Journal of Basics and Applied Sciences Research (JOBASR)
ISSN (print): 3026-9091, ISSN (online): 1597-9962
Volume 4(4) July 2026
DOI: https://dx.doi.org/10.4314/jobasr.v4i4.4

Partial Shading Mitigation on DC link Stability based on Deep Reinforcement Learning Grid-Forming Inverter Control

Ameze Big-Alabo[1*]
[1]Department of Electrical/Electronic Engineering, Faculty of Engineering, University of Port Harcourt, Nigeria
[*]Corresponding Author Email: ameze.odia@uniport.edu.ng

ABSTRACT
Partial shading on photovoltaic (PV) arrays induces nonlinear power–voltage characteristics that cause rapid, multi-peaked fluctuations in DC-link power injection. Conventional grid-forming (GFM) inverter controllers, which rely on fixed-gain PI regulators tuned around a nominal operating point, struggle to track these fast, nonlinear disturbances: their linear control action cannot adapt quickly enough to abrupt irradiance changes, leading to sluggish transient response and, under severe shading, DC-link voltage collapse. This paper proposes a hybrid deep reinforcement learning (DRL) based GFM inverter control strategy that replaces these fixed-gain regulators with adaptive, learning-based controllers to overcome this limitation. A Deep Deterministic Policy Gradient (DDPG) agent governs the DC–DC boost converter for maximum power point tracking (MPPT), while a Twin Delayed Deep Deterministic Policy Gradient (TD3) agent replaces the conventional inner current PI regulators of the droop-controlled GFM inverter, enabling the controller to continuously learn and adjust its response to nonlinear, time-varying shading conditions rather than relying on static gains. Performance is evaluated under four irradiance scenarios, from uniform irradiation to severe partial shading, and compared against a conventional PI–PI GFM controller. Results show that the proposed GFM (PI–TD3) controller achieves faster DC-link voltage stabilization and superior voltage regulation under all shading conditions, maintaining the DC-link at 779.1 V under severe shading compared to a collapse to 290 V with the conventional controller. These findings indicate that TD3-based DRL GFM inverter control offers a robust, adaptive solution for standalone PV systems under variable irradiance conditions.

Keywords:
Deep Reinforcement Learning;
DC – Link stability;
Grid forming inverters;
Maximum Power Point Tracking;
Photovoltaic systems

INTRODUCTION

There has been significant increase in photovoltaic (PV) systems in the last few decades, especially in standalone and off-grid configurations. This has led to the need for robust power conversion and control architectures capable of maintaining voltage and frequency stability without reliance on the utility grid (Liserre et al., 2006). A key challenge with PV system is partial shading whereby portions of the portions of the PV array receive non uniform solar irradiance due to passing clouds or surrounding structures. Partial shading induces nonlinear power–voltage (P–V) characteristics that produces multiple local maxima on the power curve that can cause instability in the DC-link (Shahgholian et al., 2025). As a result, power oscillations propagate to the inverter stage, thereby degrading the quality of the AC output voltage. It can further destabilize reactive power sharing among parallel units and ultimately compromise the reliability of energy delivery to sensitive local loads (Naderipour et al., 2018).

Mitigating the effects of partial shading on DC-link stability therefore requires two complementary solutions which are an effective maximum power point tracking (MPPT) strategy that resolves the global maximum under multi-peak P–V characteristics, and a robust inverter control framework that actively regulates DC-link voltage. MPPT algorithms such as perturb and observe, incremental conductance and metaheuristic variants (Hassan et al., 2026; Mohanty et al., 2016), including adaptive P&O variants with improved convergence and tracking accuracy (Majia et al., 2026), have received considerable research attention. However, the control of the inverter stage under shading induced DC disturbances remains comparatively underexplored, particularly in islanded configurations.

Existing approaches to GFM control under partial shading can be grouped into four broad methodological categories, each with distinct unresolved limitations for adaptive DC-link regulation.

Classical model-based GFM control. Droop control (Shahgholian et al., 2025), Lyapunov-based stability design (Dasgupta et al., 2013), and dq-frame transient response strategies (Ninad & Lopes, 2012) established the theoretical foundation for autonomous voltage/frequency formation (Rocabert et al., 2012). However, these methods rely on linearized small-signal models tuned around a fixed operating point. Their gains are not re-derived online and so cannot track the abrupt, multi – peaked power excursions that partial shading induces on the DC side.

Multi-loop and feedforward-enhanced control. Cascaded current/voltage loops with reference feedforward (Lei et al., 2010; Lakshmanan et al., 2014) and adaptive neural feedforward for online impedance estimation (Chand et al., 2022; Gutiérrez-Escalona et al., 2024) improved disturbance rejection and steady-state accuracy. These schemes still depend on an accurately identified plant model or impedance estimate. Under fast – changing shading, the estimation lag limits how quickly the feedforward term can compensate, so transient DC-link excursions persist.

Inertia-emulation paradigms (VSM/VES/VSG/dVOC). Virtual Synchronous Machine (Lin et al., 2022), Virtual Energy Storage (Yang et al., 2022), VSG, and dVOC (Jia et al., 2021; Colombino et al., 2019) approaches emulate synchronous-machine inertia to improve frequency/voltage support (Yazdani et al., 2020). Their inertia and damping coefficients, however, are again fixed design parameters which were developed primarily for AC-side disturbances such as grid faults and are not explicitly designed to regulate the DC – link against irradiance driven power fluctuations.

Data-driven and DRL-based control. Model – based limitations under parameter variation and unpredictable disturbances such as partial shading (Khan et al., 2024) have motivated learning-based alternatives. DRL agents that learn control policies through direct environment interaction, without an explicit plant model (Gutiérrez-Escalona et al., 2024), have shown promise for voltage regulation, frequency support, and fault-ride-through (Zhang et al., 2021; Cao et al., 2020). Yet this line of work has largely targeted AC-side objectives; DRL has not been systematically applied to *coordinate* DC – link regulation with MPPT under rapidly varying partial shading. DRL formulations that handle both the boost-converter and inverter-current control simultaneously remain unexplored.

Across all four categories, the common shortcoming is the same: none provide an online-adaptive mechanism that jointly resolves the multi-peak MPPT problem *and* suppresses the resulting DC-link disturbance before it propagates to the AC side. This gap motivates a hybrid, dual-agent DRL architecture that treats MPPT, and inverter current regulation as coupled control problems.

There are varying DRL algorithms, and the choice of application is critical for this present study. For example, discrete-action methods algorithms such as the DQN are unsuitable because the converters require continuous duty-cycle and modulation-index outputs. Among continuous-control actor–critic algorithms, DDPG is well suited to the slower MPPT task, where its single critic architecture is sufficient for tracking gradually varying operating points. However, DDPG is prone to Q – value overestimation and training instability that makes it less suitable for the inner current control loop, where accurate regulation directly affects DC-link voltage quality. TD3 mitigates these issues through twin critics, delayed policy updates, and target policy smoothing, providing more stable and reliable value estimates under fast, noisy transients. Consequently, DDPG is adopted for MPPT, while TD3 is selected for the faster, stability – critical inner current – control task [Fujimoto et al., 2018].

