PV) systems and their associated parameters can remedy these harmonic challenges that are associated with PV panels in various configurations. The author is compelled to seek a methodology that is independent of any specific model.

While several research studies have examined specific elements of inter-harmonics and DC offset in grid-connected solar PV systems, there is a limited availability of sophisticated approaches and comprehensive methodologies that effectively tackle both of these concerns concurrently. The research gap pertains to the lack of comprehensive solutions that can adequately address the suppression and mitigation of inter-harmonics and DC offset, taking into account their potential interactions and cumulative impact on power quality. The simultaneous management of these phenomena will greatly improve the efficiency, reliability, and performance of grid-connected solar photovoltaic (PV) systems.

System topology and control architecture

3.1 System Topology

The system is comprised of three primary components, namely photovoltaic (PV) panels or arrays, PV inverters, and the alternating current (ac) grid. PV inverters, such as the full-bridge inverter (IGBT), are integral in regulating the transmission of power from PV arrays to the AC grid. The Ensembled DRL maximum power point (MPP) based on total harmonic distortion (THD) Reward technique is employed for the purpose of producing a PWM signal that is utilized in a boost converter, with the aim of enhancing the voltage profile of the photovoltaic (PV) system and mitigating the inter-harmonic component. Figure 1 illustrates the conventional control configuration of 3-phase grid-connected photovoltaic (PV) inverters, A photovoltaic array with a power output of 250 kilowatts is linked to a 25 kV electrical grid through a three-phase converter. The photovoltaic (PV) array is composed of 88 parallel strings. A series connection of 7 SunPower SPR-415E modules is present in each string. Table 1 provides the relevant system parameters. To achieve optimal power extraction from photovoltaic arrays, the implementation of a maximum power point tracking (MPPT) algorithm is necessary.

Table 1 The parameter setting of PV

Parameters	Values	Standard Units
PV arrays (series connected)	7	-
PV arrays (parallel connected)	88	-
Current of single PV array	6.09	A
Voltage of single PV array	85.3	V
Maximum power output	250	K watts
Temperature	0-500	
Irradiation	680-1000	

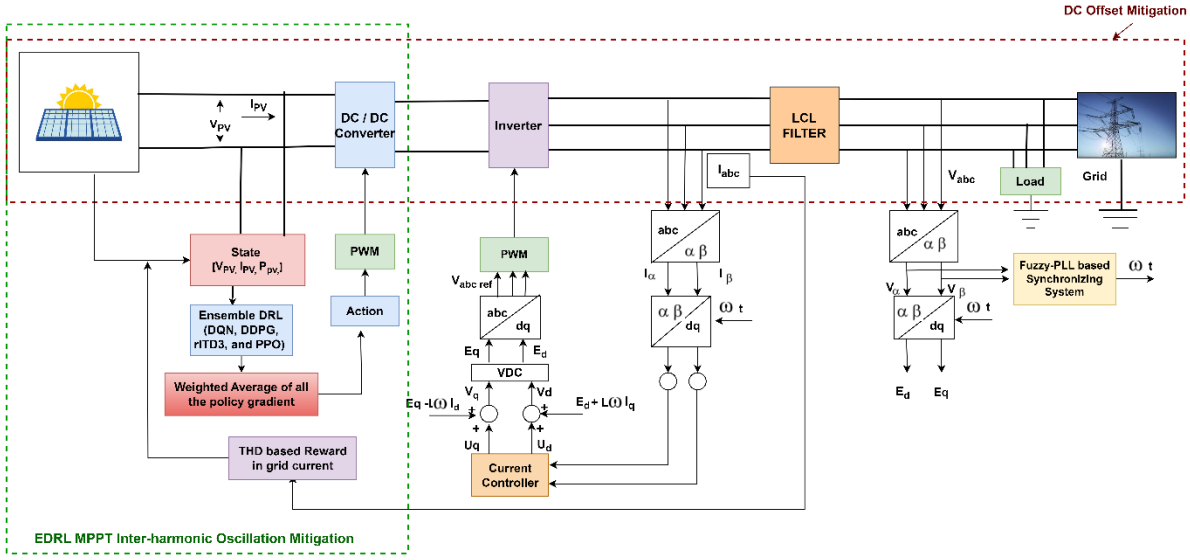

Figure 1. Proposed methodology for performance quality enhancement of three Phase grid-connected solar PV system

3.2 Proposed Methodology

The methodology employs two key techniques: Ensembled Deep Reinforcement Learning (EDRL) based Maximum Power Point Tracking (MPPT) for inter-harmonics removal and a Novel Coati Optimization Algorithm (COA)-fuzzified Phase-Locked Loop (PLL) for DC offset removal is shown in figure 1. The methodology begins by utilizing Ensembled DRL, a combination of reinforcement learning and ensembled policy gradients, to optimize the MPPT process. Multiple neural network agents are trained independently, each responsible for tracking the maximum power point of the solar PV system. These agents collaborate and share their experiences through weighted averaging, ensuring diverse exploration and improving the accuracy of the MPPT process. This ensemble approach enhances the system's performance by mitigating the effects of inter-harmonics, unwanted harmonics that can degrade power quality.

The EDRL MPPT technique is utilized to ascertain the reference dc-link voltage during operation and to aid in mitigating inter-harmonic oscillation. Following this, the EDRL controller efficiently regulates the DC-link voltage, denoted as V_{dc}, by utilizing a reward or penalty system based on the grid currents' total harmonic distortion (THD). Inter-harmonics are a type of electrical disturbance that can occur in power systems. They are non-integer harmonic frequencies that fall between the standard integer harmonic frequencies. Inter-harmonics can result from various sources, including partial shedding conditions, non-linear loads, voltage fluctuations, and disturbances in the power system. However, the presence of inter-harmonics in the output current can be attributed to the utilization of power semiconductor devices and the variable power flow of the photovoltaic (PV) panels. Moreover, it is worth noting that the grid typically provides numerous non-linear loads that have the capability to absorb distorted currents. The fluctuating currents that pass through the impedances of the power distribution system, which vary in accordance with frequency, lead to a distortion of the voltage at the system bus.

The subsequent approach incorporates a COA-fuzzified phase-locked loop (PLL) to alleviate the issue of direct current (DC) offset. The optimal performance of a fuzzy logic controller is achieved when the membership

function parameters are efficiently and accurately tuned using COA optimization. COA-fuzzified PLL is an efficient approach to identifying and eliminating DC offset in the solar PV system connected to the grid. This technique guarantees precise synchronization with the utility grid and reduces the likelihood of power quality problems. The proposed methodology integrates the ensembled DRL-based MPPT and COA-fuzzified PLL techniques to enhance the overall performance of the 3-phase grid-connected solar PV system. The ensembled DRL optimizes the MPPT process, while the COA-fuzzified PLL eliminates DC offset and ensures precise synchronization with the utility grid. By effectively addressing inter-harmonics and DC offset, the methodology aims to improve power quality, maximize energy extraction, and enhance the overall performance of grid-connected solar PV systems.

3.3 Ensembled DRL-MPPT Approach for Inter-harmonic Mitigation

The present study introduces a novel approach that involves incorporating a THD reward within the framework of ensembled deep reinforcement learning (EDRL) to address the issue of maximum power point tracking (MPPT) in photovoltaic (PV) arrays. Ensembled Deep Reinforcement learning (EDRL) provides an effective solution to this problem without requiring any parametric information regarding the dynamic parameters of the model. The objective of the algorithm is to illustrate the system analogy utilizing the DQN (discrete), DDPG (continuous), rlTD3 (continuous), and PPO (discrete) network framework. The ensembled DRL approach can be shown in Figure 2.

The DRL consists of four primary components. The three fundamental components of DRL are commonly referred to as the state space X, reward function r, and action space U. The primary principle of Maximum Power Point Tracking (MPPT) is to optimize the power extraction from photovoltaic (PV) modules by ensuring their operation at the voltage corresponding to the maximum power point, thereby maximizing the available power output. The acquisition of knowledge by the agent is facilitated by any form of engagement with the environment. This process involves the execution of an action $u_t \in U$, which triggers the evolution of the system from its current state $x_t \in X$ to the subsequent state x_{t+1}. The agent obtains feedback in the form of a THD reward, which serves as a quantitative measure of the efficacy of the action or decision made by the agent. Consequently, the incentive serves as an indication to pinpoint the attainable objective or ideal resolution. The RL approach aims to identify an optimal policy π that meets a given set of criteria.

$$J^* = \max_{\pi} J_\pi = \max_{\pi} E_\pi \{r_t | x_t = x\} \quad (1)$$

The symbol J_π denotes the cumulative expected reward under a given policy π. Assuming a policy π that absolves, the value function for a given time interval, denoted as $V^\pi(x)$, or the expected cumulative reward, is a function of $x^\pi$and is defined as $x^\pi = \{x_t\}_{t=1}^{t=n}$, where the state values are represented by $k^\pi = \{k_t\}_{t=1}^{t=n}$, which are sequences of actions taken by the agent.

