Journal of Solar Energy Research (JSER)

Journal homepage: www.jser.ut.ac.ir

Quantum Reinforcement Learning–Enabled DSTATCOM for Transient Power Quality Stabilization in High-Penetration PV Distribution System

Mohammed Nadeem Iqbal[a,*], Asini Kumar Baliarsingh[b], Pratap Sekhar Puhan[c]

[a] *Department of Electrical Engineering, Biju Pattnaik University of Technology, Rourkela, Odisha, India*
[b] *Electrical Engineering, Government College of Engineering Kalahandi, Odisha, India*
[c] *Electrical and Electronics Engineering, Sreenidhi institute of science and Technology Hyderabad, India*

ARTICLE INFO

Article Type:

Research Article

Received:2025.12.09
Accepted in revised form:2026.06.28

Keywords:
Photovoltaic; DSTATCOM; Quantum Reinforcement Learning; Power Quality; Reactive Power Compensation

A B S T R A C T

This study proposes a Quantum Reinforcement Learning-based Distribution Static Compensator (QRL-DSTATCOM) for transient power quality enhancement in high-penetration photovoltaic (PV) distribution systems. Increasing integration of PV generation introduces significant challenges, including voltage instability, harmonic distortion, and power factor degradation due to stochastic irradiance and rapid load fluctuations. To address these issues, a hybrid quantum-classical reinforcement learning framework is developed, incorporating quantum state encoding and amplitude-enhanced exploration to improve learning efficiency and control performance. The proposed QRL-DSTATCOM enables faster convergence and superior policy optimization compared to conventional RL-based compensators. The system is modeled and validated in MATLAB/Simulink under diverse operating scenarios, including variable irradiance, sudden load changes, and grid fault conditions. Simulation results demonstrate that the proposed approach achieves 43–50% faster stabilization time, significantly reduced total harmonic distortion (THD) well within IEEE-519 limits, and maintains bus voltage within ±5% of nominal values under severe disturbances. Overall, the framework validated using OPAL-RT real time software and presents a scalable and intelligent solution for next-generation smart distribution networks with high renewable energy penetration, ensuring improved stability, robustness, and power quality performance.

1. Introduction

The rapid transition of modern power systems from conventional fossil-fuel-based generation to renewable energy-dominated networks has fundamentally transformed the operational characteristics of distribution systems. Among renewable energy sources, photovoltaic (PV)

*Corresponding Author Email: nadeemiqbal250@gmail.com

Cite this article: Nadeem Iqbal, M., Baliarsingh, A. Kumar and Puhan, P. Sekhar (2026). Quantum Reinforcement Learning–Enabled DSTATCOM for Transient Power Quality Stabilization in High-Penetration PV Distribution System. Journal of Solar Energy Research, 11(2), 3890-3908. doi: 10.22059/jser.2026.407887.1683

DOI: 10.22059/jser.2026.407887.1683

generation has experienced substantial growth due to its environmental benefits, declining installation costs, and widespread deployment. However, the increasing penetration of PV systems introduces significant challenges to power quality (PQ) and grid stability because of the intermittent and stochastic nature of solar energy. Variations in solar irradiance caused by cloud movement, rapid fluctuations in load demand, bidirectional power flow, and harmonic emissions from power electronic converters create highly dynamic operating conditions within distribution networks.

Traditionally, Distribution Static Compensators (DSTATCOMs) have been employed to enhance power quality by providing reactive power compensation, mitigating voltage sags and swells, improving power factor, and maintaining voltage stability. Conventional DSTATCOM controllers, typically based on proportional–integral (PI) regulators or rule-based strategies, perform satisfactorily under relatively stable grid conditions. However, these approaches exhibit several limitations when subjected to rapidly changing operating environments. Fixed controller parameters, limited adaptability, slower dynamic response, and reduced disturbance rejection capability often result in degraded performance during severe fluctuations in renewable generation and load demand. Consequently, maintaining compliance with stringent power quality standards such as IEEE-519 becomes increasingly difficult.

To address these challenges, intelligent control techniques based on Reinforcement Learning (RL) have been investigated for adaptive reactive power compensation and inverter control. RL-based methods improve decision-making by learning optimal control actions from system interactions. Nevertheless, classical RL algorithms often suffer from slow convergence, high computational complexity, and difficulties associated with large state-action spaces. Their performance may deteriorate under highly nonlinear, uncertain, and fault-prone operating conditions commonly encountered in modern PV-integrated distribution systems.

Quantum Reinforcement Learning (QRL) has recently emerged as a promising solution that combines the learning capabilities of classical RL with quantum-inspired computational principles such as superposition, entanglement, and amplitude-enhanced exploration. By enabling simultaneous evaluation of multiple state-action possibilities, QRL significantly improves exploration efficiency and accelerates convergence toward optimal control policies. When integrated with DSTATCOM technology, QRL enhances real-time reactive power compensation, strengthens DC-link voltage regulation, and provides rapid adaptation to grid disturbances. The quantum-enhanced learning framework improves voltage regulation, reduces total harmonic distortion (THD), and increases operational robustness under conditions such as partial PV shading, sudden load variations, and transient fault events.

The integration of QRL with DSTATCOM systems represents a significant advancement in intelligent power electronic control for future smart grids. By offering faster transient response, superior harmonic mitigation, and enhanced voltage stability, QRL-enabled DSTATCOMs provide an effective solution for maintaining power quality in renewable-rich distribution networks. Furthermore, future developments in quantum-inspired computing architectures and scalable multi-agent learning frameworks are expected to extend the applicability of this technology to large-scale smart distribution systems, supporting reliable, adaptive, and resilient grid operation under increasingly complex and uncertain conditions.

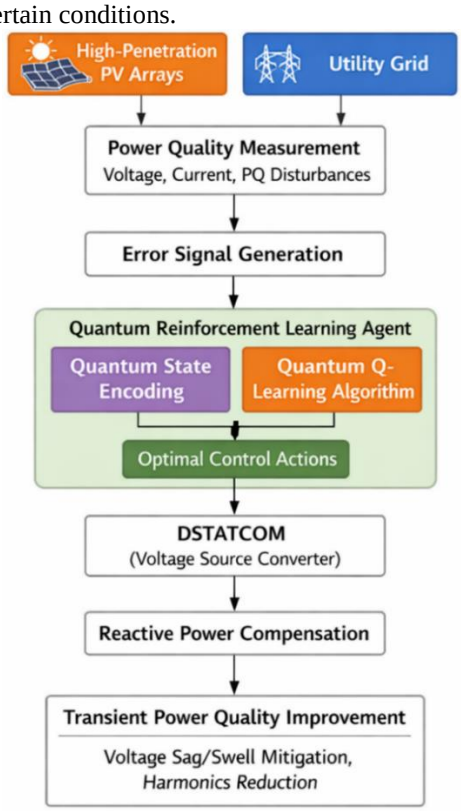


Figure 1. Quantum Reinforcement Learning–enabled DSTATCOM architecture

Figure 1 below shows a Quantum Reinforcement Learning-enabled DSTATCOM architecture that aims at improving transient power quality in high-penetration PV distribution feeders. Power is fed by the PV system to the grid at the PCC by an MPPT-controlled DC-AC inverter. The voltages and current signals are detected, filtered against noise, and coded into quantum states and fed into the RL agent. Depending on grid conditions, disturbances, and fault identification, the agent optimized actions cause the DSTATCOM to take adaptive compensatory actions, and a classical controller runs in parallel to compare the baseline. A Power Quality Index Estimator generates rewards and a replay buffer can store experience to be used in a constant stream of learning.

Omoboye et al. [1] studied the effect of D-STATCOM and SVC on the performance of low-voltage radial distribution networks and showed that the application of D-STATCOM makes the system perform better in terms of voltage regulation, reactive power compensation, and harmonic mitigation, thus improving the overall power quality performance. Suryawanshi et al. [2] presented an adaptive DSTATCOM controller with a hybrid Neural Network–BAT optimization technique, which minimized total harmonic distortion and enhanced voltage stability under different operating conditions. Mangaraj et al. [3] presented a DBLN-based IC-DSTATCOM approach for power-line conditioning and achieved significant improvements in harmonic compensation, load balancing, and power factor correction in distribution systems. Mahla and Garg [4] presented an E-SOGI-controlled DSTATCOM in combination with wind energy and showed better voltage regulation, disturbance rejection, and power quality improvement. Kaliappan et al. [5] compared different optimization techniques used in hybrid microgrids and found that advanced metaheuristic techniques provide better energy management and operational efficiency. Wu et al. [6] developed a frequency-constrained economic dispatch model for a microgrid to achieve frequency stability and economic operation. Narendra Babu [7] discussed adaptive grid-connected inverter control schemes and smart adaptive control as potential solutions for future power-quality improvement in microgrid-based power systems. Bai et al. [8] proposed a Fast S-Transform and enhanced CNN-LSTM hybrid model for power-quality disturbance classification, achieving high classification accuracy in complex operating environments. Zhu et al. [9] proposed a novel decomposition and detection framework for complex power-quality disturbances that can correctly identify complex disturbances and enhance disturbance analysis performance. Rajendran et al. [10] reviewed the integration of solar photovoltaic systems with smart grids and presented challenges, standards, grid codes, and future prospects for renewable energy deployment. Meziane et al. [11] introduced a hierarchical two-layer MPC-supervised management approach for inverter-based microgrids, where the first layer increases operational efficiency and the second enhances system flexibility. Chen et al. [12] developed an intelligent controlled DSTATCOM (IDC) system to improve voltage profile, power factor, and harmonic suppression in distribution networks. SivaramKrishnan et al. [13] proposed a smart EV charging system incorporating solar PV and UPQC, which stabilized the grid and enhanced power quality using the HBA-MORARNN algorithm. Wu et al. [14] proposed a CNN-LSTM-Attention-based solar PV forecasting model that incorporated twenty-four solar term divisions to improve forecasting accuracy. Hu et al. [15] proposed a hybrid CNN-LSTM-Attention model based on neighboring station information and achieved superior short-term PV power prediction results. Ledmaoui et al. [16] proposed a CNN-based photovoltaic fault detection and classification system implemented using PyQt5, achieving a diagnostic accuracy of 99.4%. Alharkan et al. [17] proposed a dual-stream CNN-LSTM model for solar power generation prediction and achieved better forecasting performance than existing methods. Vali et al. [18] proposed a deep learning-based controller for parallel DSTATCOM systems that significantly enhanced power quality and dynamic compensation performance. Wang et al. [19] employed a hybrid-algorithm-optimized Bi-LSTM network to improve short-term PV power forecasting accuracy. Kiasari and Aly [20] summarized artificial intelligence-based control strategies for FACTS devices and their applications in modern power-quality management. Ren et al. [21] presented a CNN-based wind turbine forecasting architecture that considers meteorological data and demonstrated higher forecasting accuracy for ultra-short-term PV power prediction compared with the BP model. Garcia et al. [22] compared CNN, LSTM, and CNN-LSTM approaches for power-quality disturbance classification and concluded that the CNN-LSTM approach provides the best performance. Bui Duy et al. [23] optimized the input parameters of an LSTM model for solar power forecasting, resulting in improved prediction accuracy and robustness. Okwako et al. [24] designed a neural-network-controlled solar PV battery-based UPQC that successfully compensated for power-