This paper proposes a DRL-based GFM inverter control strategy specifically designed to mitigate the impact of partial shading on DC-link stability in standalone PV systems. A droop controlled GFM inverter provides the AC voltage and frequency reference, while a DRL agent augments the conventional control loop to actively suppress DC – link voltage oscillations arising from partial shading. The key contributions of this work are as follows:

- A DRL augmented GFM control architecture that maintains DC-link voltage stability and AC voltage quality under partial shading conditions without requiring prior knowledge of shading patterns.
- Comparative evaluation against conventional GFM controllers, thereby demonstrating improved transient performance and steady-state regulation under dynamic partial shading profiles.

The remainder of this paper is organized as follows. Section II presents the system overview and describes the standalone PV system model including the PV array, DC–DC boost converter, and three-phase GFM inverter. Section III presents the proposed DRL – GFM control framework, simulation results and discussion under four partial shading scenarios, comparing the proposed DRL – GFM control framework against conventional GFM controller. Section IV concludes the paper.

MATERIALS AND METHODS

Nomenclature

System & DC-Link	**Description**	**Unit**
V_{dc}	DC-link voltage	V
$V_{dc,ref}$	DC-link voltage reference	V
V_{oc}	PV module open-circuit voltage	V
V_{mp}	PV module voltage at maximum power point	V
V_{nom}	Nominal AC phase voltage (rms)	V
V_{LL}	Nominal AC line-to-line voltage (rms)	V
V_{peak}	Peak AC output voltage	V
I_{sc}	PV module short-circuit current	A
I_{mp}	PV module current at maximum power point	A
P_{rated}	Rated power of the PV system	W
P_{pv}	Instantaneous power supplied by the PV array	W
P_{inv}	Instantaneous power drawn by the inverter	W
P	Active power at the PCC	W
P_{ref}	Active power reference (droop control)	W
Q	Reactive power at the PCC	var
Q_{ref}	Reactive power reference (droop control)	Var
PV Array		
N_{ser}	Number of PV modules in series per string	—
N_{par}	Number of PV strings in parallel	—
N_{cell}	Total number of cells per module	—
P_{mp}	PV array maximum power point power	W
DC/DC Boost Converter		
D	Boost converter duty cycle	—
V_{pv}	PV array terminal voltage	V
i_L	Boost converter inductor current	A
L_b	Boost converter inductance	H
R_L	Boost converter inductor internal resistance	Ω
C_{dc}	DC – link capacitance	F
i_{conv}	DC current drawn by the inverter from the DC link	A
GFM Inverter and Droop Control (dq-Frame)		

i_d ,i_q	dq-axis inverter output currents	A
v_d ,v_q	dq-axis inverter-side voltages	V
V_{od} ,V_{oq}	dq-axis PCC output voltages	V
i_{qref}	q-axis current reference	A
u_d, u_q	Normalised dq-axis PWM voltage modulation commands	—
θ	Internal voltage angle (GFM reference frame)	rad
ω	Angular frequency of the GFM inverter	rad/s
ω_0	Nominal angular frequency ($2\pi \times 50$ rad/s)	rad/s
V_0	Nominal output voltage amplitude reference	V
K_P	Active power droop coefficient	—
K_Q	Reactive power droop coefficient	—
R	LCL filter resistance	Ω
L	LCL filter inductance	H
L_1, L_2	Inverter-side and load-side LCL filter inductances	H
C_f	LCL filter capacitor	F
R_d	LCL filter damping resistor	Ω
f_s	Inverter switching frequency	Hz
S_{inv}	Inverter rated apparent power	kVA
K_{cp}, K_{ci}	Inner current-loop PI proportional	—
TD3-DRL Agent		
s_t	State vector at time step t	—
a_t	Action vector at time step t	—
r_t	Scalar reward at time step t	—
R_t	Cumulative discounted reward from time step t	—
α	Control effort weighting coefficient in reward function	—
γ	Discount factor ($\gamma = 0.99$)	—
I_{nom}	Nominal inverter current (reward normalization)	A

System Overview

The proposed 100 kW grid-forming photovoltaic, PV system consists of three main stages a PV array and DC–DC boost converter for MPPT; a DC-link capacitor serving as the energy buffer; and a three-phase voltage source inverter, VSI interfaced with the AC load through an LCL filter, as illustrated in Fig 1. The VSI operates in GFM mode generating a controlled three-phase voltage at the point of common coupling, PCC with defined amplitude and frequency.

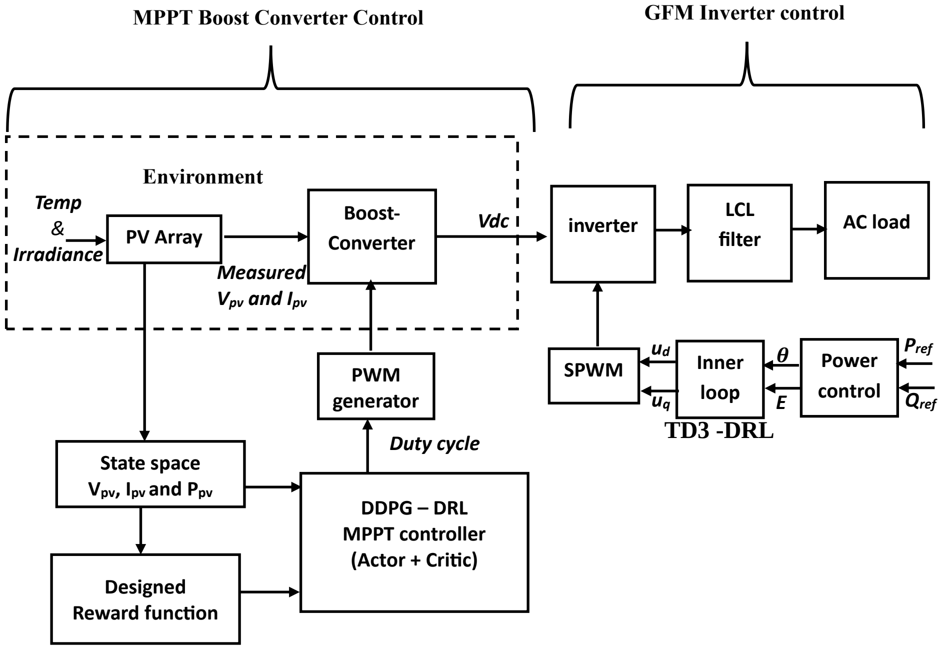

Fig 1: Schematic diagram of hybrid DRL based MPPT and droop-controlled GFM inverter framework

The DC–DC boost converter ensures stable DC-link voltage and extracts the maximum PV power, while the VSI governs AC side power exchange and voltage and frequency regulation. System ratings are as follows: Rated power, $P_{rated} = 100$ kW; AC nominal voltage $V_{LL} = 230$ V (rms), 50 Hz ; DC-link voltage Vdc = 795V; The LCL filter provides current ripple attenuation with parameters: $L_1, L_2 = 5\,mH$; $C_f = 200\,\mu F$. The inverter is switched using sinusoidal pulse-width modulation (SPWM) at 10 kHz. Table 1 describes the complete parameter ratings used.