State Space

The state-space design in Maximum Power Point Tracking (MPPT) problems involves analysing the movement of the Maximum Power Point (MPP) on the Photovoltaic (PV) curve across varying environmental conditions. The methodology employed in reinforcement learning is regulated by the photovoltaic current, power, and the direct current voltage of the coupling. The state-space comprises X, which is a vector containing the variables $X \in \left[V_{PV}, I_{PV}, P_{PV}, \Delta P_{PV}, \int \Delta P_{PV}, \Delta V_{DC} \right]$. The selection of duty cycle within the range of $[0,1]$ is governed by the variable X. The divergence of PV power from the intended generation capacity is denoted by ΔP_{PV} , while the discrepancy between the reference coupling voltage and the measured V_{DC} is represented by ΔV_{DC}.

Action Space

The MPPT problem is typically associated with a discrete action space. The aforementioned approach ensures a notable degree of accuracy and serves as a potent pedagogical strategy, rendering it a computationally expedient methodology. The EDRL-MPPT agent's action involves a predetermined duty cycle. The duty cycle D_c is determined within the range of $D_c = (0,1]$ through a sequence of actions, with an incremental interval of 0.01 in case of discrete action space. Consequently, a matrix comprising of one hundred potential actions is generated.

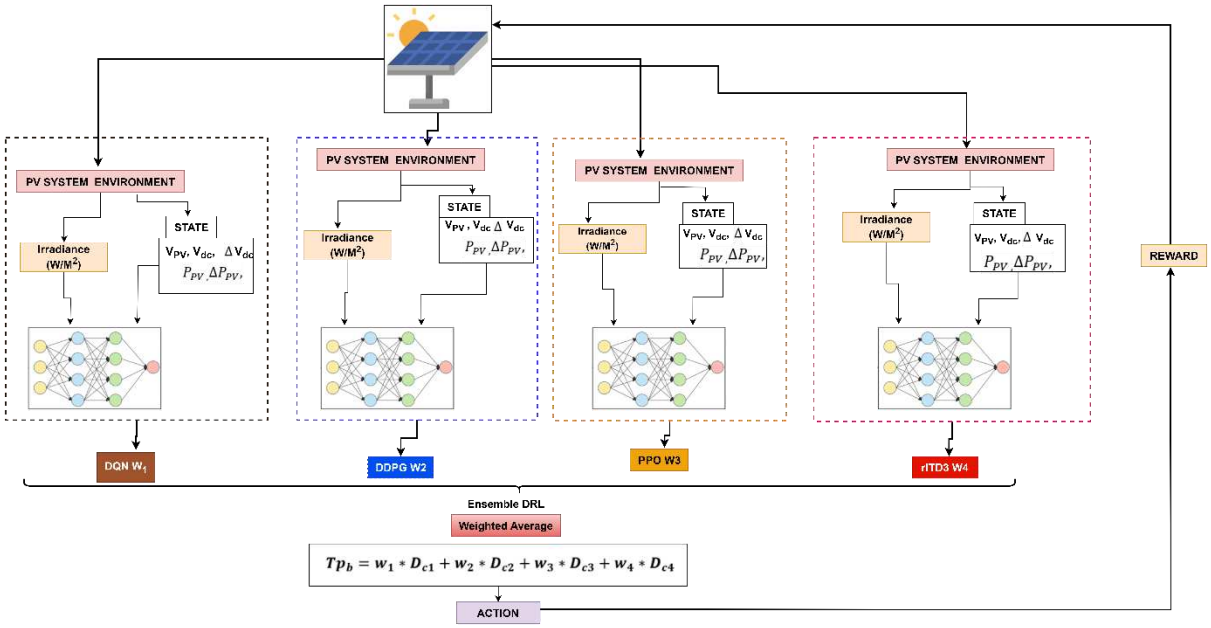

Figure 2. Ensembled DRL for MPPT control in a three-phase grid connected PV system

Reward

The agent receives a reward contingent upon the action executed. The reward signals have been designed to achieve the objective of minimizing the total harmonic distortion in grid current. Total harmonic distortion (THD) is a crucial metric for assessing power quality as it measures the level of harmonic distortion that exists in the waveform of the grid current. Through the incorporation of Total Harmonic Distortion (THD) as a form of incentive in ensembled Deep Reinforcement Learning (EDRL), the agents are capable of acquiring knowledge to enhance the efficiency of the system's operations and control tactics. This enables the minimization of harmonic distortion and guarantees a seamless transfer of power to the grid. The utilization of the ensemble approach facilitates the agents to collaboratively explore diverse control parameters, strategies, and switching techniques with the aim of mitigating total harmonic distortion (THD). By means of iterative learning and collaborative efforts, the ensemble agents are capable of identifying optimal solutions that effectively reduce harmonic distortion and improve the overall power quality within the grid-connected photovoltaic system.

The total harmonic distortion (THD) in grid current is a measure of the harmonic distortion present in the current waveform of a three-phase grid-connected PV system. When designing a reward function for Deep Reinforcement Learning (DRL) applied in such a system, the THD can be incorporated as a component to incentivize the reduction of harmonic distortions.

The mathematical formula for calculating THD in the grid current can be expressed as follows:

$$THD = \sqrt{\left(\frac{I_{rms}}{I_1}\right)^2} \times 100 \tag{2}$$

Where, THD is the total harmonic distortion of the grid current, I_{rms} represents the rms value of the individual harmonic current components, I_1 represents the rms value of the fundamental current component. When incorporating THD in the reward function for DRL, the goal would be to minimize the THD value. This can be achieved by assigning a negative reward proportional to the THD value. For instance, the THD-based reward function can be defined as:

$$Reward = -K \times THD \tag{3}$$

Where, Reward is the reward value assigned to the DRL agent, K is a scaling constant that determines the weight or importance of the THD in the overall reward calculation. By using this reward function, the DRL agent is encouraged to learn policies that result in lower THD values, thus promoting a reduction in harmonic distortions in the grid current of the three-phase grid-connected PV system.

The proposed ensembled deep reinforcement learning (EDRL) involves the construction of a collective of DRL models that cooperate in order to enhance the overall efficacy and resilience of the learning system. The weighted ensembled DRL learning process involves training multiple DRL models independently, each with its own set of weights. During the decision-making phase, the models' outputs are combined using weighted averaging or another aggregation method that considers the assigned weights. The weights can be adjusted dynamically based on the models' performance, exploration-exploitation trade-offs, or other criteria. This approach aims to leverage the diversity and expertise of individual agents while assigning different degrees of importance to their contributions based on their performance or confidence levels. Algorithm-1 presents the proposed weighted average deep reinforcement learning (DRL) approach. Four distinct neural network agents (namely DQN, DDPG, rlTD3, and PPO) were utilized in this study. Each agent was trained independently, with unique sets of parameters, exploration strategies, and architectures, as outlined in the subsequent section. The aforementioned agents engage with their surroundings, obtain incentives, and revise their strategies through reinforcement learning methodologies. Upon completion of training the four Deep Reinforcement Learning (DRL) models, the resultant policy is utilized to obtain the duty cycle denoted as D_c. Therefore, the overall probability denoted as Tp_b can be expressed as a weighted average of the duty cycle obtained from individual DRLs, as shown in Equation (4).

$$Tp_b = w_1 * D_{c1} + w_2 * D_{c2} + w_3 * D_{c3} + w_4 * D_{c4} \quad (4)$$

The weights of DQN (discrete), DDPG (continuous), rlTD3 (continuous), and PPO (discrete) are denoted as w_1, w_2 , w_3 , and w_4, respectively. Additionally, the duty cycle for each of the DRL models are represented as D_{c1}, D_{c2}, D_{c3}, and D_{c4}. Based on the Algorithm of DQN (discrete), DDPG (continuous), rlTD3 (continuous), and PPO [23], the models are trained and saved. The approach employed to obtain output from a weighted average deep reinforcement learning (DRL) method is derived from the research paper referenced as [23]. The method is applied to each model, resulting in the acquisition of D_{c1}, D_{c2}, D_{c3}, and D_{c4}. The final weighted average action is obtained using Equation (4). The model structure that has been suggested is depicted in Figure 2. The algorithm

Algorithm 1: EDRL THD Reward MPPT PV Control

1. Establish a connection to the solar PV array SunPower SPR-415E.
2. Determine the magnitude of the current and voltage that results from a short circuit and open circuit
3. Calculate the maximum power for $N_S = 7$ (PV's in series) and $N_p = 88$ (PV's in parallel) using $P_{npp} = \left(N_S \times V_{mpp}\right) \times \left(N_P \times I_{mpp}\right)$
4. Choose the DC-link voltage.
5. Set the EDRL agent's initial state, action, and reward.
6. State-space X=$\left[V_{PV}, I_{PV}, P_{PV}, \Delta P_{PV}, \int \Delta P_{PV} \right]$
7. Action space U= $(0,1]$
8. Update duty cycle D_c
9. calculate: $e(t)$ and $\Delta e(t)$
10. Pass the error and $\Delta e(t)$ through
11. Provides these values to the network DQN, PPO, TD3, and DDPG
12. Initialize / Load Q, α 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 $(x_t, u_t, u_{t+1}, u_{t+1})$
20. IF $|R| > N$
21. Update the network using weighted average EDRL:
22. **end if**
23. Set $x_t = x_{t+1}$
24. **end for**
25. **end for**

The present study employs an EDRL methodology that involves assessing a weighted average of the behaviors exhibited by four distinct models. These models, which will be elaborated upon in the subsequent section, consist of two models operating within a continuous action space and two operating within a discrete action space. The average reward per episode during training of EDRL can be shown in figure 3.