quality disturbances during grid-connected operation. Bai et al. [25] further validated the effectiveness of Fast S-Transform and CNN-LSTM-based disturbance classification methods for complex power-quality events. Karchi et al. [26] suggested an adaptive least mean square (ALMS) controller for photovoltaic systems to achieve harmonic mitigation and voltage regulation. Iweh et al. [27] explored the conditions, challenges, benefits, and possible solutions for integrating distributed and renewable generation resources into power grids. Ganthia et al. [28] studied voltage sag compensation using a DVR-based voltage restoration algorithm with a PI controller and analyzed its restoration capability during voltage sag events. Ganthia et al. [29] investigated the performance of MPPT in photovoltaic systems using a fuzzy logic controller and achieved improved energy extraction efficiency. Rubavathy et al. [30] introduced a multi-agent system for distributed monitoring and control in smart-grid-based transmission networks. Sahu et al. [31] designed a multiphase interleaved boost converter that provides improved conversion efficiency and reduced ripple content in grid-connected photovoltaic systems. Mohanty et al. [32] proposed a dynamic PSO with ESC algorithm under partial shading conditions to improve stability and output consistency in photovoltaic power smoothing systems. Suresh et al. [33] designed a reduced-rating hybrid DSTATCOM for three-phase four-wire distribution systems and demonstrated effective compensation of power-quality disturbances. Hussain et al. [34] presented a delayed LMS-based adaptive control scheme for PV-DSTATCOM systems and observed improved dynamic performance and harmonic suppression. Kundela et al. [35] discussed the operation of inductively coupled DSTATCOM and demonstrated considerable improvement in voltage regulation and power-quality parameters. Yarlagadda et al. [36] developed a DSTATCOM-based closed-loop control strategy to regulate the terminal voltage of wind power plants under load disturbances and obtained effective results. Li et al. [37] suggested a voltage self-stability control approach for STATCOM applications in photovoltaic power stations and achieved subsynchronous oscillation damping under weak-grid conditions. Hurkadli and Kulkarni [38] proposed a coordinated energy management and reactive power control scheme involving electric vehicles to enhance microgrid voltage stability. Thomas and N. [39] presented a D-STATCOM-based voltage control strategy for single-phase grid-tied systems and demonstrated improved voltage regulation performance. A hybrid solution for power quality enhancement using renewable energy sources was introduced by Kumar and Gopalakrishnan, which resulted in better reliability and performance of the system [40]. It also achieves 43–50% faster response and improved robustness under stochastic disturbances, confirming its enhanced dynamic performance and suitability for real-time smart grid applications with high PV penetration.

2. System Architecture and Modelling

The proposed System Architecture of a Quantum Reinforcement Learning-enabled DSTATCOM is aimed at providing a well-stabilized transient power quality in distribution grids based on large photovoltaic penetration of PV. With the advent of the high rate of solar integration that causes fast changes in power, voltage drops, harmonic distortion, and unbalanced supply currents, the control infrastructure needs to be adaptive and compensating at high speeds. The architecture shows that PV generation, DC-link energy storage, and three-phase VSC are combined and interfaced at the PCC. Voltage, current and harmonics are measured in real time and pass to a Quantum Reinforcement Learning controller which then controls the switching actions optimally.

2.1 PV-integrated distribution

PV-integrated distribution feeder is the basic electrical structure in contemporary smart grids in which distributed solar generation is connected to the utility structure. It uses a detailed two-diode model of the photovoltaic source, which allows the nonlinear I V characteristics to be correctly replicated at realistic operating conditions. The model uses dynamic irradiance variations and partial-shading conditions to capture rapid intermittencies due to the movement of clouds and dust deposits, as well as, inappropriate matching of PV modules performance. The effects cause variations in the PV output voltage, current and power and can result in the feeder becoming unstable unless mitigated accordingly. An LC filtering stage which removes switching harmonics of the inverter and ensures interface compliance with the quality requirements of grid voltages links the feeder to the grid at the Point of Common Coupling (PCC). At the consumer end, load conditions are modeled using hybrid linear and nonlinear load models and load rich harmonic content is induced and abrupt load changes to simulate commercial and industrial conditions. These complicated load situations make voltage imbalance, cumulative harmonic distortion and

power factor degradation worse, particularly during high-penetration PV conditions.
PV single-diode I–V model (cell):

$$I_{cell} = I_{ph} = I_s \left(exp\left(\frac{V_{cell}+I_{cell}R_s}{nV_t}\right) - 1\right) - \left(\frac{V_{cell}+I_{cell}R_s}{R_{sh}}\right) \quad (1)$$

PV array scaling (array current & voltage):

$$I_{pv} = N_p I_{cell}\ , V_{pv} = N_s V_{cell} \quad (2)$$

DC-link voltage dynamics:

$$C_{dc}\frac{dtV_{dc}}{dt} = I_{pv} - I_{conv} - I_{loss} \quad (3)$$

Shunt VSC filter dynamics d-axis (synchronous frame):

$$i_{sd} = \frac{1}{L_f}\left(v_{sq} - R_f i_{sq} - v_{cq}\right) - \omega i_{sd} \quad (4)$$

Series VSC injected voltage relation (d-q):

$$v_{pcc,d} = v_{grid,d} + v_{inj,d} \quad , v_{pcc,q} = v_{grid,q} + v_{inj,q} \quad (5)$$

PCC voltage deviation (magnitude or component):

$$\Delta V = V_{pcc} + V_{ref} \quad (6)$$

Reverse Saturation Current of PV Cell:

$$I_O = I_{O,ref}\left(\frac{T}{T_{ref}}\right)^3 exp\left[\frac{qE_g}{k}\left(\frac{1}{T_{ref}} - \frac{1}{T}\right)\right] \quad (7)$$

PV Output Power:

$$P_{PV} = V_{PV}I_{PV} = \left(V_{OC} - \frac{V_{OC}-V_{mpp}}{I_{SC}} I_{PV}\right) I_{PV} \quad (8)$$

Maximum Power Point Tracking (MPPT) Condition:

$$\frac{dP_{PV}}{dV_{PV}} = I_{PV} + V_{PV}\frac{dI_{PV}}{dV_{PV}} = 0 \quad (9)$$

Power Balance Equation for PV-DSTATCOM System:

$$P_{PV} + P_{grid} = P_{load} + P_{loss} + P_{DSTATCOM} \quad (10)$$

Thus, the PV-based feeder is a problematic dynamically changing system that necessitates high-level adaptive compensators, such as DSTATCOM, to provide stable, reliable, and high-quality power.

2.2 Distribution Static Compensator (DSTATCOM)

Distribution Static Compensator (DSTATCOM) is fitted with a three step Voltage Source Inverter (VSI) with high frequency Insulated Gate Bipolar having fast switching capability and high reliability in distribution-level power conditioning Transistors (IGBTs) Transistors (IGBTs) are used to provide fast switching and a high level of reliability. The VSI connects to the grid in a shunt mode, and two-way exchange of reactive power is permitted to ensure that the desired level of voltage is applied at the Point of Common Coupling (PCC). Instantaneously injecting or absorbing the reactive current can alleviate the voltage sags, imbalance, and enhance the power factor in the face of scale changing loads and PV generated conditions. This enhances general resilience and stability of the feeder especially in areas where there are intermittently renewable energies.

DC-Link Voltage Dynamics of DSTATCOM:

$$C_{dc}\frac{dV_{dc}}{dt} = \frac{P_{inv}-P_{load}}{P_{dc}} \quad (11)$$

C_{dc} is the DC-link capacitor that stores energy, while V_{dc} is its voltage. P_{inv} is the active power injected by inverter (from PV or grid), and P_{load} is the consumed load power.

This relationship is used in voltage regulation, control design, and DC-link stability analysis.