Table 1: System parameters for standalone PV system

Subsystem	Parameter	Symbol	Value
PV array	Rated Power	P_{pv}	100 kW
	Nominal voltage	V_{pv}	638 V
DC – Link	Nominal Voltage	V_{dc}	795 V
	Capacitance	C_{dc}	20 mF
Inverter	Apparent Power	S_{inv}	$\geq 110\,kVA$
	Voltage	V_{L-L}, V_{RMS}, V_{peak}	400V, 230V, 325 V
	Switching frequency	f_s	10 kHz
	Inner current–loop PI gains	K_{cp}, K_{ci}	30, 200
LCL filter	Inverter/ Load–side inductance	L_1, L_2	5 mH
	Capacitor	C_f	200 μF
	Damping resistor	R_d	0.6 Ω

PV Array and DC–DC Boost Converter Model

The PV array consists of two series connected strings each rated at ~50 kW, while the DC-DC boost converter interfaces the PV panels with ratings of 0.623 mH, and 1888 μF for inductance and capacitance respectively. The PV array is designed based on the manufacturer ratings shown in Table 2 with 28 Modules connected in parallel and 6 modules connected in series for each string.

Table2: Data specification: SunPower SPR-305E-WHT-D

Parameter	**Ratings**
Maximum Power (W)	305.23
Voltage at Maximum Power (Vmp)	54.70
Current at Maximum Power (Imp)	5.58
Open Circuit Voltage (Voc)	64.20
Short circuit Current (Isc)	5.96
Total number of cells in Series (Ns)	96
Total number of cells in parallel (Np)	1

The DC–DC boost converter regulates the PV operating point using a Deep Deterministic Policy Gradient, DDPG MPPT controller, which adjusts the duty ratio D to track the maximum power point, MPP. The converter dynamics are given in Eqns (1) and (2):

$$\frac{di_L}{dt} = \frac{1}{L_b}\left(V_{pv} - (1-D)V_{dc} - R_L i_L\right) \quad (1)$$

$$\frac{dV_{dc}}{dt} = \frac{1}{C_{dc}}\left((1-D)i_L - i_{conv}\right) \quad (2)$$

Where i_L is the inductor current, L_b is the boost inductance and i_{conv} is the converter DC current. The design of the DDPG - DRL agent system for the MPPT control of the 100 kW PV system is described in a previous study (Big – Alabo, 2026).

For all comparative simulations presented in this work, the pretrained DDPG MPPT controller, reward function, network architecture, and training parameters were kept identical. The only modification between the two GFM control configurations was the replacement of the inverter inner current PI controller with the TD3 controller.

Three-Phase Grid-Forming VSI Model

In the dq-reference frame aligned with the converter voltage vector, the averaged VSC equations are given by Eqn (3) and (4):

$$L\frac{di_d}{dt} = v_d - Ri_d + \omega L_{iq} - V_{od} \quad (3)$$

$$L\frac{di_q}{dt} = v_q - Ri_q - \omega L_{id} - V_{oq} \quad (4)$$

where i_d, i_q, are dq-axis currents; v_d, v_q, are inverter-side dq voltages; V_{od}, V_{oq} are PCC voltages; R and L are the filter resistance and inductance; and ω is the angular frequency. The instantaneous active and reactive powers at the PCC are expressed by Eqn (5):

$$P = \frac{3}{2}\left(v_d i_d + v_q i_q\right);\ Q = \frac{3}{2}\left(v_q i_d - v_d i_q\right) \quad (5)$$

The outer control loop defines the voltage amplitude and frequency references using droop equations described by Eqn (6).

$$\omega = \omega_0 - K_P\left(P - P_{ref}\right);\ V = V_0 - K_Q\left(Q - Q_{ref}\right) \quad (6)$$

where $\omega_0 = 2\pi \times 50$ rad/s and $V_0 = 230\, V_{rms}$ are nominal values, and K_P, K_Q are active and reactive power droop coefficients. The updated frequency and voltage references are integrated to yield the internal angle as shown in Eqn (7):

$$\theta(t) = \int_0^t \omega(\tau)d\tau. \quad (7)$$

This internal angle θ defines the dq reference frame for current control, thereby eliminating the need for a phase lock loop, PLL. The droop/VSM outer loop outputs the reference dq currents i_d^{ref} and i_q^{ref} as functions of power and voltage setpoints as shown in Eqns (8) and (9) respectively.

$$i_d^{ref} = \frac{2}{3}\frac{P_{ref}}{v_d} \quad (8)$$

$$i_q^{ref} = \frac{2}{3}\frac{Q_{ref}}{v_d} \quad (9)$$

TD3-Based Adaptive Current Control

DRL controllers are trained to optimize control actions such as current reference tracking directly in response to observed states such as grid conditions or disturbances. They help in stabilizing the transient performance of GFM inverters (Das, S. C, 2020). In the proposed hybrid scheme, the inner dq current control loop traditionally implemented using PI regulators is replaced by a TD3 - DRL agent. The agent observes the system state vector in Eqn (10) and generates the control action in Eqn (11):

$$s_t = \left[i_d^{ref}, i_q^{ref}, i_d, i_q\right] \quad (10)$$

$$a_t = \left[u_d, u_q\right] \quad (11)$$

which corresponds to normalized voltage modulation commands for the PWM modulator in Eqn (12):

$$-1 \le u_d, u_q \le 1 - \quad (12)$$

Reward Function

The reward function encourages accurate current tracking and penalizes excessive control effort. It is defined by Eqn (13):

$$r_t = -\sqrt{\left(\frac{i_d^{ref} - i_d}{I_{nom}}\right)^2 + \left(\frac{i_q^{ref} - i_q}{I_{nom}}\right)^2} - \alpha\sqrt{u_d^2 + u_q^2} \quad (13)$$

where $\alpha = 0.01$ is a control effort weighting coefficient. The agent maximizes the cumulative discounted reward given by Eqn (14)

$$R_t = \sum_{k=0}^{\infty} \gamma^k r_{t+k} \quad (14)$$

Where $\gamma = 0.99$

The TD3 agent was trained offline within the MATLAB/Simulink environment through episodic interactions with the simulated grid-forming inverter as shown in Fig 2.