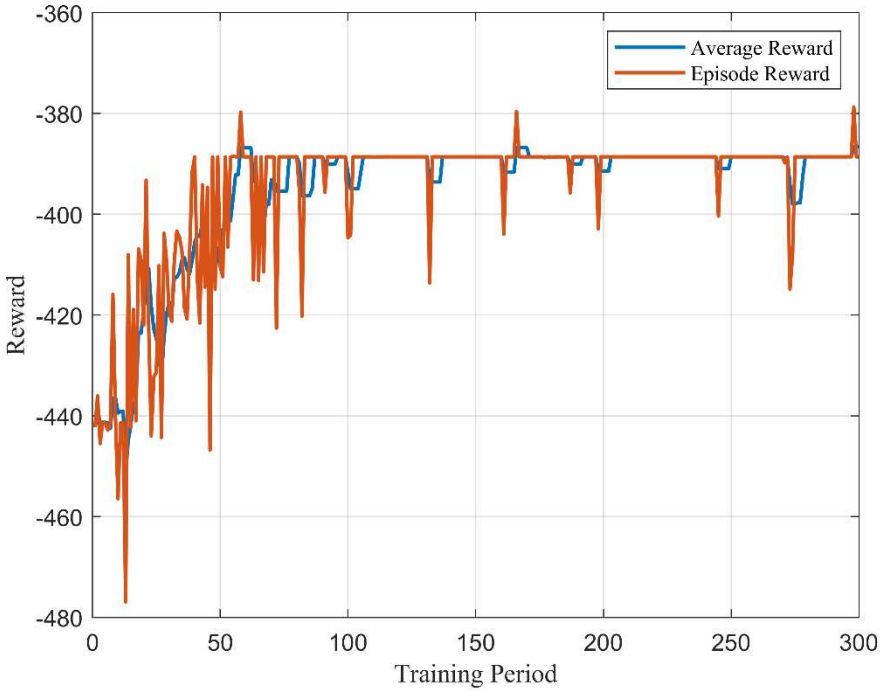

Figure 3. The average reward per episode during training of EDRL

3.3.1 DQN (Discrete)

DQN and its associated variant, double Q, are commonly employed to manage a discrete action space [16-19]. As per reference, the selection of a policy to implement in DQN is contingent upon identifying the action that will result in the highest attainable reward for the present state. The primary concern raised in DQN pertains to an overestimation of the Q-value in action, which poses a hindrance to the development of an optimal strategy. The utilization of a policy-based paradigm is advantageous in addressing the intricacy of DQN. The DDPG and PPO models exhibit versatility in their ability to operate effectively in both continuous and discontinuous action spaces. The present study employs a sole DQN in the context of DDPG and a discrete action space, as opposed to a continuous one. Consequently, the model computes the ensembled average of the behaviour with appropriate weighting.

DQN stores state-action pairings as $< S_t, A_t, S_{t+1}, A_{t+1} >$.. S_{t+1}, A_{t+1} are the states and actions at $t+1$. DQN provides extremely uncorrelated data from random prior events.

Weight loss function (θ):

$$L_i(\theta_i) = \mathbb{E}_{S,A\sim\rho(.)}[(yS_i - Q(S,A;\theta_i))^2] \tag{5}$$

Equation (5) gives the target as y_i and prediction as $Q(S,A;\theta_i)$. In the previous iteration, the weights were θ_{i-1}, and the desired output from Equation (5) was

$$y_i = \mathbb{E}_{S',\sim\varepsilon}[r + \gamma \max_{A'} Q(S',A';\theta_{i-1})] \tag{6}$$

The stochastic gradient of equation 5 with the replacement of the target $y_i$using equation 6 is

$$\nabla_{\theta_i} L_i(\theta_i) = \mathbb{E}_{S,A\sim\rho(.)}[(r + \gamma \max_{A'} Q(S',A';\theta_{i-1}) - Q(S,A;\theta_i))\nabla_{\theta_i} Q(S,A;\theta_i)] \tag{7}$$

The variable γ denotes the discount factor, while r represents the reward. The computational procedure executes for a total of M iterations. The agent selects a set of tuples from the existing data based on the chosen action. The execution of the task is obtained by the agent via the values indicated in Equation (7). The incentive is contingent upon behaviour and can manifest as either a favourable or unfavourable outcome.

3.3.2 DDPG (Continuous)

The DDPG model is categorized as a policy-based approach that is applicable to action spaces that are either continuous or discrete. The present approach is proposed as a solution to address the challenges encountered in the Deep Q-Network (DQN) and to mitigate the impact of densely connected neural networks. The approach described in reference [21] has demonstrated efficacy in managing environments characterized by continuous action spaces. Therefore, it is highly appropriate for the current academic research setting.

The DDPG architecture comprises two distinct components, namely the actor and critic. The representation of the actor is denoted as $\mu(S|\theta^{\mu}))$. The representation of the critic can be denoted as $(S, A|\theta^{Q}))$.

The gradient update incorporates the deterministic policy function denoted by θ^{μ} and the Q network denoted by θ^{Q}. During each iteration of the training process, the actors and critic engage in the exchange of information. In the context of soft update networks in Deep Deterministic Policy Gradient (DDPG), the actor and critic networks are denoted as $\mu(S|\theta^{\mu'}))$ and $(S, A|\theta^{Q'}))$ respectively.

The notation used in the text denotes that $\theta^{\mu'}$ refers to the Target policy network, while $\theta^{Q'}$ refers to the Target Q network. The policy function with direction, denoted as θ^{μ}, is commonly represented as $J(\theta^{\mu})$, while its gradient is typically denoted as follows:

$$\frac{\partial J(\theta^{\mu})}{\partial \theta^{\mu}} = E[\nabla_{\mu(S)} Q(S, \mu(S|\theta^{\mu})|\theta^{Q}) \nabla_{\theta^{\mu}} \mu(S|\theta^{\mu})] \tag{8}$$

The critic network in the Deep Deterministic Policy Gradient (DDPG) algorithm minimizes the loss function, which is the Mean Square Error (MSE), with respect to the action (A). The resulting expression for the loss function is obtained as a result.

$$(\theta^{Q}) = E[\left(Q_{target} - Q_{predict}\right)^{2}] \tag{9}$$

Where, $Q_{target} = r + \gamma Q\left(S_{t+1}\mu\left(S_{t+1}\middle|\theta^{\mu'}\right)\middle|\theta^{Q'}\right)$ and $Q_{predict} = Q(S, A|\theta^{Q})$

In contrast to the DQN approach, the target networks undergo updates at each time and step through the utilization of soft updates.

3.3.3 PPO (Discrete)

The computation of PPO is an integral component of the weighted ensembled strategy employed in this context. The model in question is a policy-based approach that regulates the gradient of the policy update, as described in reference [21]. This is carried out to ensure alignment between the policies. The utilization of this can be applied to either a discrete or continuous action space. The implication of this statement is that the on-policy model in a discrete action space exhibits a limited capacity for policy modifications during implementation. The evaluation of the performance of a chosen action is conducted through the utilization of the critic network, as per the advantage function. Equation (7) provides the advantage function.

$$\hat{A}_t = \delta_t + (\lambda_{\gamma})\delta_{t+1} + \cdots + (\lambda_{\gamma}{}^{T-t+1})\,\delta_{T-1} \tag{10}$$

$$\delta_t = r_t + \gamma V_{\pi}(S_{t+1}) - V_{\pi}(S_t) \tag{11}$$

In Equation (8), the state-value function is denoted as $V_{\pi}(S)$ and serves as a representation.

$$V_{\pi}(S) = \mathbb{E}_{\pi}[\textstyle\sum_{k=0}^{\infty} \gamma^{k} r_{t+k+1} | S_t = S] \tag{12}$$

The policy utilized for environmental sampling is denoted as $\pi_{\theta old}$, while the policy subject

t to optimization is represented as π_θ. PPO utilizes the clipped surrogate objective to establish the bounding constraints on the policy updates, thereby ensuring the stability of the training process. The target function utilized in PPO undergoes a transformation, as depicted in Equation (13).

$$J_{\theta} \approx \textstyle\sum_{(S_t, A_t)} \min\left(\frac{\pi_{\theta}(A_t/S_t)}{\pi_{\theta old}} \hat{A}_t, clip\left(\frac{\pi_{\theta}(A_t/S_t)}{\pi_{\theta old}}, 1-\epsilon, 1+\epsilon\right)\hat{A}_t\right. \tag{13}$$

3.3.4 rlTD3 twin-delayed deep deterministic policy gradient (Continuous)

TD3, an actor-critic structure, combines action as a policy function and value function on the present policy. TD3 needs continuous action space [22]. TD3, the upgraded DDPG, eliminates value function overestimation. Reducing value function overestimation improves accuracy and reduces variance. The TD3 network forms the

target network using the smallest of two critic networks. TD3's delayed actor network updates every two-time step for stability and efficiency throughout training. Clipped noise calculates targets when the action is selected. High action value makes the model robust in all scenarios (continuous action-space). DDPG overestimates the Q-function, but TD3 employs dual Q-functions, delayed policy updates to ensure stability, and smoothens the target policy. Twin delay DDPG. For continuous action-space environments.