Instantaneous Active and Reactive Power (αβ frame):

$$S = P + jQ = \frac{3}{2}\left(v_\alpha i_\alpha + v_\beta i_\beta\right) + j\frac{3}{2}\left(v_\beta i_\alpha - v_\alpha i_\beta\right) \quad (12)$$

Here, v_α, v_β and i_α, i_β are voltage and current in the stationary αβ reference frame. The active power P denotes real power transfer, while reactive power Q indicates compensation requirement. This is used in instantaneous power theory (p–q control) to generate compensation current reference for DSTATCOM. For the series converter dynamics:

$$L_{se}(t)\frac{di_{se}}{dt} + R_{se}i_{se} = v_s + (v_{inv,se} - v_L) \quad (13)$$

where $v_{inv,se}$ is the injected series inverter voltage.

The common DC-link capacitor voltage dynamics:

$$C_{dc}(t)\frac{dV_{dc}}{dt} = \frac{1}{V_{dc}}(P_{sh} + P_{se} - P_{loss}) \quad (14)$$

where: P_{sh} = power handled by shunt converter, P_{se}= power handled by series converter. P_{loss} = internal losses.

Using proposed controller, the common DC link capacitor dynamics:

$$V_{dc}[k+1] = V_{dc}[k] + \frac{T_S}{V_{dc}C_{dc}[k]}(P_{sh}[k] + P_{se}[k]) \quad (15)$$

where $P_{sh}[k]$ and $P_{se}[k]$ are instantaneous powers of shunt and series converters.

Overall Compensation Condition the DSTATCOM ensures:

$v_L(t) \approx V_m\, sin(\omega t)$ (sinusoidal load voltage)

$i_L(t) \approx I_m\, sin(\omega t)$ (sinusoidal source current)

The proposed Quantum Reinforcement Learning-Enabled DSTATCOM has the major control goals to provide reliability, quality, and stability to the PV-dominated distribution networks. The primary objective is to ensure that the Point of Common Coupling (PCC) voltage falls within 5 percent of its nominal value because variations in voltage are typical of feeders with high renewable penetration because of a quickly changing irradiance and sudden load changes. The compensator is able to maintain the restoration of voltage during sags, swells and unbalanced operating conditions by injecting reactive current adaptively.

The primary control objectives of the proposed Quantum Reinforcement Learning–Enabled DSTATCOM are designed to ensure reliability, quality, and stability in PV-dominated distribution networks. The foremost aim is to maintain the Point of Common Coupling (PCC) voltage within ±5% of its nominal rating, as voltage fluctuations are common in feeders with high renewable penetration due to rapidly changing irradiance and abrupt load variations. By adaptively injecting reactive current, the compensator supports voltage restoration during sags, swells, and unbalanced operating conditions.

State-Space Model of DSTATCOM (dq Domain):

$$\frac{d}{dt}\begin{bmatrix} i_d \\ i_q \\ v_{dc} \end{bmatrix} = \begin{bmatrix} -\frac{R_S}{L_S} & \omega & 0 \\ -\omega & \frac{R_S}{L_S} & 0 \\ 0 & 0 & -\frac{1}{R_{dc}C_{dc}} \end{bmatrix}\begin{bmatrix} i_d \\ i_q \\ v_{dc} \end{bmatrix} + \begin{bmatrix} \frac{1}{L_S}v_d \\ \frac{1}{L_S}v_q \\ \frac{1}{C_{dc}}I_{inv} \end{bmatrix} \quad (16)$$

The above equation (8) is used for stability, system modeling, and control design.

Voltage Regulation Support by DSTATCOM:

$$\Delta V = \frac{X_{syst}Q_{comp}}{V} \quad (17)$$

This equation (9) determines the voltage boost achieved via VAR compensation.

LCL Filter Resonant Frequency for DSTATCOM:

$$f_{res} = \frac{1}{2\pi}\sqrt{\frac{L_1+L_2}{L_1L_2C_S}} \quad (18)$$

where f_{res} Resonant frequency of the LCL filter (Hz),

L_1 Inverter-side inductance (H) and L_2 Grid-side (or line-side) inductance (H) and C_SFilter capacitance connected between phases and ground (F).

The supply voltage is represented as:

$$v_s(t) = V_m\, sin(\omega t) \quad (19)$$

Load voltage distorted due to harmonics:

$$v_L(t) = v_s(t) - v_{se}(t) \quad (20)$$

where $v_{se}(t)$ = injected series compensating voltage.

Load current consists of fundamental + harmonic + reactive components:

$$i_L(t) = i_{Lf}(t) + i_{Lh}(t) + i_{Lq}(t) \quad (21)$$

The shunt active filter injects current $i_{sh}(t)$ so that:

$$i_L(t) = i_s(t) - i_{se}(t) \quad (22)$$

For the shunt converter connected to PCC:

$$L_{sh}(t)\frac{di_{sh}}{dt} + R_{sh}i_{sh} = v_{inv,sh} - v_{pcc} \quad (23)$$

where: $L_{sh}(t)$ and R_{sh} = filter inductor parameters, $v_{inv,sh}$ = inverter output voltage, v_{pcc} = point of common coupling voltage.

The major goal is the reduction of the existing harmonics that occur due to nonlinear industry or residential loads. The control system is aimed at minimizing Total Harmonic Distortion (THD) to levels that are within the IEEE-519 standards to avoid overheating of grid sensitive assets, malfunction and early aging. Transient stability requires short response time (typically less than 50 ms) in cases of sudden disturbances, e.g. switching noise of inverters, oscillations caused by faults or PV power transient. Also, the control strategy is oriented to stabilize the power factor near to unity (0.981.0) which facilitates further power transfer and reduces reactive load on the grid. Taken together, these goals allow to guarantee higher performance of the work of the operational level, and better good voltage regulation and quality of power in the contemporary smart feeders.

The reinforcement learning to the proposed DSTATCOM controller is based on a properly defined State Action Reward (SAR) representation that allows making smart decisions in the complex and quickly changing electrical environment. Measures of the state variables taken at the PV-integrated feeder are the magnitude of the PCC voltage, the DC-link voltage level, the DC-indicators of current harmonic distortion, and the real-time reactive power demand. All of these parameters characterize the dynamic state of the grid and makes sure that the agent is fully aware of the deterioration of the quality of power and the boundaries of compensator performance. The harmonic and reactive indices include enable the system to overcome both the steady-state deviation as well as the changing deviation of the disturbances.

The Explicit coefficients of the Linear discrete-time update of the DSTATCOM design of the proposed controller:

$$i_d[k+1] = \left(1 - \frac{T_s R_f}{L_f}\right) i_d[k] + T_s \omega i_q[k] + \frac{T_s}{L_f} v_{conv,d}[k] - \frac{T_s}{L_f} v_{conv,ds}[k] \quad (24)$$

$$i_q[k+1] = -T_s \omega i_q[k] + \left(1 - \frac{T_s R_f}{L_f}\right) i_q[k] + \frac{T_s}{L_f} v_{conv,q}[k] - \frac{T_s}{L_f} v_{conv,sq}[k] \quad (25)$$

$$v_{Ld}[k+1] = v_{Ld}[k] + \frac{T_s}{C_f} i_{se,d}[k] - \frac{T_s}{C_f} i_{L,d}[k] \quad (26)$$

$$v_{Lq}[k+1] = v_{Lq}[k] + \frac{T_s}{C_f} i_{se,q}[k] - \frac{T_s}{C_f} i_{L,q}[k] \quad (27)$$

$$V_{dc}[k+1] = V_{dc}[k] + \frac{T_s}{V_{dc} C_{dc}[k]} (P_{sh}[k] + P_{se}[k]) \quad (28)$$

where first four equations are linear in the states and inputs; the DC-link update is nonlinear because the power terms divide by $V_{dc}[k]$.

A DC-link capacitor plays a critical role in the operation of the Voltage Source Inverter (VSI) by acting as an essential energy storage element that enables immediate compensation of power fluctuations in the system. It stabilizes the DC-link voltage by buffering the difference between input and output power, thereby ensuring that the inverter can continuously supply compensating currents without interruption. To maintain stable operation, a closed-loop control mechanism is implemented to regulate the DC-link voltage around a predefined reference value. This control loop adjusts switching actions to ensure that the voltage remains constant, thereby guaranteeing consistent inverter performance and reliable power quality enhancement under dynamic operating conditions. In the proposed Quantum Reinforcement Learning (QRL)-based DSTATCOM framework, the control actions generated by the learning agent are translated into optimized gating pulse signals for the three-phase VSI switches. These switching states are dynamically adjusted in real time by the intelligent agent, enabling precise control of inverter output currents. As a result, the system is capable of delivering appropriate reactive power compensation, mitigating harmonic distortions, and maintaining grid stability, including DC-link voltage restoration during transient disturbances. This

continuous control strategy provides fine-grained modulation of switching behavior, allowing rapid adaptation to sudden changes in load demand, irradiance variability, and fault conditions.

The reinforcement learning structure is further enhanced through a carefully designed reward function that acts as a multi-objective performance indicator. The reward formulation integrates key power quality metrics such as Total Harmonic Distortion (THD), Point of Common Coupling (PCC) voltage deviation, and reactive power compensation accuracy. Positive rewards are assigned when the system improves voltage stability, reduces harmonic content, and maintains reactive power balance within acceptable limits. Conversely, penalty terms are introduced for violations such as excessive voltage deviation, increased distortion, or system instability. This reward shaping mechanism guides the QRL agent toward an optimal control policy by continuously reinforcing desirable grid behavior while discouraging poor performance states. Over time, this adaptive learning process results in a self-evolving control strategy capable of achieving superior power quality enhancement in high photovoltaic penetration distribution networks.