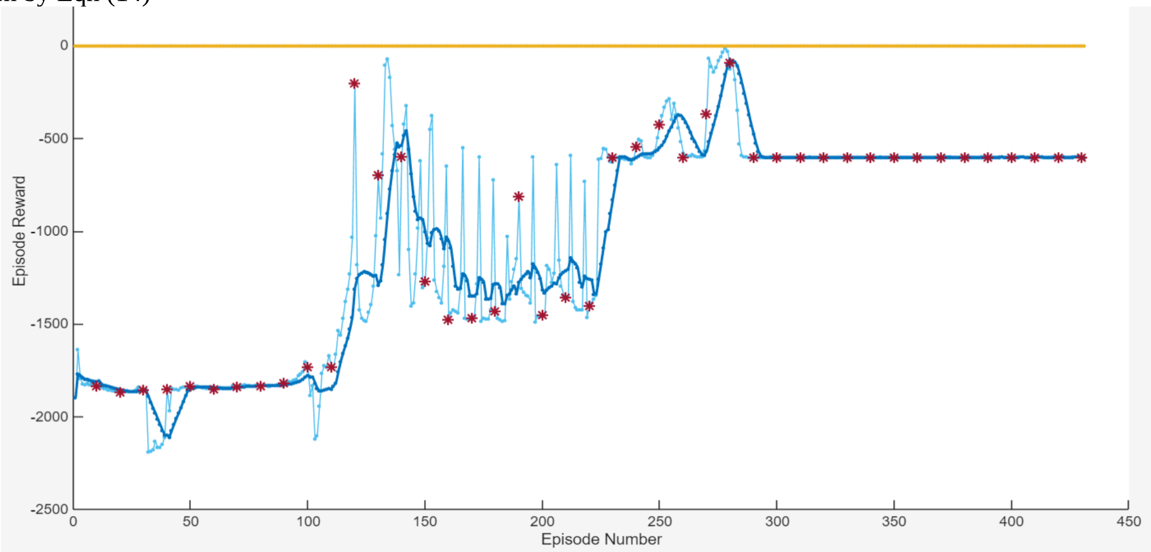

Fig 2: TD3 – DRL Agent Training Plot

The TD3 – DRL agent observed a four-dimensional state vector comprising the dq-axis reference currents and the measured dq-axis currents, as well as generated normalized voltage modulation commands as continuous actions for the PWM modulator. Training progressed over 400 episodes, during which the agent explored the inverter's environment and iteratively refined its policy. As reflected in the training monitor, the agent exhibited an initial phase of poor performance with high reward variance, followed by a period of active policy improvement characterized by significant exploration and reward fluctuations. By approximately episode 300, the agent converged to a stable policy with an average reward of approximately −601 which is an indication of effective current tracking behaviour. The converged TD3 policy was subsequently deployed in the inner current control loop of the grid-forming inverter without further online learning which provides real time voltage modulation commands that replaced the conventional inner current PI regulators for the GFM controller.

Mechanism for DC-link improvement.

It is important to note that the proposed TD3 controller does not directly regulate the DC-link capacitor voltage but rather replaces only the inverter inner current regulator. The DC-link dynamics are governed by the instantaneous power balance between the PV array and the inverter, which may be expressed by Eqn (15)

$$C_{dc}\frac{dV_{dc}}{dt} = \frac{P_{PV} - P_{inv}}{V_{dc}} \quad (15)$$

or equivalently, Eqn (16)

$$\frac{1}{2}C_{dc}\frac{dV_{dc}^2}{dt} = P_{PV} - P_{inv} \quad (16)$$

where P_{PV} is the power supplied by the PV array and P_{inv} is the instantaneous power drawn by the inverter. Therefore, the DC-link voltage remains constant only when the power injected into the DC link is balanced by the power extracted by the inverter.

Thus, the proposed TD3 controller improves the tracking of the d- and q-axis current references during rapid irradiance transients. The TD3 controller enables the inverter to follow the commanded active and reactive power references more accurately than the conventional PI controller. Since the PV - side power is determined primarily by the irradiance level and the MPPT controller, improvements in DC – link regulation arise from more accurate realization of the inverter side power command through improved current tracking.

RESULTS AND DISCUSSION

The GFM controllers are evaluated under four operating scenarios to investigate the impact of partial shading on the DC-link voltage V_{dc} as well as the grid-forming voltage and current signals. Different irradiance conditions are applied to Panel A and Panel B to emulate varying levels of partial shading.

In Case 1, the simulation begins under uniform irradiance conditions of (1000, 1000) W/m^2, representing the irradiance levels on Panel A and Panel B, respectively. Subsequently, degraded irradiance patterns are introduced to represent partial shading conditions as follows: Case 2 (1000, 800) W/m^2, Case 3 (800, 600) W/m^2 and Case 4 (600, 400) W/m^2. For each case, the performance of two GFM control strategies is compared. The first controller is based on the conventional droop control approach with an outer voltage PI control loop and inner current PI control loops. The second controller employs the same GFM structure but replaces the inner current PI control loops with a TD3-based deep reinforcement learning (TD3-DRL) controller.

Case 1: Uniform Irradiation (1000, 1000) W/m^2

Under uniform irradiation conditions where Panel A and Panel B are both operating at 1000 W/m^2, both the GFM (PI–PI) and GFM (PI–TD3) control strategies achieve the expected DC-link voltage of approximately 795 V. This results in a corresponding stable grid-forming inverter voltage and current supplied to the load. For the GFM (PI–PI) control strategy, the simulation plots of the P–V characteristics, DC-link voltage, and the resulting inverter voltage and current supplied to the load are presented in Fig. 3a Although some oscillatory behavior is observed during the initial startup period, the system eventually stabilizes