Equation (14) represents the target action of the policy $\mu_{\theta targ}$, which has been modified by the addition of clipped noise.

$$A'(S') = clip\big(\mu_{\theta targ}(S') + clip(\in, -c, c), A_{Low}, A_{High}\big), \in \sim \mathcal{N}(0, \sigma) \tag{14}$$

All instances of the target action A are found to be satisfactory, $A_{Low} \leq A \leq A_{High}$

Equation (15) presents the clipped double Q-learning function.,

$$y(r, S', d) = r + \gamma(1 - d) \min_{i=1,2} Q_{\phi_{i,targ}}(S', A'(S')) \tag{15}$$

The policy is acquired through the process of maximizing the Q_{ϕ_1}:

$$\max_{\theta} \mathrm{E}[Q_{\phi_1}(S, \mu_\theta(S))] \tag{16}$$

Equation (17) provides the expression for the Q-function through one-step gradient descent.

$$\nabla_{\phi_i} \frac{1}{|B|} \sum_{(S,A,r,S',d)} (Q_{\phi_1}(S, A) - y(r, S', d))^2 \tag{17}$$

for $i = 1,2$

Equation (18) depicts the policy update utilizing one step gradient ascent,

$$\nabla_{\phi_i} \frac{1}{|B|} \sum_{S \epsilon B} Q_{\phi_1}(S, \mu_\theta(S)) \tag{18}$$

3.4 COA Fuzzified-PLL based Controller for DC offset Removal

The output of a PV inverter typically exhibits a direct current (DC) offset voltage component. This phenomenon arises due to various factors, such as discrepancies among power modules, asymmetry in driving pulses, and errors in current detection. According to reference [3], the presence of induced DC offset in the measured grid voltage can result in undesirable fluctuations in the estimated values of both the amplitude and frequency of the grid voltage. The amplitude of the induced ripple is contingent upon both the percentage of the DC offset value and the fundamental frequency of the power grid. Thus, the presence of DC offset in the measured grid voltage renders the estimation procedure of the grid parameters virtually infeasible. This paper presents a novel approach for mitigating induced DC offset. The proposed method employs COA Fuzzified-PLL-based controller, can be shown in Figure 3. The Fuzzified-Phase-Locked Loop (PLL) was developed through the replacement of the Proportional-Integral (PI) controller with a Fuzzy Logic Controller (FLC) within the context of the PLL. The Coati Optimization algorithms are employed to optimize the parameters of the fuzzy, including membership functions, rule base, and scaling factors. Fuzzification is applied to handle the uncertainty and imprecision associated with DC offset, transforming the input signals into fuzzy sets.

The phase-locked loop (PLL) operates by minimizing the phase discrepancy between a reference signal and a feedback signal within the control loop. The process involves modifying the phase of a voltage-controlled oscillator (VCO) until the two signals achieve phase alignment. Phase-locked loops (PLLs) are vulnerable to various factors such as noise, non-linearities, and abrupt disturbances, which can adversely impact their operational efficiency. The introduction of fuzzy logic to the PLL architecture helps overcome these limitations. Changes in angular phase angle $(\Delta\theta)$ can be observed subsequent to each phase angle jump. The application of phase angle change is intended to address subtle and abrupt changes. The proposed model inserts the fuzzy controller between the phase detector and the low-pass filter in a conventional PLL device.

In the conventional phase-locked loop (PLL), the three-phase voltage vector is converted from the abc natural reference frame to the $\alpha\beta$ stationary reference frame through the utilization of Clarke's transformation. Subsequently, it is further converted to the dq rotating frame using Park's transformation, as depicted in Figure 1. The proposed modification being suggested for the PLL framework pertains to its existing control mechanism. The difference between the reference currents I_d and I_q and the measured currents I_d and I_q is utilized as the input

signal for the fuzzy logic controller. The equation depicts the modelling design of Fuzzified-PLL. The variables $V_a, Vb,$ and Vc represent the magnitudes of the three-phase voltage.

$$\begin{bmatrix} V_a \\ V_b \\ V_c \end{bmatrix} = \begin{bmatrix} V_m \cos\theta \\ V_m \cos(\theta - {}^{2\pi}/_{3}) \\ V_m \cos(\theta + {}^{2\pi}/_{3}) \end{bmatrix} \tag{19}$$

The conversion of these signals into the stationary reference frame signals V_α and V_β is achieved through the application of the Clarke transformation, while the Park transformation is utilized for the conversion to the dq frame.

$$\begin{bmatrix} V_\alpha \\ V_\beta \end{bmatrix} = {}^{2}/_{3} \begin{bmatrix} 1 & {}^{-1}/_{2} & {}^{-1}/_{2} \\ 0 & {}^{-\sqrt{3}}/_{2} & {}^{\sqrt{3}}/_{2} \end{bmatrix} \begin{matrix} V_a \\ V_b \\ V_c \end{matrix} \tag{20}$$

$$\begin{bmatrix} V_d \\ V_q \end{bmatrix} = \begin{bmatrix} \cos\theta^* & -\sin\theta^* \\ \sin\theta^* & \cos\theta^* \end{bmatrix} \begin{bmatrix} V_\alpha \\ V_\beta \end{bmatrix} \tag{21}$$

$$V_d = V_\alpha \cos\theta^* - V_\beta \sin\theta^* \approx V_{d,Offset} + V_m \tag{22}$$

$$V_q = V_\alpha \sin\theta^* + V_\beta \cos\theta^* \approx V_{q,offset} - V_{me} \tag{23}$$

The symbols $V, \theta,$ and θ^* represent the, magnitude of voltage, input angle, and estimated angle, respectively. e is the error in phase angle. The main objective of the proposed controller is to eliminate the direct current (DC) offset component that exists in the synchronous q-axis and d-axis components, denoted as $V_{d,offset}$ and $V_{q,offset}$ respectively which is shown equation (23).

The fuzzy controller employs optimized parameters and a rule base to make decisions based on the fuzzy sets of the input signals. The controller generates control signals that adjust the PLL's parameters dynamically to cancel out the DC offset component effectively. The fundamental stages of fuzzy logic control consist of three distinct phases, which are fuzzification, decision-making, and defuzzification. The procedure of fuzzification pertains to the transformation of quantitative measurements of input variables into a precisely defined linguistic variable boundary, which is denoted by a fuzzy set. The performance of the optimized fuzzy-PLL is evaluated by assessing the accuracy of DC offset removal and the overall improvement in power quality.