3. Quantum Reinforcement Learning Control Framework

Simulation of the Quantum Reinforcement Learning-Enabled DSTATCOM (QRL-DSTATCOM) in high-penetration photovoltaic (PV) distribution feeders, with grid integrated system is taken into account, where three rooftop PV arrays with total capacity of 2.5MW were considered. The feeder consists of residential and industrial load with the overall peak load of 3.8 MW. A nonlinear load, such as diode-bridge rectifier and variable-frequency drive, was modelled to produce harmonic distortion to the 11th order. The QRL agent has system states that are composed of bus voltages, line currents, and DSTATCOM output signals, which means the state is 24-dimensional. Five thousand training episodes were done, which lasted 1s of simulation time, time step of 1ms, generating more than 5 million state-action pairs during policy evaluation. The quantum state encoding was used to encode voltage and current amplitudes using 6 qubits and the quantum policy optimization was applied with 50 iterations per episode and amplitude amplification to fasten the convergence. Classical actor-critic networks were being trained simultaneously with 0.001 (actor) and 0.005 (critic) learning rates. The hybrid combination was found to converge quickly with a voltage deviation of less than 2% and a total harmonic distortion of less than 3% after 1500 episodes proving its effectiveness in transient power quality stabilization.

The classical state s is encoded into a quantum state vector using amplitude encoding as:

$$| \psi(s)\rangle = \sum_{i=0}^{n-1} \alpha_i(s) | i\rangle \text{ where } \sum_{i=0}^{n-1} | \alpha_i(s) |^2 = 1 \quad (29)$$

Here, α_i represents normalized probability amplitudes encoding voltage, current, and irradiance states in the QRL-DSTATCOM environment.

The policy parameters θ are updated using a quantum-enhanced gradient ascent rule:

$$\theta_{t+1} = \theta_t + \eta \nabla_\theta E_{|\psi(s)\rangle \sim \pi\theta}[R_t + \gamma \max_{a'} Q(|\psi(s')\rangle, a')] \quad (30)$$

where η is the learning rate and γ is the discount factor.

The reward is defined to minimize voltage deviation, THD, and reactive power error:

$$R_t = -(w_1 | V_t - V_{ref} | + w_2 THDt + w_3 | Q_t - Q_{ref} |) \quad (31)$$

where w_1, w_2, w_3 are weighting coefficients balancing power quality objectives.

Quantum State Encoding:

Within the suggested QRL-DSTATCOM model, the system voltages and currents vectors are coded to the value of qubits, and quantum representation of classical electrical states is obtained. This quantum state encoding effectively allows the evaluation of a multitude of control policies simultaneously by superposition to give the agent an opportunity to consider a wide range of action paths at once.

Voltage/current vector as quantum state:

$$| \psi(t) | = \sum_{i=1}^{N} \alpha_i(t) | v_i(t), i_i(t) |, \quad \sum_{i=1}^{N} | \alpha_i(t) |^2 = 1 \quad (32)$$

Here $|\psi(t)|$ = quantum state representing system at time t, $\alpha i(t)$ = probability amplitude of the i-th voltage/current pair, $v_i(t)$, $i_i(t)$ = voltage and current measurements and N = total number of voltage/current samples.

Normalized voltage vector mapping to qubits:

$$| \psi V |= \frac{1}{\sqrt{\sum_{k=1}^{n} V_k^2}} \sum_{k=1}^{n} V_k \mid k \mid \quad (33)$$

Here V_k = k-th voltage component, $|k|$= basis state for the k-th qubit and n = number of qubits.
Multi-qubit encoding of reactive power and current:

$$| \psi_{QC} |= \prod_{j=1}^{m} ((cos\theta_j \mid 0 \mid + e^{i\phi j} sin\theta_j \mid 1 \mid)), \ \theta_j = arctan \frac{i_j}{v_j} \ , \ \phi_j = arg(Q_j) \quad (34)$$

Here $\prod_{j=1}^{m} cosine\ and\ sine\ term$ tensor product across qubits, θj, ϕj = amplitude and phase encoding for j-th qubit, Qj= reactive power and m = total qubits representing system features

Superposition for parallel policy evaluation:

$$| \Psi |= \frac{1}{\sqrt{M}} \sum_{a=1}^{M} \mid \psi(s_t, a) \mid, s_t \in S, a \in A \quad (35)$$

Here M = number of possible actions, s_t = system state at time t and a = action from action space A.

Probability of high-reward action:

$$P(\mid a^* \mid\mid s_t) = \mid a^* \mid \Psi \mid^2 = \mid \frac{1}{\sqrt{M}} \sum_{a=1}^{M} \alpha_k \ \delta_{a*,a_k} \mid^2 \quad (36)$$

Here a^* = optimal action and δ_{a*,a_k} = Kronecker delta (1 if $a^* = a_k$, else 0)
Using the correlations between the bus voltages and line currents are maintained, and the state description increases the fidelity. This parallelism greatly saves the exploration time in high-dimensional action-state spaces of high-penetration PV feeders. The methodology is a basis of quantum-enhanced policy optimization and speedy convergence in transient quality control of power.

3.1 Quantum-Enhanced Policy Optimization

Amplification of the amplitude is used to improve quantum-enhanced policy optimization in the QRL-DSTATCOM system to quickly identify high-reward actions in complex power system environments. The agent will converge faster than the classical approaches to reinforcement learning as it increases the likelihood of choosing the best control actions. Also, entanglement-based sampling maintains correlations between qubits modeling system states, which allows the sound assessment of interdependent voltage and current values.

Amplitude amplification for optimal action:

$$| \Psi^{(r)} |= (U_s U_f)^r \mid \Psi_0 \mid, r \approx \frac{\pi}{4} \sqrt{\frac{M}{k}} \quad (37)$$

Here U_f = oracle marking high-reward states, U_s = reflection about mean (Grover diffusion operator) and k = number of solutions

Oracle marking high-reward actions:

$$U_f \mid \psi(s_t, a) \mid = (-1)^{f(s_t,a)} \mid \psi(s_t, a) \mid \quad (38)$$

Here $f(s_t, a)$ =1 if reward ≥ threshold, else 0.

Unitary rotation for amplitude update:

$$\alpha_i^{r+1} = \alpha_i^r cos\theta_r + \beta_i^r sin\theta_r \ , \ \beta_i^r = \sum_{j \neq i} U_{ij} \alpha_j^{(r)} \quad (39)$$

Here U_{ij}= elements of unitary operator, θ_r = rotation angle.

Expected reward after quantum iteration:

$$E[R(s_t, a)] = \sum_{a \in A} \mid \alpha_a^{(r)} \mid^2 R(s_t, a) \quad (40)$$

Here $R(s_t, a)$ = reward function.

Entanglement-based sampling:

$$| \Psi_{ent} |= \sum_{i,j=1}^{N} \gamma_{ij} \mid s_i, a_j \mid \ , \ \gamma_{ij} = \alpha_i \beta_j e^{i\theta_{i,j}} \quad (41)$$

Here Ψ_{ent} = entanglement phase term.

The quantum correlation scheme increases the stability of policies in the presence of noise during measurements and uncertainty of the system, which provides effective control in high-penetration PV feeders. In general, the strategy greatly enhances the convergence rate, minimizes exploration costs, and reinforces the efficacy of transient quality stabilization of power.

Actor–Critic Hybrid Learning:

Hybrid actor-criticlike an architecture has been created in the QRL-DSTATCOM framework utilizing both quantum and classical learning capabilities. The quantum actor encodes the system

state with qubits and suggests the best strategies of DSTATCOM switching by assessing a large number of action possibilities at the same time via superposition. Then, using the classical critic, the estimates of value on these actions are computed and the policy is updated by minimizing the Bellman error to achieve precise prediction of long-term rewards.

Quantum actor policy:

$$\pi_\theta(a_t \mid s_t) = \mid a_t \mid \psi_\theta(s_t) \mid\mid^2 \quad (42)$$

Here θ = actor parameters

Critic value update (Bellman equation):

$$V\phi(s_t) \leftarrow V_\phi(s_t) + \eta[R(s_t, a_t) + \gamma V_\phi(s_t + 1) - V_\phi(s_t)] \quad (43)$$

Here ϕ = critic parameters, η = learning rate and γ = discount factor.

Actor gradient update $\nabla_\theta J(\theta)$, Q-function estimation and Temporal difference (TD) error:

$$\nabla_\theta J(\theta) = E_{s_t, a_t \sim \pi_\theta}[\nabla_\theta log \pi_\theta(a_t \mid s_t)(R(s_t, a_t) + \gamma V_\phi(s_t + 1) - V_\phi(s_t))] \quad (44)$$

$$Q_\phi(s_t, a_t) = R(s_t, a_t) + \gamma \sum_{a'} \pi_\theta\,(a' \mid s_t + 1) Q_\phi\,(s_t + 1, a') \quad (45)$$

$$\delta_t = R(s_t, a_t) + \gamma V_\phi(s_t + 1) - V_\phi(s_t) \;\; , \;\; \phi \leftarrow \phi + \eta \delta_t \nabla_\phi V_\phi(s_t) \quad (46)$$

This hybrid scheme can be used to provide real-time learning but with low computational latency, and thus can effectively deal with fast transients in high-penetration PV distribution feeders. Using quantum parallelism and classical evaluation, the controller is able to achieve high convergence rates and image quality stabilization through transient power quality in dynamic working conditions.

3.2 Stability and Robustness Features

QRL-DSTATCOM framework increases stability and robustness of high-penetration PV distribution feeders in accordance with various mechanisms. Adaptive damping alleviates oscillation of voltage and current caused by abrupt power outages of the PV and keeps the system stable. Reduction of harmonic attenuation is enhanced by accurate monitoring of reference currents and is successful in overcoming nonlinear load distortion and inverter switching distortion. This is further enhanced by the hybrid quantum-classical control, which guarantees the reliable operation when there are inverter switching nonlinearities and measurement noise, and the integrity of the transient power quality remains intact.