***B.* GFM (PI – PI) control** ***A.* GFM (PI – TD3) control**

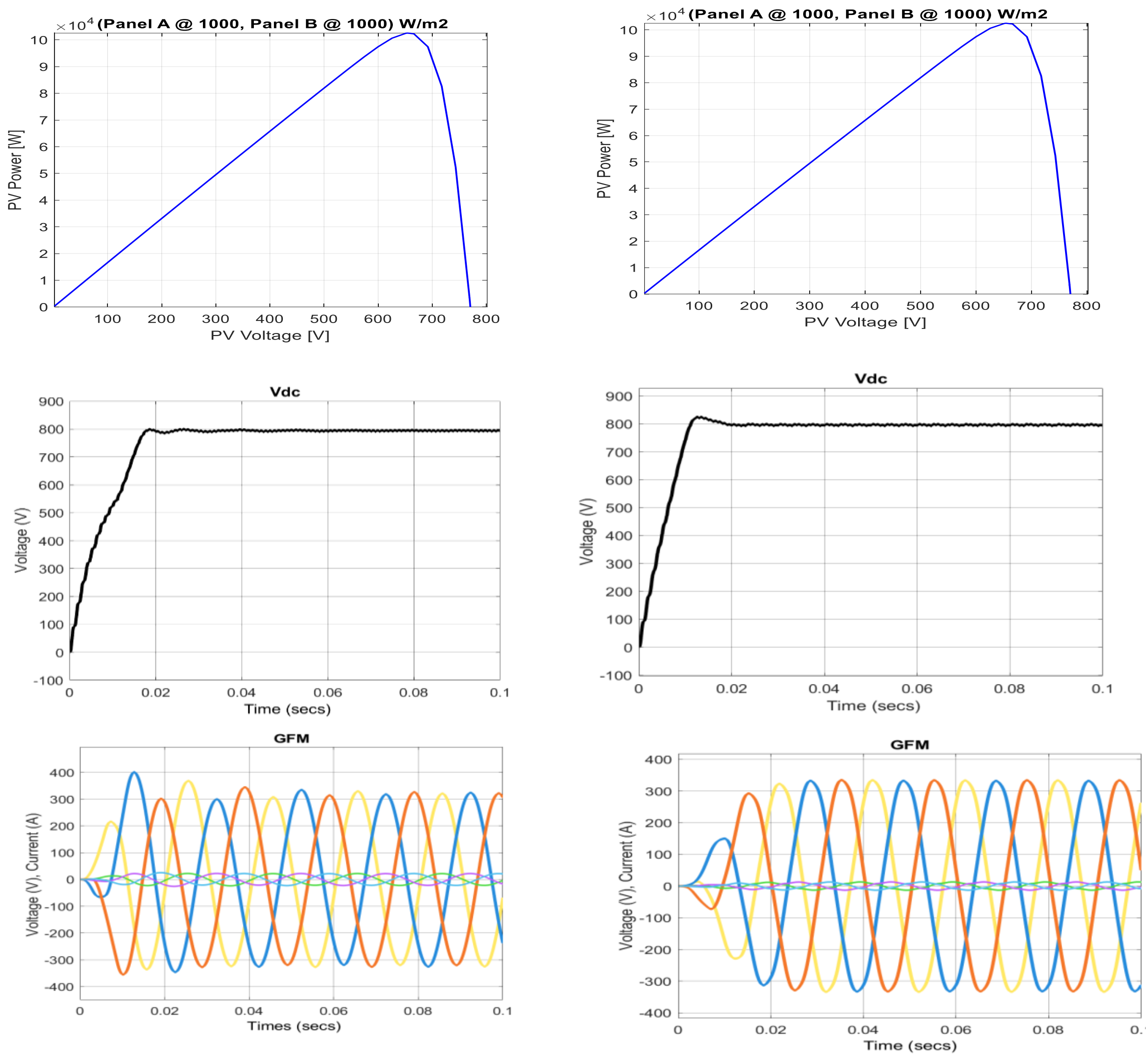

Fig 3: Case 1: for P-V, Vdc, Vac$_{(peak)}$ at uniform irradiance

Based on the simulation results in Fig 3, the DC-link voltage requires approximately 0.016 secs to rise and settle to its steady-state value while the P–V characteristics curve is smooth and reaches the rated peak power of approximately 100 kW.

Similar results are obtained for the GFM (PI–TD3) control strategy, as shown in Fig. 3b. By comparing the conventional GFM (PI–PI) controller with the proposed GFM (PI–TD3) controller, the latter demonstrates a faster and more stable dynamic response with reduced startup oscillations. The DC-link voltage rises and stabilizes in less than 0.01 secs indicating improved transient performance. Furthermore, the P–V characteristics curve remains smooth and reaches the rated peak power of approximately 100 kW confirming the effectiveness of the TD3 – DRL control approach under uniform irradiation conditions.

Case 2: Partial Irradiation (1000, 800) W/m²

Under partial shading conditions where Panel A operates at 1000 W/m² and Panel B operates at 800 W/m², the GFM (PI–PI) control strategy is still able to maintain the expected DC-link voltage of approximately 795V. Consequently, the inverter delivers a stable grid-forming voltage and current to the load. The simulation plots for the P–V characteristics, DC-link voltage, and the resulting inverter voltage and current supplied to the load are presented in Fig. 4a.