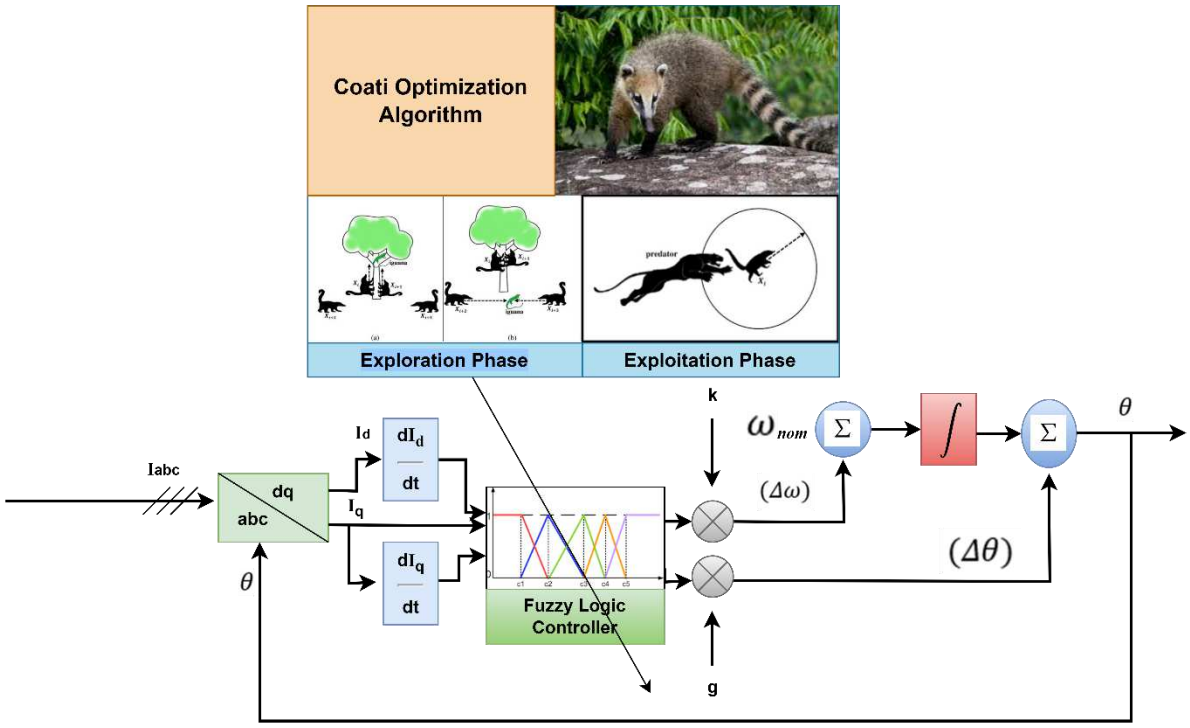

Figure 4. COA Fuzzified-PLL based Controller

The Novel Coati Optimization algorithm is used to optimize the parameter of fuzzified-PLL. Figure 4 illustrates a comprehensive model of a fuzzy phase-locked loop (PLL). The global optimization COA approach offers numerous advantages, including the absence of control parameters, the ability to solve complex high-dimensional problems across various domains, superior research and search process balancing, and effective handling of optimization applications. Given the presence of two inputs and one output, 18 positions of membership functions must be optimized by utilizing the Coati Optimization (COA) algorithm. An objective function is selected to minimize the total harmonic distortion in the optimization process. The fuzzy controller being examined employs linguistic variables, specifically negative big (NB), negative medium (NM), small negative (NS), zero (ZE), small positive (PS), medium positive (PM), and positive big (PB), to establish the error and error rate. These variables are differentiated by their corresponding memberships. The concept of memberships pertains to mathematical curves that facilitate the process of mapping each point within the input space to a corresponding membership value that falls within the range of 0 to 1.

The study involves the modification of eighteen membership function positions for each coati, followed by evaluating a fuzzified-PLL system based on the resulting error. The membership range of two inputs is defined as [-10 to 10] and [-1 to 1], and one output as [-10 to 10]. The above ranges remain constant throughout the simulation and do not undergo any modifications. The trapezoidal function has a fixed range that spans from negative infinity to positive infinity, both at the initial and final points. Additionally, it is worth noting that two points of intersection exist between each membership function and other membership functions. Table 3 provides clear evidence of this assertion.

Table 2 The range values for the input/output of fuzzy controller.

Membership functions	Range parameters
Trapezoidal	$[-\infty, -0.032, x\ (1), x\ (2)]$
Triangular	$[x\ (1), x\ (2), x\ (3)]$
Triangular	$[x\ (2), x\ (3), 0]$
Triangular	$[\ x\ (3), 0, x\ (4)]$
Triangular	$[0, x\ (4), x\ (5)]$
Triangular	$[x\ (4), x\ (5), x\ (6)]$
Trapezoidal	$[x\ (5), x\ (6), 0.032, \infty]$

The most common values can be readily anticipated based on the data presented in Table 2. In the case of a single variable, the tuning process involves adjusting four specific values. However, when dealing with three variables, the total number of values to be tuned increases to eighteen. There exist certain constraints that necessitate consideration when adjusting these values.

The given problem is subject to certain constraints, which is given n table 3. require that every value must adhere to the specified inequality criteria:

Table 3. Constraints for Fuzzy Controller

Input	**Boundary Condition**
Order	$x(1) < x(2) < x(3)$
Error E	-0.032 to 0.032
Change in Error ΔE	-10 to 10
Phase angle output $\Delta\theta$	-1 to 1

The membership function values of the fuzzy controller are modified and the model is then executed using these new values. The total harmonic distortion (THD) will serve as the target to be minimized, representing the value of the objective function.

$$objective\ function = min\ THD \tag{24}$$

Table 4 presents elaborate variations of the expressions mentioned above for alterations in phase angle. The q-axis component of the grid voltage, along with its derivative, are regarded as inputs to the FLC. In order to incorporate variables with distinct ranges, it is necessary to multiply two variables by the outputs of Fuzzy Logic Controllers (FLCs). The parameters depicted in Figure 4 are denoted by k and g. Figure 4 illustrates a comprehensive model of a fuzzy phase-locked loop (PLL).

Table 4 Fuzzy rule for output variable $\Delta\theta$

$\Delta\theta$	NB	NM	NS	ZE	PS	PM	PB
NB	NB	NB	NB	NM	NM	NS	ZE
NM	NB	NB	NM	NM	NS	ZE	PS
NS	NB	NM	NM	NS	ZE	PS	PM
ZE	NM	NM	NS	ZE	PS	PM	PM
PS	NM	NS	ZE	PS	PM	PM	PB
PM	NS	ZE	PS	PM	PM	PB	PB
PB	ZE	PS	PM	PM	PB	PB	PB

The parameters of the fuzzy controller are tuned using coati Optimization. Following the tuning process, the values of the parameters are modified from their original values, resulting in a change in their overall shape. The figures presented in Figure 5. (a) - 5. (c) illustrate the updated fuzzy logic membership functions, which have been modified to encompass a new range. These modifications have been made to facilitate the identification of the minimum error through coati optimization algorithm.

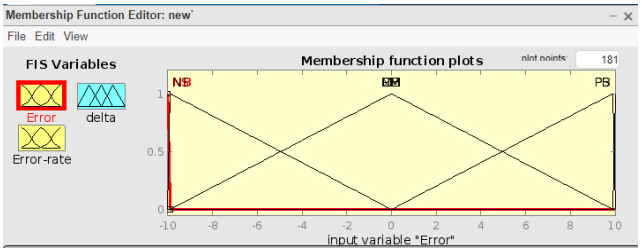

Figure 5. (a): Coati optimised membership function for input E

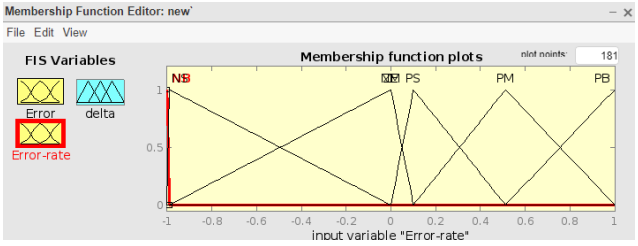

Figure 5 (b): Coati optimised membership function for input ΔE

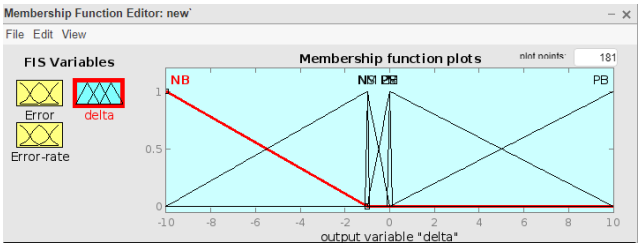

Figure 5 (c): Coati optimised membership function for output $\Delta\theta$

Simulation Results

The effectiveness of the proposed EDRL-MPPT technique for inter-harmonics elimination and COA-Fuzzified PLL synchronization approach, as illustrated in Figure 1, has been evaluated and simulated. The simulation has been conducted in three distinct cases as follows:

- The photovoltaic (PV) inverter functions under a consistent level of solar irradiance, specifically known as steady-state maximum power point tracking (MPPT), which corresponds to its rated power, equivalent to 250 kW. It is worth noting that during this operating condition, the emission of inter-harmonics is particularly noticeable. In the first case, this study examines the effect of inter-harmonics during (P & O) Maximum Power Point Tracking (MPPT) technique and conventional Phase-Locked Loop (PLL) synchronization technique in a three-phase grid-connected solar photovoltaic (PV) system.
- In second case, the implementation of the EDRL-MPPT technique and conventional PLL synchronization technique has been proposed for the purpose of inter-harmonics mitigation in a three-phase grid connected solar PV system.
- In the third scenario, it is suggested that the abrupt load variation results in the generation of a DC offset in the grid current. In order to address the disruption caused by this DC offset in the grid current, the EDRL-MPPT technique with COA-Fuzzified PLL synchronization approach have been proposed.

As discussed, introduction section, due to partial shedding condition, the transient response of the dc-link voltage controller is identified as one of the sources of inter-harmonics in the grid current. To mitigate inter-harmonics in the PV inverter, it is imperative to prevent perturbations in the dc-link voltage during operation. The attainment of this objective is facilitated through the utilization of an (EDRL-MPPT). The present study utilizes an ensembled deep reinforcement learning approach that incorporates a reward system based on total harmonic distortion in grid current. The objective is to effectively train the agent to execute actions that based on the minimal harmonic distortion in grid current. Consequently, the inter-harmonics can be effectively circumvented. The figure 6 illustrates a comparison between Perturb and Observe MPPT and EDRL-MPPT techniques in terms of their impact on DC coupling voltage output and inter-harmonic mitigation in waveform.