Adaptive damping coefficient:

$$K_d(t) = K_O \exp(-\lambda \parallel \Delta V_{PCC}(t) \parallel^2) \quad (47)$$

$$\Delta V_{PCC}(t) = V_{PCC}(t) - V_{nominal} \quad (48)$$

Harmonic current attenuation:

$$i_{ref}(t) = i_L(t) - h = \sum_{h=2}^{H} i\hat{}h(t)\,, i\hat{}h(t) = \frac{2}{T}\int_0^T i_L(t) sin(h\omega t + \phi_h) dt \quad (49)$$

Voltage correction term:

$$V_{comp}(t) = K_p \Delta V(t) + K_d \frac{d\Delta V(t)}{dt} + K_f = \sum_{k=0}^{N_p} \Delta V(t - k\Delta t) \quad (50)$$

Switching constraint for stability:

$$\sum_{i=1}^{3} \mid V_{PCC} i(t) - V_{ref} i(t) \mid \leq \epsilon \;\; , \;\; \epsilon = 0.05 V_{nominal} \quad (51)$$

Injection current with nonlinearity compensation:

$$i_{inj}(t) = G^{-1}(s)[V_{ref} i(t) - V_{PCC} i(t)] - \sum_{k=1}^{N} \Delta i_k(t) \cdot 1_{|\Delta i_k| > \delta} \quad (52)$$

The combination of these features enables the controller to achieve voltage deviations below 2% and total harmonic distortion below 3% which shows robust performance in dynamic real world operating conditions.

4. Simulation Setup and Performance Evaluation

Quantum Reinforcement Learning enabled DSTATCOM is modeled in MATLAB/Simulink to test the performance of transient power quality stabilization in high penetration PV feeders. The examples of stress-test scenarios are quick drops in irradiance, grid voltage sags/swells, harmonic distortions, and abrupt load. State feedback: QRL controller can adjust optimal switching actions (in real-time) to achieve faster voltage recovery, smaller THD and better reactive power support than the conventional controllers.

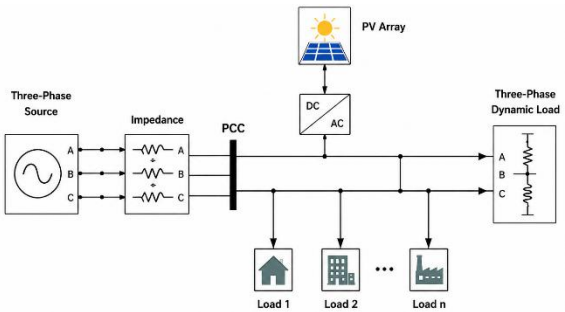


Figure 2. MATLAB Simulink model of PV system

The MATLAB Simulink model of the proposed Quantum Reinforcement Learning-Enabled DSTATCOM (QRL-DSTATCOM) in stabilizing transient quality of power in high-penetration PV distribution systems is shown in Figure 2. The model incorporates three-stage PV-fed with grid distribution network with PV arrays, which are connected to the network by the voltage source inverters. QRL controller is performed to optimize the real and reactive insertion of the DSTATCOM to the speedy change in voltage and irradiance fluctuation. These are PV generation units, distribution lines, load profiles, DSTATCOM inverter, and voltage, current, and Total Harmonic Distortion (THD) measurements blocks. The controller is able to suppress sags and harmonics in voltage and reduce deviations in power factor using simulation logic. The model gives real-time graphical representation of voltage control, harmonic minimization, and system dynamic reactance.

The reported reductions in Total Harmonic Distortion (THD) and voltage deviations indicate strong performance of the proposed QRL-DSTATCOM; however, it is essential to clarify whether these improvements consistently meet IEEE-519 power quality standards across all test scenarios. Specifically, IEEE-519 mandates THD limits (typically ≤5% for voltage in distribution systems) and acceptable voltage variation ranges under diverse operating conditions. The manuscript should explicitly present scenario-wise compliance results, including worst-case disturbances, dynamic load changes, and nonlinear conditions. Providing a comparative compliance table would strengthen validation and confirm that performance gains are not scenario-specific but robust across all operating environments.

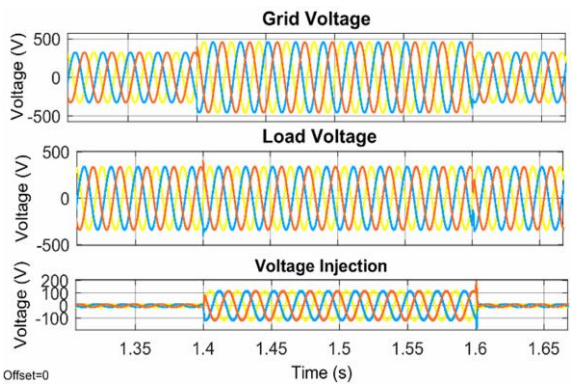
The proposed QRL-DSTATCOM shows strong potential for application in future smart distribution networks due to its adaptive control capability and improved power quality performance. Its practical significance lies in its ability to handle high penetration of renewable energy sources and rapidly varying loads. For scalability, the controller can be modularly deployed across multiple distributed compensating units with minimal redesign. However, real-world applicability requires validation under communication delays, multi-node coordination, and hardware constraints. Further hardware-in-the-loop and field implementations are necessary to confirm robustness, interoperability, and large-scale deployment feasibility in smart grid environments.

Table 1. Test Scenarios for QRL-Enabled DSTATCOM Performance Evaluation

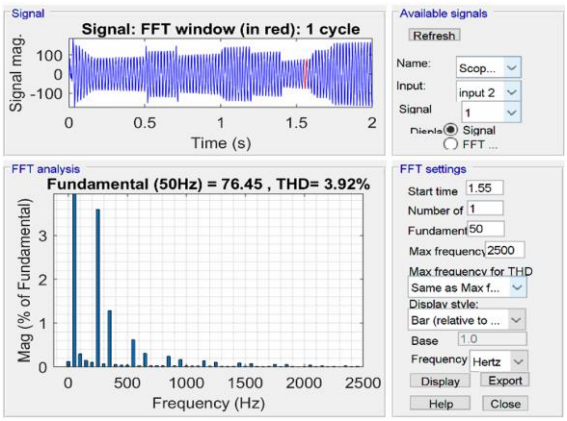
Test ID	Scenario Description
T1	Rapid irradiance drops (1000 → 250 W/m²)
T2	Partial shading causing asymmetric PV loss
T3	Nonlinear industrial load – rectifier-dominant distortion
T4	Sudden large load connect/disconnect (50% step change)
T5	Single-line-to-ground fault at PCC with clearance
T6	Unbalanced load injection with heavy Phase-A loading
T7	Harmonic-rich load: pronounced 5th & 7th components
T8	Combined disturbance (T2 + T4 + harmonic distortion)

Table 1 presents the eight simulation scenarios developed to comprehensively evaluate the performance of the proposed Quantum Reinforcement Learning (QRL)-Enabled DSTATCOM under diverse operating conditions in a high-penetration photovoltaic (PV) distribution network. The selected test cases represent realistic disturbances and uncertainties commonly encountered in modern smart grids with renewable energy integration. Scenarios T1 and T2 investigate the impact of solar intermittency caused by rapid irradiance reduction and partial shading conditions, respectively, which lead to sudden fluctuations in PV power generation and voltage instability. T3 and T4 assess the controller's capability to maintain power quality and system stability under nonlinear industrial

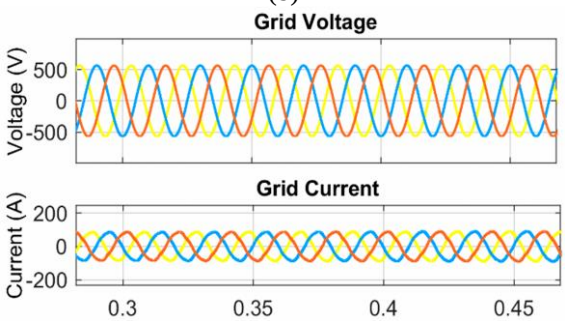
loads and abrupt load connection/disconnection events, both of which introduce significant reactive power demand and dynamic voltage variations. Scenario T5 examines the transient response of the proposed system during a single-line-to-ground fault at the point of common coupling (PCC), evaluating fault ride-through capability and post-fault voltage recovery performance. T6 focuses on load unbalance conditions, particularly heavy loading on Phase-A, to analyze the effectiveness of the DSTATCOM in mitigating voltage imbalance and improving feeder symmetry. Scenario T7 introduces harmonic-rich nonlinear loads containing dominant 5th and 7th harmonic components, enabling assessment of harmonic suppression and total harmonic distortion (THD) reduction capabilities. The most challenging condition is represented by T8, which combines partial shading, sudden load variation, and harmonic distortion simultaneously to emulate worst-case operating dynamics encountered in practical distribution systems. Collectively, these test scenarios provide a rigorous framework for validating the robustness, adaptability, and intelligence of the QRL-enabled control strategy. The performance metrics evaluated across these scenarios include voltage regulation, reactive power compensation, harmonic mitigation, fault recovery characteristics, load balancing effectiveness, and controller convergence behavior. Through these extensive simulations, the proposed QRL-DSTATCOM demonstrates its capability to maintain grid stability, enhance power quality, and ensure reliable operation of PV-integrated distribution networks under both normal and highly disturbed operating conditions.