A. GFM (PI – PI) control **B. GFM (PI – TD3)**

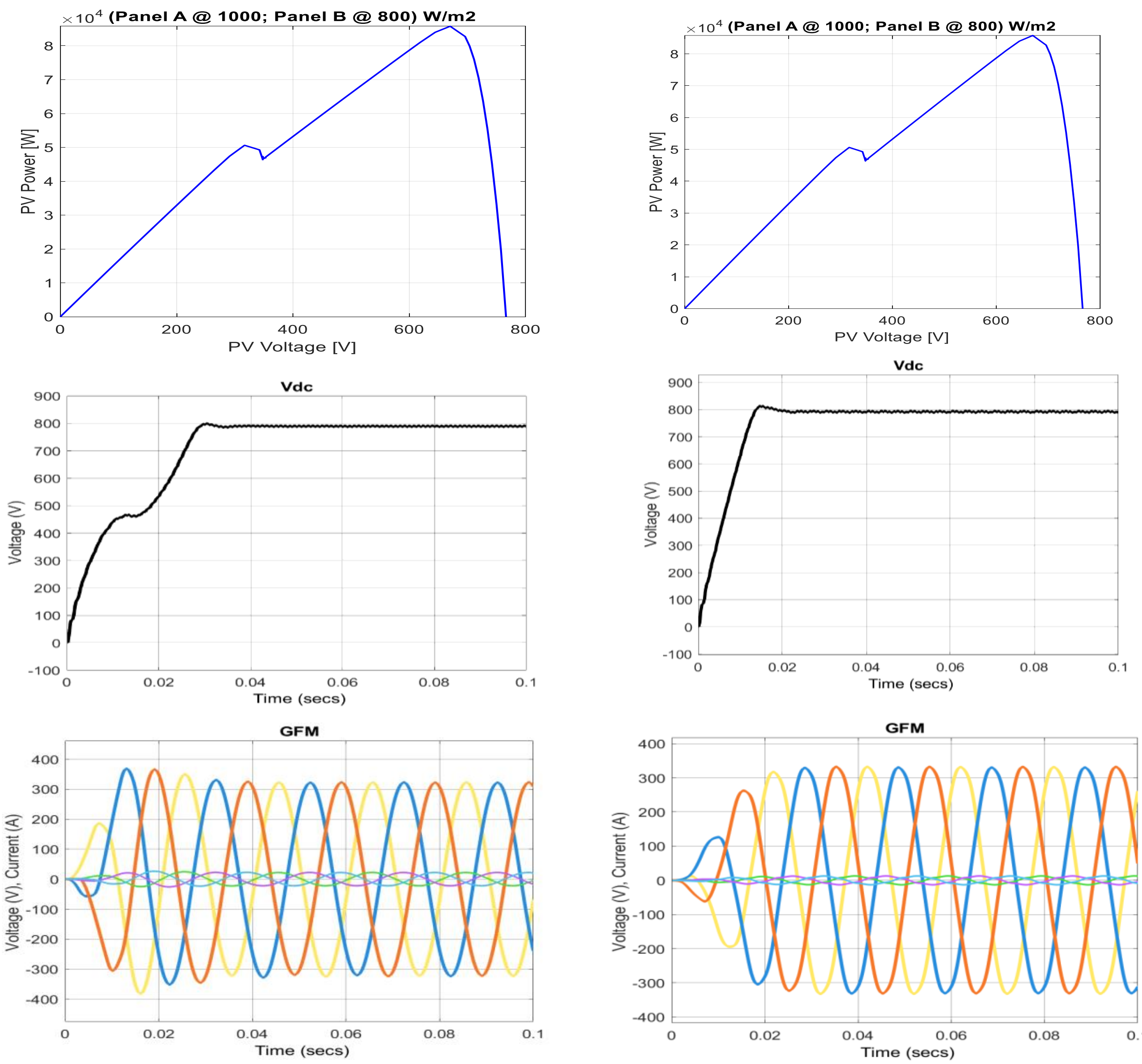

Fig 4: Case 2: P-V, Vdc, $Vac_{(peak)}$ @ 1000, 800 W/m2

However, compared to the uniform irradiation condition, more pronounced oscillations are observed during startup, and the system requires a longer settling time of approximately 0.03 secs before reaching steady-state operation. The influence of partial shading is also evident in the P–V characteristics curve, where the global maximum power point decreases to approximately 85 kW. For the proposed GFM (PI–TD3) control strategy under the same partial shading conditions (1000, 800) W/m^2. The expected DC-link voltage of approximately 795 V is also achieved, but with a faster and more stable dynamic response compared to the conventional GFM (PI–PI) controller. This results in a corresponding stable inverter voltage and current supplied to the load.

The corresponding simulation plots for the P–V characteristics, DC-link voltage, and inverter voltage and current are shown in Fig. 4b. From the results, the system stabilizes in less than 0.02 secs demonstrating improved transient performance and reduced startup oscillations. Although the effect of partial shading causes the global maximum power point to reduce to approximately 85 kW, the overall system performance remains stable due to the robustness and adaptability of the proposed TD3-DRL-based control architecture.

Case 3: Partial Irradiation (800, 600) W/m^2

Under more severe partial shading conditions where Panel A operates at 800 W/m^2 and Panel B operates at 600 W/m^2, the performance differences between the GFM (PI–PI) and GFM (PI–TD3) control strategies become more pronounced. The conventional GFM (PI–PI) control strategy fails to maintain the DC-link voltage, which collapses from approximately 795 V to approximately 290 V. This indicates a loss of DC-link voltage regulation and suggests that the inverter can no longer sustain normal operation under the available PV power.

A. GFM (PI – PI) control **B. GFM (PI – TD3) control**

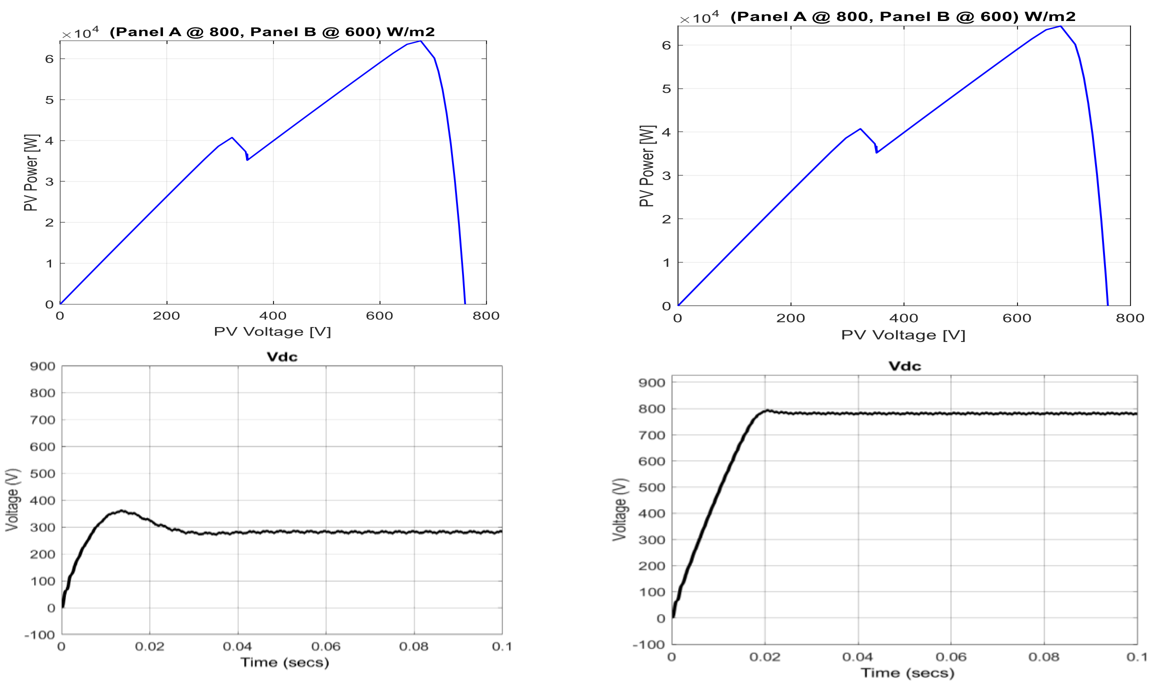

Fig 5: Case 3: P-V, Vdc @ 800, 600 W/m^2

The simulation plots for the P–V characteristics, DC-link voltage, and the resulting inverter voltage and current are presented in Fig. 5 for case 3. More severe oscillations are observed during startup, and the system requires approximately 0.03 secs to stabilize. In addition, the impact of partial shading is clearly reflected in the P–V characteristics curve, where the global maximum power point decreases to approximately 65 kW.

In contrast, the proposed GFM (PI–TD3) control strategy in Fig 5b demonstrates improved resilience under the same partial shading condition. The DC-link voltage experiences a slight reduction, but it is still maintained at approximately 779.1 Vdc, which remains within the acceptable design specification range.

Case 4: Partial Irradiation (600, 400) W/m^2

Further analysis of the proposed GFM (PI–TD3) control strategy was conducted under more severe partial shading conditions, where Panel A operates at 600 W/m^2–and Panel B operates at 400 W/m^2. The impact of partial shading is also reflected in the P–V characteristics curve, where the global maximum power point decreases further to approximately 45 kW.