In the conventional case, to achieve optimal power extraction from photovoltaic system, an MPPT algorithm such as Perturb and Observe (P&O) is utilized to ascertain the reference DC-link voltage (i.e., PV voltage) during operation. Subsequently, the dc-link voltage (V_{dc}) regulation is achieved by the dc-link voltage controller utilizing a PLL controller. This controller operates by controlling the grid current. The outcome denoted as Vdc, obtained through the employment of the normal MPPT technique with conventional phase-locked loop control, is depicted in Figure 6. The Perturb and Observe (P&O) technique was employed to evaluate the (MPPT) operation in the initial scenario. However, it is important to note that the injected grid current that is linked to this operation displays a considerably higher level of distortion, as illustrated in Figure 7. The disparity in waveform can also be observed in a magnified plot ranging from 0.1 to 0.2 seconds.

The diagram in Figure 8 displays the frequency spectrum of the grid current with total harmonic distortion of 8.90%. It is observed that the inter-harmonic level exhibits a higher absolute value when the PV inverter is functioning at its rated power. The inter-harmonic emission originating from the photovoltaic inverter holds considerable significance in this scenario. In order to mitigate the presence of harmonics, we have put forth the

EDRL-MPPT technique, which incorporates a reward and penalty system based on the distortion observed in the grid current. The subsequent section of our discussion pertained to the MPPT-EDRL system, which incorporates PLL synchronization.

The second scenario involves using EDRL-MPPT to attain maximum power output from the photovoltaic system, aiming to determine the reference DC-link voltage (i.e., PV voltage) during its operation. The direct current (DC) link voltage, denoted as Vdc, is regulated by utilizing a phase-locked loop (PLL) controller by the DC-link voltage controller. Figure 6 (a) illustrates the system's response to a decrease in irradiance from 1000 to 0 W/m2 in the presence of a load. The EDRL-MPPT sustains a constant output of dc voltage with less waveform distortion compared to normal MPPT. At a time, interval of 0.55 seconds to 1 sec, there was a significant decrease in solar irradiance to a value of 0 W/m2.

Figure 6 (b) illustrates the grid current for the proposed (EDRL-MPPT) system. Figure 6 (c) depicts the amplitude of the primary frequency constituent of grid current. After analyzing the frequency spectrum of the output current depicted in Figure 6 (c), it becomes apparent that the dominant inter-harmonics present in the output current exhibit a significant reduction in magnitude. The total harmonic distortion is 7.069%, measured in this case. Additionally, we clearly see less damping in the angular frequency compared to the conventional MPPT in Figure 6 (d). The utilization of an EDRL MPPT in conjunction with the PLL method demonstrates reduced inter-harmonic levels. A comparative analysis has been conducted to assess the phase error between the Perturb and Observe (P&O) MPPT algorithm and the EDRL MPPT algorithm with a conventional Phase-Locked Loop in typical operational circumstances. The EDRL MPPT with conventional PLL generates precise control signals to accurately track the phase angle of the grid and achieve less phase error between the output signal and that of a reference signal which can be seen in Figure 6 (e).

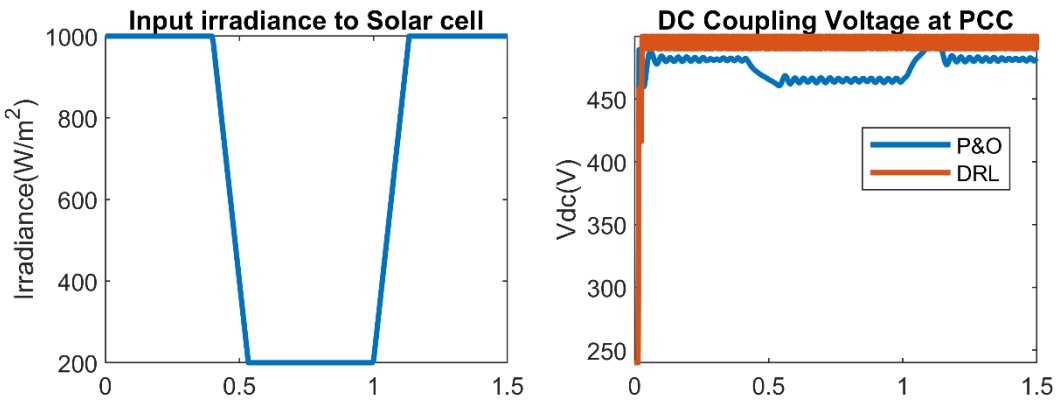

Fig 6 (a), The output of DC coupling voltage of (P&O) and EDRL

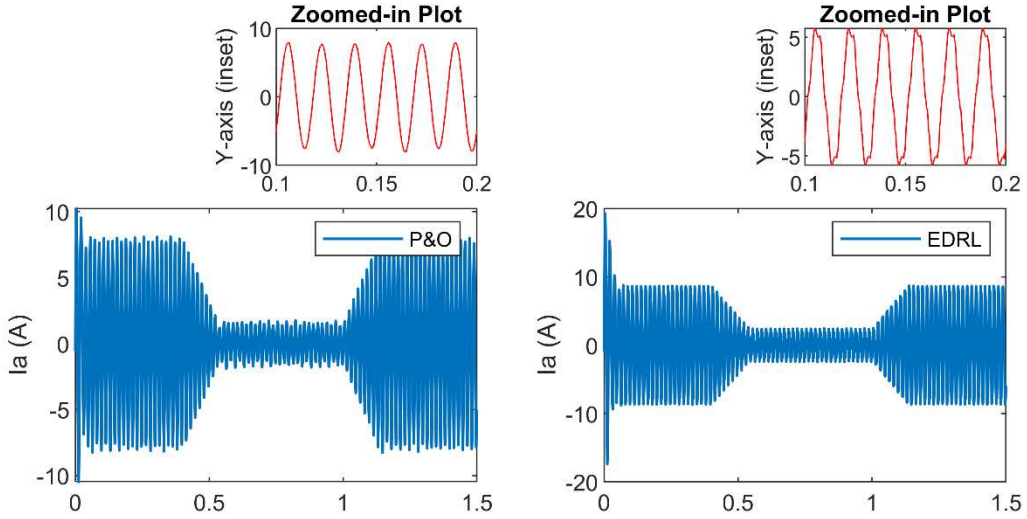

Fig 6 (b), The output of grid current during (P&O) and EDRL MPPT Technique

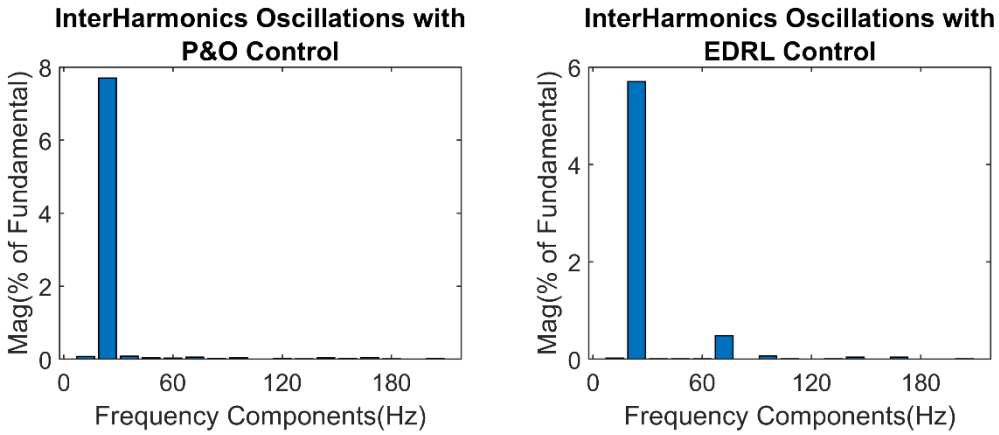

Fig 6 (c), The magnitude of fundamental frequency component during (P&O) and EDRL MPPT Technique

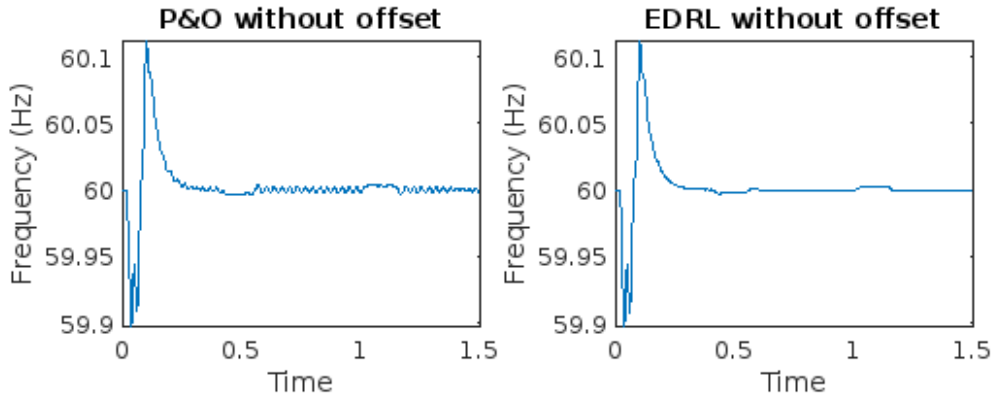

Fig 6 (d), The frequency variation during (P&O) and EDRL MPPT Technique with conventional PLL