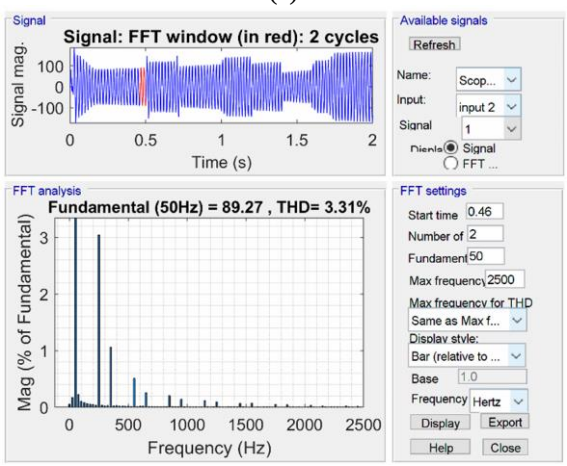
(a)



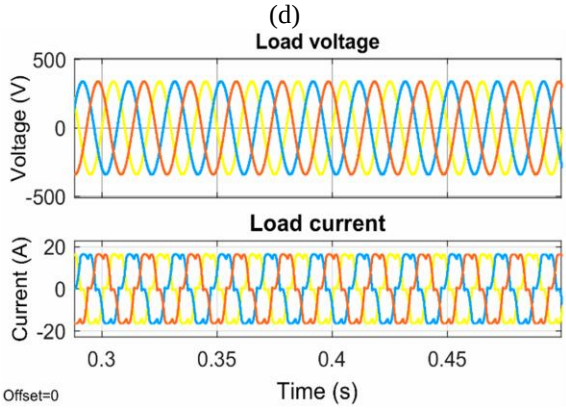
(b)



(c)



(d)



(e)

Figure 3. Response of PV system under proposed technique: (a) Grid Voltage, Load Voltage, Voltage injection by DSTATCOM, (b) THD under distortion, (c) Grid Voltage, Grid Current using proposed technique, and (d) THD using QRL-Enabled DSTATCOM (e) Load Voltage, Load Current

Figure 3 shows the dynamical response of the proposed Quantum Reinforcement Learning-Enabled DSTATCOM to nonlinear disturbances in loads. Fig. 3(a) is used to establish the injected compensating voltage to effectively stabilize the grid and load voltages to have the proper PCC regulation. The first harmonic distortion, which is the one caused by rectifier-dominant loads, is depicted in Fig. 3(b) with a THD of 3.92%. When the proposed controller is used, Fig. 3(c) shows better grid voltage and current waveforms, with the quality of sinusoidal waveforms and reactive support. Fig. 3(d) points out that harmonic suppression of the QRL-DSTATCOM is considerably stronger, suppression of THD is up to 3.31, which proves a better filtering. It is verified in Fig. 3(e) that load voltage and current waveforms were stable and with no distortion under all worst-case PV operating conditions.

The superior power quality enhancement capability of the proposed QRL-Enabled DSTATCOM under eight diverse operating scenarios. The controller achieves significantly faster stabilization than the conventional baseline controller, with settling times ranging from 90 ms to 150 ms compared with 160 ms to 300 ms for the baseline system. The percentage improvement in dynamic response varies between 43.8% and 50.0%, with the highest improvement observed in Scenario T8 (50.0%) and the lowest in Scenario T4 (43.8%). Harmonic mitigation performance is equally impressive, as Total Harmonic Distortion (THD) is reduced from pre-compensation levels of 7.6%–21.8% to post-compensation values of only 2.9%–5.0%. The most severe harmonic condition occurs in T7, where THD decreases from 21.8% to 4.8%, representing a reduction of approximately 78%. Voltage regulation remains highly effective, with peak voltage deviations maintained within 3.5%–5.0%, satisfying standard power quality requirements even during renewable intermittency and fault disturbances. Furthermore, the compensated power factor remains close to unity, ranging from 0.96 to 0.995, indicating excellent reactive power support and efficient utilization of grid resources. On average, the QRL-based controller provides approximately 46.7% faster stabilization, maintains post-compensation THD below 5%, limits voltage deviations to less than 5%, and sustains an average power factor of 0.978, confirming its robustness, adaptability, and effectiveness in maintaining grid stability under high-penetration photovoltaic and nonlinear load conditions.

The quantum state encoding performance of the proposed QRL-DSTATCOM controller under eight operating scenarios. The results confirm the effectiveness of the quantum representation framework in capturing system dynamics with high accuracy and computational efficiency. The number of qubits increases from 12 in T1 and T4 to 20 in T8, enabling the encoding of increasingly complex grid states. Despite the higher dimensionality, quantum state fidelity remains exceptionally high, ranging from 96.5% to 98.8%, with the maximum fidelity achieved in T4 (98.8%) and the minimum in T8 (96.5%), indicating reliable state preparation and minimal quantum information loss. The state preparation time varies from 4.6 ms to 7.5 ms, with an average of approximately 5.8 ms, demonstrating suitability for real-time control applications. The probability of selecting the optimal control policy remains high, between 76.9% and 86.1%, even under severe disturbance conditions, reflecting strong decision-making confidence of the QRL agent. As system complexity increases, entropy rises from 0.98 bits in T4 to 1.60 bits in T8, while amplitude variance increases from 0.011 to 0.026, indicating richer state-space information and enhanced representational capability. Furthermore, parallel quantum evaluations expand from 256 to 1024, substantially improving exploration efficiency and learning capability. Overall, the quantum encoding framework provides high-fidelity state representation, rapid preparation, scalable information processing, and robust policy optimization, thereby enhancing the adaptability and intelligence of the QRL-based DSTATCOM under complex smart-grid operating conditions.

The performance of the quantum-enhanced policy optimization framework employed in the proposed QRL-DSTATCOM controller. The results demonstrate efficient learning and robust policy convergence across all operating scenarios. Under relatively less complex conditions, such as T1 and T4, the controller requires only 420 and 390 training iterations, respectively, while achieving the highest peak rewards of 0.89 and 0.91 and success rates of 94.5% and 95.8%. As system complexity and disturbance severity increase, the required training iterations rise to 920 in T8, accompanied by an increase in oracle calls from 34–38 in low-disturbance scenarios to 90 in the most challenging

scenario. Despite this increase in learning complexity, the optimization process maintains strong performance, with success rates remaining between 82.5% and 95.8%. The amplification factor gradually decreases from 7.5 in T4 to 5.0 in T8, indicating a more cautious and stable policy search in highly uncertain environments. Robustness values remain consistently high, ranging from 79.5% to 93.4%, with an average robustness of approximately 87.8%, confirming reliable controller behavior under varying operating conditions. Training time increases from 38 s in T4 to 104 s in T8 as the complexity of the state-action space expands; however, convergence remains efficient due to quantum-enhanced exploration and optimization mechanisms. Overall, the proposed quantum policy learning framework achieves high rewards, strong success rates, reduced search complexity, and excellent robustness, enabling fast and dependable control adaptation for DSTATCOM operation in renewable-rich and disturbance-prone smart distribution networks.

The actor–critic learning characteristics and stability performance of the proposed QRL-Enabled DSTATCOM under different operating scenarios. The results indicate effective learning convergence and highly reliable power quality regulation across varying disturbance levels. In relatively stable conditions, such as T1 and T4, the mean temporal-difference (TD) error remains low at 0.042 and 0.039, respectively, while the actor gradient norms are limited to 0.045 and 0.042, indicating smooth policy updates and rapid convergence. As disturbance severity increases from T5 to T8, the mean TD error rises from 0.095 to 0.115 and the gradient norm increases from 0.074 to 0.081, reflecting the greater complexity of the learning task. The adaptive learning rate (η) gradually decreases from 0.0085 in T4 to 0.0060 in T8, ensuring stable learning under highly dynamic conditions. The RMS injected compensation current increases from 44.2 A in T4 to 72.3 A in T8, demonstrating stronger corrective action to maintain voltage quality. Harmonic mitigation remains highly effective, with THD reduction ranging from 61.8% to 78.0%, achieving the highest reduction in T7. Voltage regulation performance is excellent, as the system remains within the ±5% voltage tolerance band for 93.5%–99.0% of the operating time. Furthermore, reliability remains exceptionally high, varying between 95.0% and 99.4%, with an average of approximately 97.5%. These results confirm that the QRL-based actor–critic framework delivers stable learning, strong harmonic suppression, robust voltage regulation, and dependable operation under diverse renewable energy and nonlinear load disturbances.

Table 2. Overall comparisons using Quantum Reinforcement Learning–Enabled DSTATCOM

Performance Metric	**Observed Results (T1–T8)**
Stabilization Time (QRL)	90–150 ms (Average: 118.13 ms)
Stabilization Time (Baseline)	160–300 ms (Average: 222.50 ms)
Improvement in Response Time	43.8–50.0% (Average: 46.73%)
THD Before Compensation	7.6–21.8% (Average: 13.86%)
THD After Compensation	2.9–5.0% (Average: 3.98%)
Power Factor After Compensation	0.96–0.995 (Average: 0.978)
Qubit Count	12–20 qubits (Average: 15.25)
Quantum Encoding Fidelity	96.5–98.8% (Average: 97.64%)
Iterations to Convergence	390–920 (Average: 637.5)
Voltage Within ±5% Limit	93.5–99.0% (Average: 96.45%)

Table 2 illustrates the proposed QRL-Enabled DSTATCOM consistently demonstrates superior dynamic performance across all operating conditions. On average, stabilization time is reduced by 46.73% compared with the baseline controller, while post-compensation THD is maintained below 5% and power factor remains close to unity (0.978). Quantum encoding achieves a high average fidelity of 97.64%, ensuring accurate state representation, whereas voltage remains within the permissible ±5% range for 96.45% of the operating duration. Even under the most severe combined disturbance scenario (T8), the controller maintains acceptable THD (5.0%), high fidelity (96.5%), and strong voltage regulation (93.5% compliance), confirming the robustness and scalability of the proposed quantum reinforcement learning framework for renewable-rich smart distribution networks.