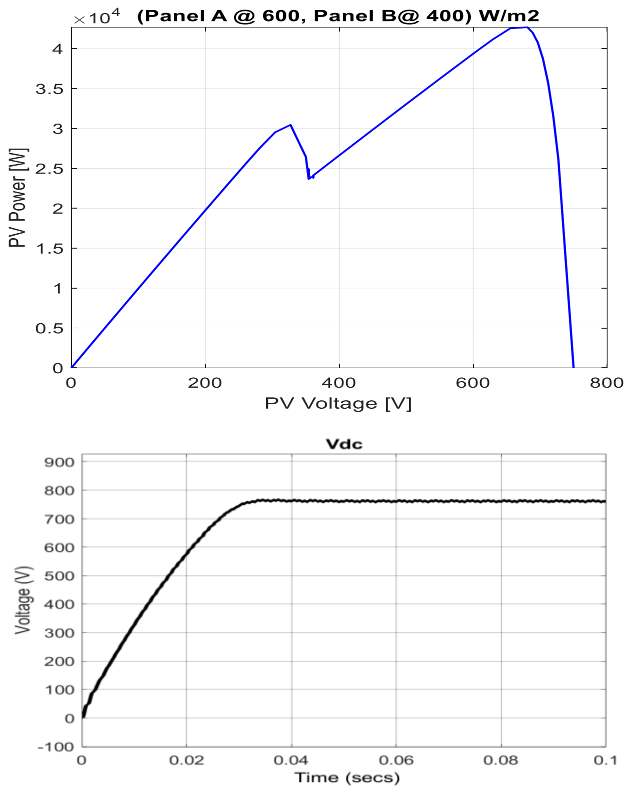

Fig 6: Case 4: P-V, Vdc @ 600, 400 W/m^2

Despite this substantial reduction in available power, the proposed TD3 – DRL control architecture demonstrates enhanced robustness and voltage regulation capability under heavily degraded operating conditions. The DC - link voltage is still maintained at a relatively stable value of approximately 759.8 Vdc which remains within the acceptable range of the designed specifications. The simulation plots for the P–V characteristics and DC-link voltage are presented in Fig. 6. From the simulation results, it is observed that the system requires approximately 0.03 secs to stabilize during startup due to the increased severity of the shading condition.

The comparative evaluation between both controller strategy includes quantitative transient metrics in terms of DC-link voltage settling time and estimated peak overshoots for all simulated irradiance cases. The proposed PI–TD3 controller consistently reduced the estimated peak DC-link voltage overshoot from approximately 7–12% with the conventional PI–PI controller to approximately 2–6%, while reducing settling time across all comparable operating conditions. Table 3 summarizes the principal transient performance metrics extracted from the simulation cases.

Table 3: Comparison of the transient performance metrics

Case	**Irradiance (W/m²)**	**Controller**	**Settling Time (s)**	**Estimated Overshoot (%)**	**Steady state V_{dc} (V)**
1	(1000,1000)	PI–PI	0.016	7–8	795
		PI–TD3	<0.01	2–3	795
2	(1000,800)	PI–PI	0.03	10–12	795
		PI–TD3	<0.02	4–5	795
3	(800,600)	PI–PI	0.03	Not applicable*	290
		PI–TD3	0.02	3–4	779.1

For Case 3, the conventional PI–PI controller does not exhibit a transient overshoot because the DC-link voltage collapses to an undesirable steady-state value of about 290 V. Therefore, the objective of the proposed TD3-based inner current controller is to improve transient current tracking and consequently voltage regulation so that the DC – link remains stable while the system transitions to a new operating point dictated by the available solar power.

Overall, the proposed PI–TD3 controller consistently demonstrated faster settling times and lower DC-link voltage overshoot than the conventional PI–PI controller across all comparable irradiance conditions. Under severe partial shading (800, 600 W/m²), the proposed controller preserved DC-link regulation near its nominal operating point. These observations are consistent with the improved current-tracking capability discussed in Section 3, which reduces transient power imbalance across the DC-link capacitor during irradiance disturbances.

Scope and Limitation

The primary objective of this study is to isolate the effect of replacing the conventional inner current PI controller with a TD3-based controller within the same droop – controlled grid-forming (GFM) inverter architecture. Accordingly, the comparative evaluation is limited to a conventional PI–PI baseline, while the outer droop controller and the DDPG based MPPT controller remain unchanged throughout all simulations. Consequently, the observed differences in DC-link voltage regulation can be attributed solely to the replacement of the inner current controller.

A comprehensive power-quality assessment was beyond the scope of the present study. Metrics such as steady-state voltage error, AC voltage total harmonic distortion (THD), RMS voltage deviation, current-limit compliance, LCL-filter resonance characteristics, and statistical variability over repeated irradiance events were not evaluated and will be investigated in future work. Comparison of the proposed controller with other advanced control strategies can be done as future investigations.

Furthermore, the simulations employ approximately 0.1 s time windows to capture the fast transient response of the inverter inner current loop immediately following irradiance changes. These short-duration simulations are appropriate for evaluating controller dynamics during disturbance initiation but do not represent the longer timescales associated with realistic cloud-driven irradiance variations. Long duration simulations incorporating realistic irradiance profiles will therefore be considered in future work to further assess controller robustness.

The TD3 controller employed in this work is an unconstrained reinforcement learning policy trained to optimize transient current control performance through the defined reward function. The stability of the proposed controller is demonstrated empirically through simulation. However future work will consider the integration of stability – aware reinforcement learning techniques and formal closed-loop stability analysis.

CONCLUSION

This paper presented a hybrid DRL – GFM inverter control architecture for a standalone PV system under partial shading conditions. A DDPG agent handled MPPT at the boost converter stage, while a TD3 agent replaced the conventional inner current PI regulators of the droop-controlled GFM inverter. The proposed GFM (PI–TD3) controller was compared against a conventional GFM (PI–PI) controller across four irradiance scenarios of increasing shading severity. Results consistently showed that the proposed controller achieved faster DC-link stabilization, reduced startup oscillations, and superior voltage regulation. Although the available PV power decreases with irradiance, as dictated by the PV source characteristics, the proposed controller enables the inverter to transition to the corresponding lower-power operating point while maintaining significantly better DC-link voltage stability than the conventional PI-based controller.

As this study is based on offline MATLAB/Simulink simulations, future work will focus on sensitivity analyses, hardware-in-the-loop (HIL) validation and real-time embedded implementation to evaluate the computational feasibility and practical performance of the proposed hybrid DDPG–TD3 control architecture. Further investigations will also consider dynamic load conditions, battery supported standalone systems, and grid-connected microgrid operation to assess the controller's robustness under more realistic operating scenarios.

ACKNOWLEDGEMENT

I would like to express my sincere gratitude to Prof. Akuro Big-Alabo for his invaluable encouragement, guidance, and support throughout the course of this research. I also wish to acknowledge the staff of the Department of Electrical/Electronic Engineering for their continuous encouragement and assistance during the period of this study, of which this work forms a part.