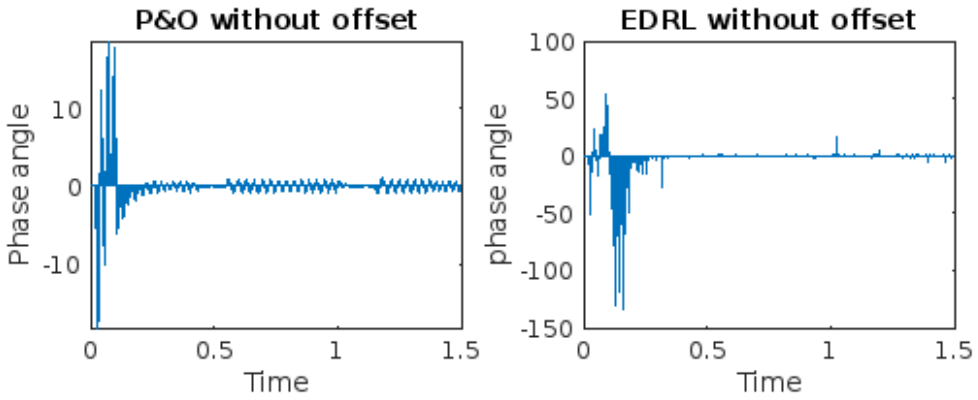

Fig 6 (e), The Phase angle error during (P&O) and EDRL MPPT Technique with conventional PLL

The third case utilizes the EDRL-MPPT and CAO-Fuzzified PLL for grid synchronization. Implementing a control algorithm in a grid-tied photovoltaic system, which aims to eliminate harmonic distortion and compensate DC offset, necessitates the calculation of both the synchronizing signal and the amplitude of the grid current. At the start of operation, the grid functions at its maximum operational capacity. The distribution of load current is divided between the grid and photovoltaic (PV) system based on the amount of solar radiation present. In the event of a sudden reduction in load, the surplus power generated by the photovoltaic (PV) system will be transmitted to the electrical grid. At full load capacity, the grid voltage and grid current exhibit a phase alignment. When there is a sudden decrease in load, it results in a phase shift. This phase shift is mitigated using fuzzy-PLL, and the hyperparameters of fuzzy are tuned using COA results mitigation of DC offset and less harmonics distortion in the grid current as seen in Figure 7 (a).

The mitigation of phase shift is achieved through the utilization of fuzzy-PLL, wherein the hyperparameters of the fuzzy system are adjusted using the COA technique. This approach effectively reduces the presence of DC offset and minimizes the distortion of harmonics in the grid current, as depicted in Figure 7 (a). The respective magnitude of the fundamental frequency component can be seen in Figure 7 (b), with less reduction in total harmonic distortion of 2.89%. The effectiveness of the proposed Fuzzified-PLL under the change in load is depicted in Figure 7 (b). The observation of Figure 7 (b) reveals that the introduction of an offset in the system within the time range of 0.28 to 0.33 seconds results in significant frequency distortion in the conventional phase-locked loop (PLL).

In contrast, the fuzzified-PLL exhibits superior performance during this period. The proposed method has superior performance in various aspects, including dc offset rejection, grid synchronization, inter-harmonic rejection, and stable V_{dc}. A comparison of the two approaches was performed to evaluate the phase error between the Perturb and Observe (P&O) technique with conventional PLL and EDRL MPPT algorithm with a COA-Fuzzified PLL under offset conditions introducing from 0.1 to 0.3 seconds. The EDRL MPPT with COA-Fuzzified PLL is capable of producing highly accurate control signals for effectively monitoring the phase angle of the grid. This results in minimal phase error between the output signal and a reference signal, as depicted in Figure 7 (c).

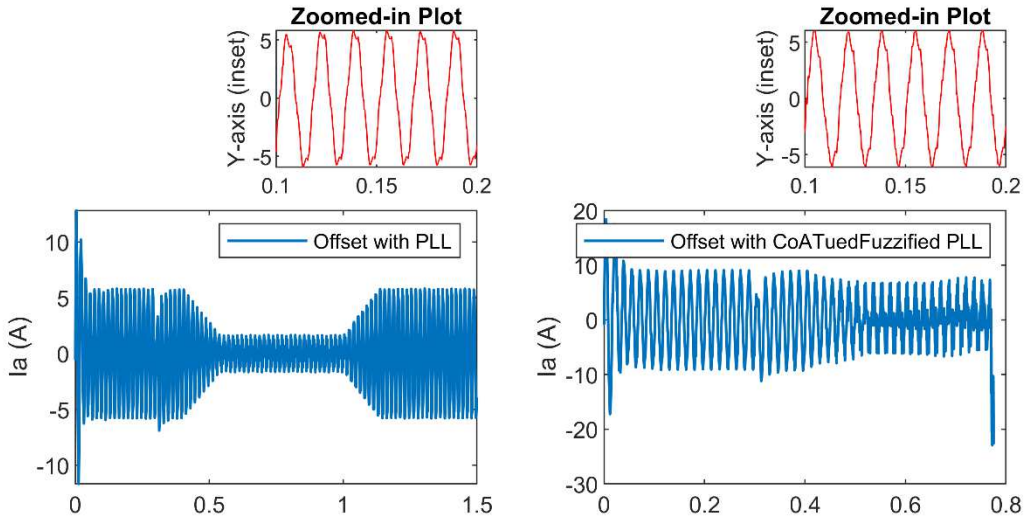

Fig 7 (a), The output of grid current during EDRL-COA Fuzzified PLL synchronization technique

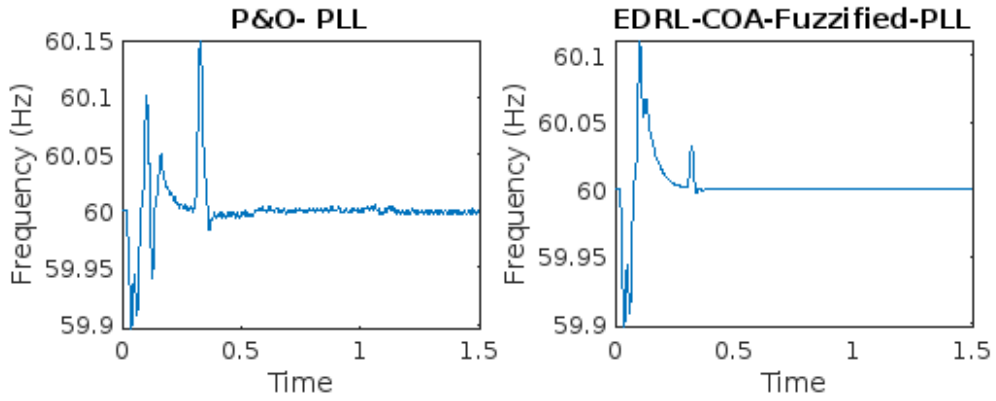

Fig 7 (b), The frequency variation during EDRL MPPT with COA-Fuzzified PLL Technique

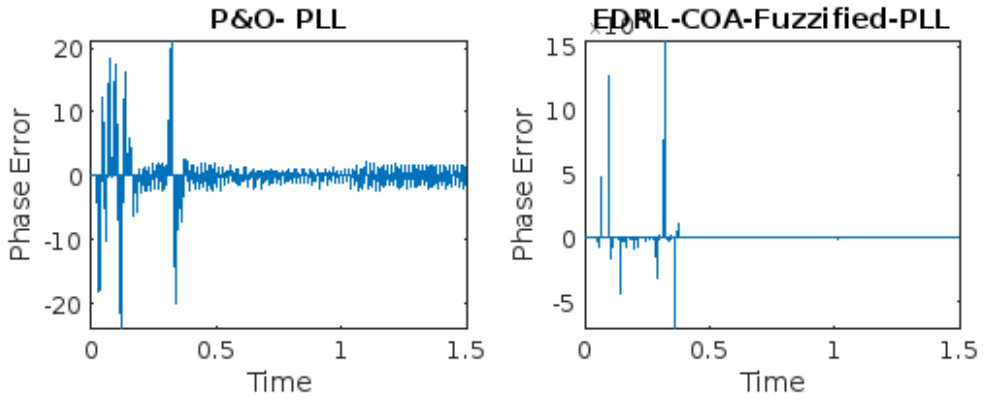

Fig 7 (c), The Phase angle error during EDRL MPPT Technique with COA-Fuzzified PLL

4.1 State-of-art comparison

Table 5 displays a comparative analysis of the proposed COA-fuzzified-PLL controller alongside existing control techniques and a limited number of adaptive control techniques. The provided comparative table illustrates the efficacy of the proposed hybrid control scheme with respect to multiple attributes such as rejection of DC offset and THD of grid current. The proposed scheme demonstrates superior performance compared to existing control algorithms such as MCCF (Multiple complex coefficient filter), MCCF-SOGI (Multiple complex coefficient filter second-order generalized integrator) and fuzzy logic proportional–integrator–derivative multiple complex coefficient filter multiple second-order generalized integrator frequency-locked loop (FLPID-MCCF-MSOGI-FLL) as evidenced by the results presented in Table 5.