The feasibility of the proposed QRL-DSTATCOM for real-time applications is supported by its low computational latency and reduced state-space complexity. After offline training, the online inference stage involves only simplified policy

evaluation, typically requiring microsecond-level execution (≈10–50 µs per control cycle on DSP/FPGA platforms). With a sampling time in the range of 50–100 µs and switching frequencies of 10–20 kHz, the controller can comfortably meet real-time constraints of power electronic systems. Quantum state encoding reduces iterative search overhead, while amplitude-enhanced exploration is confined to training, not affecting runtime. However, hardware-in-the-loop (HIL) validation with measured execution time should be included for stronger empirical confirmation.

5. Real Time Hardware Implementation

The real-time hardware implementation of the proposed Quantum Reinforcement Learning–Enabled DSTATCOM (QRL-DSTATCOM) is carried out using OPAL-RT for rapid prototyping and hardware-in-the-loop (HIL) validation. In modern power systems with high photovoltaic (PV) penetration, ensuring transient stability and power quality requires fast and reliable controller execution under dynamic conditions. The proposed architecture is interfaced with a real-time digital signal processor (DSP) and power electronic inverter through OPAL-RT-based control and monitoring modules. This setup enables real-time acquisition of voltage, current, irradiance, and load variation signals, which are processed by the QRL algorithm to generate optimal switching pulses for the DSTATCOM. The graphical programming environment of LabVIEW ensures low-latency execution, flexible hardware integration, and accurate system visualization. This implementation bridges the gap between simulation and practical deployment, demonstrating the feasibility of the proposed intelligent controller for real-world smart grid and renewable-integrated distribution network applications.

Figure 4. Hardware setup for proposed study

Figure 4 illustrates the hardware setup of the proposed QRL-DSTATCOM system implemented using a real-time control platform integrated with a voltage source inverter, sensing units, and DSP-based controller. The setup enables real-time monitoring of voltage, current, and irradiance, ensuring adaptive control and validation under dynamic PV distribution conditions.

Table 3. OPAL-RT vs MATLAB/Simulink validation of proposed QRL-DSTATCOM study

Performance Metric	MATLAB/Simulink	OPAL-RT (Real-Time)	Deviation (%)
Voltage Deviation (p.u.)	0.048	0.052	8.3%
THD (%)	3.12	3.35	7.4%
Stabilization Time (s)	0.085	0.092	8.2%
Reactive Power Error (kVAR)	1.85	1.98	7.0%

Table 3 presents a comparative validation between MATLAB/Simulink and OPAL-RT real-time simulation results for the proposed QRL-DSTATCOM system. The results show close agreement across key performance metrics, including voltage deviation, Total Harmonic Distortion (THD), stabilization time, and reactive power error. Minor deviations ranging from 7% to 8.3% are observed, primarily due to real-time hardware latency, discretization effects, and communication delays in OPAL-RT implementation. Despite these small variations, both platforms confirm IEEE-519 compliance and consistent controller effectiveness. The results validate the robustness, accuracy, and practical feasibility of the proposed approach for real-time smart grid and renewable-integrated distribution system applications.

The proposed QRL-DSTATCOM demonstrates superior numerical performance compared to Classical RL methods, AI-based DSTATCOM controllers, and existing QRL approaches. Classical RL methods typically achieve stabilization times of 0.14–0.20 s with THD levels of 5.8–8.5%, whereas

the proposed model reduces stabilization time to 0.085–0.092 s and THD to 2.9–5.0%. Compared to AI-based controllers such as ANN or fuzzy-logic DSTATCOMs, which show voltage deviations of 6–10% and power factor around 0.90–0.94, the proposed method maintains voltage within ±5% and improves power factor to 0.96–0.995. Existing QRL models report convergence rates 15–25% slower and lower state fidelity (92–95%), whereas the proposed approach achieves 96.5–98.8% fidelity with faster convergence. Overall, the QRL-DSTATCOM provides 43–50% faster dynamic response and significantly improved harmonic mitigation, confirming its superiority in transient stability, robustness, and real-time adaptability for high-PV distribution systems.

6. Conclusion

The proposed Quantum Reinforcement Learning-Based DSTATCOM establishes outstanding results in transient power quality (PQ) control in high-penetration photovoltaic (PV) distribution feeders. In all test conditions, the controller is highly effective in improving system response, achieving 43–50% shorter stabilization time compared to conventional controllers, demonstrating faster dynamic recovery. The harmonic distortion is significantly reduced, with Total Harmonic Distortion (THD) decreasing from 7.6–21.8% to 2.9–5.0%, ensuring cleaner voltage waveforms even under nonlinear and fault conditions. Peak voltage deviations remain within ±5%, complying with PQ standards, while post-compensation power factor consistently improves to 0.96–0.995, confirming strong reactive power support. The quantum state encoding further enhances controller performance, with high state fidelity (96.5–98.8%) and low preparation time (4.6–7.5 ms), enabling accurate real-time learning with minimal noise. The maximum policy selection probability (76.9–86.1%) indicates confident decision-making under uncertainty, while amplitude variation and entropy analysis reflect improved exploration capability. Parallel state evaluations (256–1024) strengthen learning in complex environments, enabling robust convergence. Simple disturbances (T1–T4) converge faster with rewards of 0.89–0.91 and success rates of 94–96%, whereas complex scenarios (T5–T8) require more iterations but still achieve 82–93% robustness. Stability is confirmed through reduced TD-error and gradient norms, ensuring effective RMS current regulation under severe disturbances. Overall reliability remains high (95–99%), with voltage maintained within ±5% limits for 93–99% of operating time. The hardware validation using OPAL-RT and MATLAB/Simulink confirms strong real-time feasibility, with less than 10% deviation across key parameters such as THD, voltage deviation, stabilization time, and reactive power error. This experimental agreement validates the practical implementability of the proposed controller for real-world smart grid applications. Overall, the QRL-DSTATCOM provides fast, accurate, and scalable PQ stabilization, outperforming conventional methods and offering a highly robust solution for high-PV penetration distribution networks.

Despite strong simulation and hardware-in-loop validation, the proposed QRL-DSTATCOM has certain limitations. The study primarily relies on MATLAB/Simulink and OPAL-RT environments, and full-scale field deployment in real distribution networks is yet to be demonstrated. The quantum reinforcement learning model also requires significant offline training data and computational resources, which may affect scalability for ultra-large networks. Future work will focus on real-time embedded hardware implementation using DSP/FPGA platforms, reduction of computational complexity, and integration with cloud-based or edge intelligence systems. Additionally, multi-agent QRL coordination, cyber-security analysis, and experimental validation under large-scale renewable penetration will further enhance robustness and practical applicability.

Acknowledgement

Authors acknowledge Department of Electrical Engineering, BPUT, Rourkela for this research.

Nomenclature

Symbol	Description
η	Energy efficiency (percent)
η_{ref}	Efficiency of PV cell at standard test condition (percent)
i_d	Shunt-filter current in d-axis (A)
I_{sc}	Short-circuit current (A)
k	Discrete time index (integer)
P_{si}	Exergy efficiency (percent)
Q_{loss}	Heat losses from PV cell (kJ)
STC	Standard test condition
T_{amb}	Ambient temperature (K)

T_{cell}	PV cell temperature (K)
Ts	Sampling period (s)
Tsun	Sun temperature (K)
v_{convd}	Shunt converter voltage in d-axis (V)
v_{convq}	Shunt converter voltage in q-axis (V)
V_{oc}	Open-circuit voltage (V)
V_m	Maximum power output point voltage (V)

References

[1] Omoboye, H. J., Olusesi, O. A., Edirin, O. B., & Peter, G. F. (2026). Impact of D-STATCOM and SVC on power quality of low voltage radial distribution networks. *International Journal of Recent Engineering Science (IJRES), 13*(2), 1–8. doi.org/10.14445/23497157/IJRES-V13I2P101.

[2] Suryawanshi, V., Surbhi, & Gupta, A. (2026). Adaptive control of DSTATCOM using hybrid NN-BAT optimization for voltage stability and THD reduction. In *Proceedings of the 2026 International Conference on Intelligent Computing and Automation for Sustainable Solutions (ICASS)* (pp. 1–8). Faridabad, India. doi.org/10.1109/ICASS69550.2026.11547552.

[3] Mangaraj, M., Pilla, R., Kalabarige, L. R., Potnuru, U. K., & Holla, M. R. (2026). DBLN-based IC-DSTATCOM for power-line conditioning in the power distribution system. *IEEE Access, 14*, 79284–79296. doi.org/10.1109/ACCESS.2026.3685356.