REFERENCE

Big-Alabo, A. (2026). Maximum power point tracking of solar-PV under partial shading and variable load resistance using deep reinforcement learning algorithm. *Discover Electronics*, *3*(1), 41. https://doi.org/10.1007/s44291-026-00187-x

Cao, J., Du, W., Wang, H., & Bu, S. (2020). Optimal distribution network restoration using deep reinforcement learning. *IEEE Transactions on Power Systems*, *35*(4), 3059–3073. https://doi.org/10.1109/TPWRS.2020.2980843

Chand, S. S., Prasad, R., Mudaliar, H., Kumar, D., Fagiolini, A., Di Benedetto, M., & Cirrincione, M. (2022). Enhanced current loop PI controllers with adaptive feed-forward neural network via estimation of grid impedance: Application to three-phase grid-tied PV inverters. *2022 IEEE Energy Conversion Congress and Exposition (ECCE)*, 1–8. https://doi.org/10.1109/ECCE50734.2022.9947752

Colombino, M., Groß, D., Brouillon, J.-S., & Dörfler, F. (2019). Global phase and magnitude synchronization of coupled oscillators with application to the control of grid-forming power inverters. *IEEE Transactions on Automatic Control*, *64*(11), 4496–4511. https://doi.org/10.1109/TAC.2019.2898549

Das, S. C., Vu, T., Ramasubramanian, D., Farantatos, E., Zhang, J., & Ortmeyer, T. (2025). Deep reinforcement learning for optimizing inverter control with fixed and adaptive gain tuning strategies for power system stability. *IEEE Transactions on Smart Grid*, *16*(6), 5099–5110. https://doi.org/10.1109/TSG.2025.3604554

Dasgupta, S., Mohan, S. N., Sahoo, S. K., & Panda, S. K. (2013). Lyapunov function-based current controller to control active and reactive power flow from a renewable energy source to a generalized three-phase microgrid system. *IEEE Transactions on Industrial Electronics*, *60*(2), 799–813. https://doi.org/10.1109/TIE.2012.2206356

Fujimoto, S., Hoof, H., & Meger, D. (2018, July). Addressing function approximation error in actor – critic methods. In *International conference on machine learning* (pp. 1587-1596) PMLR.

Gutiérrez-Escalona, J., Roncero-Clemente, C., Husev, O., Matiushkin, O., & Blaabjerg, F. (2024). Artificial intelligence in the hierarchical control of AC, DC, and hybrid AC/DC microgrids: A review. *IEEE Access*, *12*, 157227–157246. https://doi.org/10.1109/ACCESS.2024.3486382

Hassan, J., Ahmed, J., Papadopoulos, S., Kahwash, F., & Goh, K. (2026). A comprehensive review of frequency response and control strategies for grid-connected solar photovoltaic systems. *Renewable and Sustainable Energy Reviews*, *226*, 116324. https://doi.org/10.1016/j.rser.2025.116324

Jia, B., Wu, H., & Li, Y. (2021). Flexible on-grid and off-grid control strategy of photovoltaic energy storage system based on VSG technology. *2021 IEEE 5th Conference on Energy Internet and Energy System Integration (EI2)*, 1978–1985. https://doi.org/10.1109/EI252483.2021.9712963

Khan, M., Wu, W., & Li, L. (2024). Grid-forming control for inverter-based resources in power systems: A review on its operation, system stability, and prospective. *IET Renewable Power Generation*, *18*, 887–907. https://doi.org/10.1049/rpg2.12991

Lakshmanan, S. A., Jain, A., & Rajpourhit, B. S. (2014). A novel current controlled technique with feed forward DC voltage regulator for grid connected solar PV system. *2014 Eighteenth National Power Systems Conference (NPSC)*, 1–6. https://doi.org/10.1109/NPSC.2014.7103892

Lei, Q., Yang, S., & Peng, F. Z. (2010). High-performance and cost-effective multiple feedback control strategy for standalone operation of grid-connected inverter. *2010 Twenty-Fifth Annual IEEE Applied Power Electronics Conference and Exposition (APEC)*, 854–860. https://doi.org/10.1109/APEC.2010.5433567

Lin, Q., Shijo, T., Ogawa, K., Uno, H., Kanekiyo, Y., & Arai, J. (2022). Inertia evaluations on grid forming inverters with virtual synchronous generator control applied to photovoltaic power systems. *2022 IEEE Energy Conversion Congress and Exposition (ECCE)*, 1–6. https://doi.org/10.1109/ECCE50734.2022.9947992

Liserre, M., Teodorescu, R., & Blaabjerg, F. (2006). Stability of photovoltaic and wind turbine grid-connected inverters for a large set of grid impedance values. *IEEE Transactions on Power Electronics*, *21*(1), 263–272. https://doi.org/10.1109/TPEL.2005.861185

Majia, A. M., Ado, M., Wali, S. M., & Ismail, U. I. (2026). Development of an improved perturb and observe (P&O) algorithm for MPPT controller. *JOURNAL OF BASICS AND APPLIED SCIENCES RESEARCH*, *4*(3), 157-162.

Mohanty, S., Subudhi, B., & Ray, P. K. (2016). A new MPPT design using grey wolf optimization technique for photovoltaic system under partial shading conditions. *IEEE Transactions on Sustainable Energy*, *7*(1), 181–188. https://doi.org/10.1109/TSTE.2015.2482120

Naderipour, A., Abdul-Malek, Z., Miveh, M. R., Hadidian Moghaddam, M. J., Kalam, A., & Gandoman, F. H. A. (2018). Harmonic compensation strategy in a grid-connected photovoltaic system using zero-sequence control. *Energies*, *11*(10), 2629. https://doi.org/10.3390/en11102629

Ninad, N. A., & Lopes, L. A. C. (2012). Control of Δ-Y transformer based grid forming inverter for unbalanced stand-alone hybrid systems. *2012 IEEE Electrical Power and Energy Conference (EPEC)*, 176–181. https://doi.org/10.1109/EPEC.2012.6474945

Rocabert, J., Luna, A., Blaabjerg, F., & Rodríguez, P. (2012). Control of power converters in AC microgrids. *IEEE Transactions on Power Electronics*, *27*(11), 4734–4749. https://doi.org/10.1109/TPEL.2012.2199334

Shahgholian, G., Moradian, M., & Fathollahi, A. (2025). Droop control strategy in inverter-based microgrids: A brief review on analysis and application in islanded mode of operation. *IET Renewable Power Generation*, *19*(1), e13186. https://doi.org/10.1049/rpg2.13186

Yang, Y., Xiao, Y., Peng, Q., Blaabjerg, F., Wu, Y., & Ji, X. (2022). Virtual energy storage operation for smart photovoltaic inverters. *2022 IEEE 13th International Symposium on Power Electronics for Distributed Generation Systems (PEDG)*, 1–6. https://doi.org/10.1109/PEDG54999.2022.9923090

Yazdani, S., Ferdowsi, M., Davari, M., & Shamsi, P. (2020). Advanced current-limiting and power-sharing control in a PV-based grid-forming inverter under unbalanced grid conditions. *IEEE Journal of Emerging and Selected Topics in Power Electronics*, *8*(2), 1084–1096. https://doi.org/10.1109/JESTPE.2019.2959006

Zhang, Y., Wang, X., Wang, J. and Zhang, Y., (2020). Deep reinforcement learning based volt-var optimization in smart distribution systems. *IEEE Transactions on Smart Grid*, *12*(1), pp.361-371. https://doi.org/10.1109/TSG.2020.3010130