Traditional MPPT algorithms with conventional PLL synchronization techniques often exhibit overshoot in estimated phase angle, temporary deviations from desired values during transient conditions. The improved power quality approach employs ensemble DRL MPPT control and COA-fuzzified-PLL-based synchronizing system, which has less overshoot through adaptive control strategies and intelligent decision-making. Continuously optimizing control parameters based on real-time system dynamics, these advanced techniques minimize overshoot, resulting in enhanced power quality and stability.

Similarly, settling time measures the duration for the system's output to reach and remain within a specified range after a disturbance or change in operating conditions. Traditional methods tend to have longer settling times due to slower convergence or suboptimal control actions. In contrast, the ensemble DRL MPPT controller and COA-fuzzified-PLL-based synchronizing system reduce settling time through intelligent learning algorithms and adaptive control strategies. These advanced techniques facilitate faster convergence and more accurate tracking of desired operating points, leading to reduced settling time and improved power quality performance.

Table 5. The summary of performance of Different controller

Features	**MCCF [24]**	**MCCF-SOGI [24]**	**FLPID-MCCF-MSOGI-FLL [24]**	**Conventional PLL**	**Proposed COA-Fuzzified-PLL**
Oscillation	Less	Less	Less	Less	Less
DC Offset Rejection	No	No	Better	No	Good
Inter-Harmonics Removal	No	Yes	Yes	No	Yes
THD Of Grid Current	Less	Less	Better	No	Good
Steady State Performance	Good	Good	Good	Good	Good
Transient Performance	Good	Good	Good	Good	Good
Grid Synchronization	Yes	Yes	Yes	No	Yes
Overshoot	-	-	-	44.03	0
Settling Time	-	-	-	71ms	68ms

Conclusion

In conclusion, this research study aimed to improve the power quality of three-phase grid-connected photovoltaic (PV) systems by addressing inter-harmonics and DC offset. This was achieved by employing the Enhanced Droop Regulator Control (EDRL) Maximum Power Point Tracking (MPPT) technique and the COA-fuzzified-Phase-Locked Loop (PLL) synchronization system. Extensive simulations have revealed that the integration of the EDRL MPPT technique and the COA-fuzzified-PLL synchronization system yields a substantial improvement in power quality within three-phase grid-connected photovoltaic (PV) systems. The reduction of inter-harmonics and DC offset leads to a decrease in voltage distortion, an enhancement in system stability, and an improvement in grid integration. Future research and development endeavours may prioritize the optimization of control strategies, the exploration of advanced algorithms, and the consideration of real-time implementation in order to augment the performance of the proposed techniques in practical contexts.

Declarations

Ethical Approval

This declaration is not applicable.

Competing interests

Both authors hereby declare that they have no competing interests to disclose in any of the cases associated.

Authors' contributions

Both authors had equivalent contributions in the work presented. All decisions vis-à-vis conceptual as well as diplomatic or technical aspects were taken in mutual consent and thus reflect equal contribution and accountability.

Funding

This declaration is not applicable.

Availability of data and materials

The data and materials can be made available on request if there are no confidentiality restrictions associated to the same.

References

[1]. Sridharan, K. and Babu, B.C., 2021. Accurate phase detection system using modified SGDFT-based PLL for three-phase grid-interactive power converter during interharmonic conditions. *IEEE Transactions on Instrumentation and Measurement, 71,* pp.1-11.

[2]. Panwar, N.L., Kaushik, S.C. and Kothari, S., 2011. Role of renewable energy sources in environmental protection: A review. *Renewable and sustainable energy reviews, 15(3),* pp.1513-1524.

[3]. Lubura, S., Šoja, M., Lale, S.A. and Ikić, M., 2014. Single-phase phase locked loop with DC offset and noise rejection for photovoltaic inverters. *IET Power Electronics, 7(9),* pp.2288-2299.

[4]. Liu, B., An, M., Wang, H., Chen, Y., Zhang, Z., Xu, C., Song, S. and Lv, Z., 2020. A simple approach to reject DC offset for single-phase synchronous reference frame PLL in grid-tied converters. *IEEE Access, 8,* pp.112297-112308.

[5]. Han, Y., Luo, M., Zhao, X., Guerrero, J.M. and Xu, L., 2015. Comparative performance evaluation of orthogonal-signal-generators-based single-phase PLL algorithms—A survey. *IEEE Transactions on Power Electronics, 31(5),* pp.3932-3944.

[6]. Saxena, H., Singh, A. and Chittora, P., 2023. Modified LMS synchronization technique for distributed energy resources with DC-offset and harmonic elimination capabilities. *ISA transactions, 135,* pp.567-574.

[7]. Pandey, R. and Kumar, N., 2023, March. Advanced TOGI Controller for Weak Grid Integrated Solar PV System. *In 2023 IEEE IAS Global Conference on Renewable Energy and Hydrogen Technologies (GlobConHT)* (pp. 1-6).

[8]. Saxena, H., Singh, A. and Rai, J.N., 2021. Analysis of SOGI-ROGI for synchronization and shunt active filtering under distorted grid condition. *ISA transactions, 109,* pp.380-388.

[9]. Xie, M., Huiqing wen. Canyan Zhu and Yong Yang,"DC offset rejection improvement in single-phase SOGI-PLL algorithms: methods review and experimental evaluation" *IEEE Access, 5.*

[10]. Saxena, H., Singh, A. and Rai, J.N., 2019, November. Design and Analysis of Cascaded Generalized Integrators for Mitigation of Power Quality Problems. *In 2019 International Symposium on Advanced Electrical and Communication Technologies (ISAECT)* (pp. 1-6). IEEE.

[11]. Punitha, K., Devaraj, D. and Sakthivel, S., 2013. Development and analysis of adaptive fuzzy controllers for photovoltaic system under varying atmospheric and partial shading condition. *Applied Soft Computing, 13(11),* pp.4320-4332.

[12]. Liu, L., Liu, C., Wang, J. and Kong, Y.G., 2015. Simulation and hardware implementation of a hill-climbing modified fuzzy-logic for mppt with direct control method using boost converter. *Journal of Vibration and Control., 21(2),* pp.335-342.

[13]. Kermadi, M., Salam, Z., Ahmed, J. and Berkouk, E.M., 2018. An effective hybrid maximum power point tracker of photovoltaic arrays for complex partial shading conditions. *IEEE Transactions on Industrial Electronics, 66(9),* pp.6990-7000.

[14]. Rizzo, S.A. and Scelba, G., 2015. ANN based MPPT method for rapidly variable shading conditions. *Applied Energy, 145,* pp.124-132.

[15]. Singh, Y. and Pal, N., 2021. Reinforcement learning with fuzzified reward approach for MPPT control of PV systems. *Sustainable Energy Technologies and Assessments, 48,* p.101665.

[16]. Xia, Z., Wu, J., Wu, L., Yuan, J., Zhang, J., Li, J., & Wu, D. (2021, July). RLCC: Practical Learning-based Congestion Control for the Internet. In *2021 International Joint Conference on Neural Networks (IJCNN)* (pp. 1-8). IEEE.

[17]. Lan, Q., Pan, Y., Fyshe, A., & White, M. (2020). Maxmin q-learning: Controlling the estimation bias of q-learning. *arXiv preprint arXiv:2002.06487*.

[18]. Anschel, O., Baram, N., & Shimkin, N. (2017, July). Averaged-dqn: Variance reduction and stabilization for deep reinforcement learning. In *International conference on machine learning* (pp. 176-185). PMLR.

[19]. Liu, X. Y., Yang, H., Chen, Q., Zhang, R., Yang, L., Xiao, B., & Wang, C. D. (2020). FinRL: A deep reinforcement learning library for automated stock trading in quantitative finance. *arXiv preprint arXiv:2011.09607*

[20]. Liu, Q., Cheng, L., Jia, A. L., & Liu, C. (2021). Deep reinforcement learning for communication flow control in wireless mesh networks. *IEEE Network*, *35*(2), 112-119.

[21]. Yang, H., Liu, X. Y., Zhong, S., & Walid, A. (2020, October). Deep reinforcement learning for automated stock trading: An ensemble strategy. In *Proceedings of the First ACM International Conference on AI in Finance* (pp. 1-8).

[22]. Zhu, J., Wu, F., & Zhao, J. (2021, December). "An Overview of the Action Space for Deep Reinforcement Learning", In *2021 4th International Conference on Algorithms, Computing and Artificial Intelligence* (pp. 1-10).

[23]. Dehghani, M., Montazeri, Z., Trojovská, E. and Trojovský, P., 2023. Coati Optimization Algorithm: A new bio-inspired metaheuristic algorithm for solving optimization problems. *Knowledge-Based Systems, 259,* p.110011.

[24]. Babu, N., Guerrero, J.M., Siano, P., Peesapati, R. and Panda, G., 2020. An improved adaptive control strategy in grid-tied PV system with active power filter for power quality enhancement. *IEEE Systems Journal, 15(2),* pp.2859-2870.