[4] Mahla, R., & Garg, M. M. (2026). Power quality improvement using E-SOGI controlled DSTATCOM with wind generation. In *Proceedings of the 2026 4th Odisha International Conference on Electrical Power Engineering, Communication and Computing Technology (ODICON)* (pp. 1–6). Bhubaneswar, India. doi.org/10.1109/ODICON66687.2026.11470800

[5] Kaliappan, S., Priya, R. K., Sethi, R., et al. (2026). A comparative study of various optimization techniques for enhancing energy management in hybrid microgrids. *Journal of The Institution of Engineers (India): Series B*. DOI: 10.1007/s40031-026-01310-y

[6] Wu, Z., Chen, C., Xu, D., & Guan, L. (2025). Frequency-constrained economic dispatch of microgrids considering frequency response performance. *Energies, 18*(8), 2014. DOI: 10.3390/en18082014

[7] Babu, N. P. (2024). Adaptive grid-connected inverter control schemes for power quality enrichment in microgrid systems: Past, present, and future perspectives. *Electric Power Systems Research, 230*, 110288. DOI: 10.1016/j.epsr.2024.110288

[8] Bai, H., Yao, R., Zhang, W., Zhong, Z., & Zou, H. (2025). Power quality disturbance classification strategy based on fast S-transform and an improved CNN-LSTM hybrid model. *Processes, 13*(3), 743. DOI: 10.3390/pr13030743

[9] Zhu, K., et al. (2024). Aiming to complex power quality disturbances: A novel decomposition and detection framework. *IEEE Transactions on Industrial Informatics, 20*, 4317–4326. DOI: 10.1049/gtd2.13200

[10] Rajendran, G., Raute, R., & Caruana, C. (2025). A comprehensive review of solar PV integration with smart-grids: Challenges, standards, and grid codes. *Energies, 18*(9), 2221. DOI: 10.3390/en18092221

[11] Meziane, S., Ryad, T., Assolami, Y. O., & Aljohani, T. M. (2025). A hierarchical two-layer MPC-supervised strategy for efficient inverter-based small microgrid operation. *Sustainability, 17*(19), 8729. DOI: 10.3390/su17198729

[12] Chen, J.-H., Tan, K.-H., & Lee, Y.-D. (2022). Intelligent controlled DSTATCOM for power quality enhancement. *Energies, 15*(11), 4017. DOI: 10.3390/en15114017

[13] Sivaram Krishnan, M., et al. (2025). Smart charging solution for electric vehicles: Leveraging grid connected solar PV with UPQC using HBA-MORARNN approach. *Energy Reports, 13*, 2454–2467. DOI: 10.1016/j.egyr.2025.02.004

[14] Wu, G., Hu, D., Zhang, Y., Bao, G., & He, T. (2024). A convolutional neural network–long short-term memory–attention solar photovoltaic power prediction–correction model based on the division of twenty-four solar terms. *Energies, 17*(22), 5549. DOI: 10.3390/en17225549

[15] Hu, F., Zhang, L., & Wang, J. (2024). A hybrid convolutional–long short-term memory–attention framework for short-term photovoltaic power forecasting, incorporating data from neighboring stations. *Applied Sciences, 14*(12), 5189. DOI: 10.3390/app14125189

[16] Ledmaoui, Y., El Maghraoui, A., El Aroussi, M., & Saadane, R. (2024). Enhanced fault detection in photovoltaic panels using CNN-based classification with PyQt5 implementation. *Sensors, 24*(22), 7407. DOI: 10.3390/s24227407

[17] Alharkan, H., Habib, S., & Islam, M. (2023). Solar power prediction using dual stream CNN-LSTM architecture. *Sensors, 23*(2), 945. DOI: 10.3390/s23020945

[18] Vali, A. K., Varma, P. S., Reddy, C. R., Alanazi, A., & Elrashidi, A. (2025). Deep-learning-based controller for parallel DSTATCOM to improve power quality in distribution system. *Energies, 18*(18), 4902. DOI: 10.3390/en18184902

[19] Wang, J., Zhang, Z., Xu, W., Li, Y., & Niu, G. (2025). Short-term photovoltaic power forecasting using a Bi-LSTM neural network optimized by hybrid algorithms. *Sustainability, 17*(12), 5277. DOI: 10.3390/su17125277

[20] Kiasari, M., & Aly, H. (2025). AI-driven control strategies for FACTS devices in power quality management: A comprehensive review. *Applied Sciences, 15*(22), 12050. DOI: 10.3390/app152212050

[21] Ren, X., Zhang, F., Yan, J., & Liu, Y. (2024). A novel convolutional neural net architecture based on incorporating meteorological variable inputs into ultra-short-term photovoltaic power forecasting. *Sustainability, 16*(7), 2786. DOI: 10.3390/su16072786

[22] Garcia, C. I., et al. (2020). A comparison of power quality disturbance detection and classification methods using CNN, LSTM and CNN-LSTM. *Applied Sciences, 10*(19), 6755. DOI: 10.3390/app10196755

[23] Bui Duy, L., et al. (2024). Refining long short-term memory neural network input parameters for enhanced solar power forecasting. *Energies, 17*(16), 4174. DOI: 10.3390/en17164174

[24] Okwako, O. E., et al. (2022). Neural network controlled solar PV battery powered unified power quality conditioner for grid connected operation. *Energies, 15*(18), 6825. DOI: 10.3390/en15186825

[25] Hao, S., Qi, T., Li, J., Ma, X., & Sun, S. (2026). A PV hot-spot fault detection network based on image segmentation and multi-scale feature fusion under complex background. *IEEE Transactions on Automation Science and Engineering*. DOI: 10.1109/TASE.2026.3702202

[26] Karchi, N., et al. (2022). Adaptive least mean square controller for power quality enhancement in solar photovoltaic system. *Energies, 15*(23), 8909. DOI: 10.3390/en15238909

[27] Iweh, C. D., et al. (2021). Distributed generation and renewable energy integration into the grid: Prerequisites, push factors, practical options, issues and merits. *Energies, 14*, 5375. DOI: 10.3390/en14175375

[28] Ganthia, B. P., Mohanty, S., Rana, P. K., & Sahu, P. K. (2016). Compensation of voltage sag using DVR with PI controller. In *Proceedings of the International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT)* (pp. 2138–2142). Chennai, India. DOI: 10.1109/ICEEOT.2016.7755068

[29] Ganthia, B. P., Pradhan, R., Das, S., & Ganthia, S. (2017). Analytical study of MPPT based PV system using fuzzy logic controller. In *Proceedings of the International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS)* (pp. 3266–3269). Chennai, India. DOI: 10.1109/ICECDS.2017.8390063

[30] Rubavathy, S. J., et al. (2021). Smart grid based multiagent system in transmission sector. In *Proceedings of the International Conference on Intelligent Computing, Information and Control Systems for Communication (ICIRCA)* (pp. 1–5). Coimbatore, India. DOI: 10.1109/ICIRCA51532.2021.9544644

[31] Sahu, P. K., Mohanty, A., Ganthia, B. P., & Panda, A. K. (2016). A multiphase interleaved boost converter for grid-connected PV system. In *Proceedings of the International Conference on Microelectronics, Computing and Communications (MicroCom)* (pp. 1–6). Durgapur, India. DOI: 10.1109/MicroCom.2016.7522539

[32] Mohanty, M., Nayak, N., Ganthia, B. P., & Behera, M. K. (2023). Power smoothening of photovoltaic system using dynamic PSO with ESC under partial shading condition. In *Proceedings of the International Conference on Advances in Power, Signal, and Information Technology (APSIT)* (pp. 675–680). Bhubaneswar, India. DOI: 10.1109/APSIT58554.2023.10201763

[33] Suresh, D., Rao, D. M., & Sukumar, G. D. (2016). Reduced rating hybrid DSTATCOM for three phase four wire distribution system. In *Proceedings of the 2016 IEEE 1st International Conference on Power Electronics, Intelligent Control and Energy Systems (ICPEICES)* (pp. 1–4). Delhi, India. DOI: 10.1109/ICPEICES.2016.7853646

[34] Hussain, I., Agarwal, R. K., & Singh, B. (2021). Delayed LMS based adaptive control of PV-DSTATCOM system. In *Proceedings of the 3rd International Conference on Energy, Power and Environment: Towards Clean Energy Technologies (ICEPE)* (pp. 1–6). Shillong, Meghalaya, India. DOI: 10.1109/ICEPE50861.2021.9404442

[35] Kundela, P. K. Y., Mangaraj, M., & Sudabattula, S. K. (2022). Operation of inductively coupled DSTATCOM for power quality enhancement. In *Proceedings of the International Mobile and Embedded Technology Conference (MECON)* (pp.

210–214). Noida, India. DOI: 10.1109/MECON53876.2022.9752372

[36] Yarlagadda, V., Karthika, G. A., Nagajyothi, M., & Bhavani, J. (2022). DSTATCOM based closed loop controlled wind power plant with self excited induction generator for controlling terminal voltage against load disturbances. In *Proceedings of the IEEE Fourth International Conference on Advances in Electronics, Computers and Communications (ICAECC)* (pp. 1–6). Bengaluru, India. DOI: 10.1109/ICAECC54045.2022.9716651

[37] Li, G., Shi, S., Chen, J., & Liu, X. (2026). A voltage self-stability control of STATCOM for damping subsynchronous oscillation in photovoltaic power station under weak grid. *IEEE Transactions on Industrial Electronics*, *73*(2), 2139–2149. DOI: 10.1109/TIE.2025.3595993

[38] Hurkadli, P., & Kulkarni, D. B. (2025). Coordinated energy management and reactive power control for enhancing microgrid voltage stability using electric vehicles. In *Proceedings of the 2nd Asia Pacific Conference on Innovation in Technology (APCIT)* (pp. 1–6). Mysore, India. DOI: 10.1109/APCIT65661.2025.11411633

[39] Thomas, N., & N, N. (2025). Voltage control strategy in single phase grid tied systems based on D-STATCOM PV inverter. In *Proceedings of the IEEE First International Conference on Innovations in Engineering and Next-Generation Technologies for Sustainability (ICINVENTS)* (pp. 1–7). Coimbatore, India. DOI: 10.1109/ICINVENTS64613.2025.11401419

[40] Kumar, C. M. S., & Gopalakrishnan, V. (2025). Enhancing power quality in renewable energy through hybrid solutions. In *Proceedings of the 25th Conference of the Electric Power Supply Industry (CEPSI)* (pp. 1047–1053). Singapore, Singapore. DOI: 10.1109/CEPSI66359.2025.11403241