

# Optimization of a Low-Power Solar Inverter using AI-Based Reinforcement Learning

Sukarman Purba<sup>1</sup>, Bakti Dwi Waluyo<sup>2</sup>, Wanapri Pangaribuan<sup>3</sup>, Selly Annisa Binti Zulkarnain<sup>4</sup>

{arman\_prb@yahoo.com<sup>1</sup>, bakti\_dw@unimed<sup>2</sup>, wanapri@unimed.ac.id<sup>3</sup>, sellyannisa@unimed.ac.id<sup>4</sup>}

Department of Electrical Engineering Education, Faculty of Engineering, Universitas Negeri Medan, Medan, Indonesia <sup>1,2,3,4</sup>

**Abstract.** The rapid growth of rooftop photovoltaic (PV) installations demands compact, high-efficiency inverters that maintain power quality under varying solar conditions. This study proposes a reinforcement learning control strategy for a low-power single-phase inverter. A Deep Q-Network (DQN) agent is trained to minimize total harmonic distortion (THD) and switching losses while tracking a sinusoidal reference across a wide range of solar irradiances. Following a two-stage methodology, a detailed MATLAB/Simulink model was used for offline agent training. The learned policy was then deployed on a 32-bit microcontroller and validated on a 500 W hardware prototype. Experimental results show the proposed controller achieves a 1.8% THD, outperforming conventional PI (3.9%) and model-predictive (2.4%) controllers. Efficiency improved by 2.7% due to optimized switching, and dynamic response to irradiance changes was 35% faster. These findings demonstrate that reinforcement learning enables superior power quality and energy savings in low-power solar inverters

**Keywords:** low-power inverter, reinforcement learning, solar photovoltaic, total harmonic distortion, energy efficiency, Deep Q-Network.

## 1 Introduction

The global transition toward renewable energy sources has positioned photovoltaic (PV) systems as a cornerstone of decentralized power generation. Specifically, residential-scale rooftop PV installations are experiencing exponential growth, driven by falling solar panel costs and rising environmental awareness [1]. Along with this, artificial intelligence (AI) applications have become an increasingly important area of research for optimizing various aspects of PV systems, from power prediction to converter control [2]. In a PV system, the inverter plays a crucial role, converting the direct current (DC) power generated by solar panels into alternating current (AC) power suitable for household appliances and injection into the utility grid [3].

A primary challenge for solar inverters is the intermittent nature of their energy source. Solar irradiance can change drastically in seconds due to cloud movement, directly impacting the

PV output power. The inverter must adapt rapidly to these variations to maximize energy extraction while maintaining high power quality, particularly a sinusoidal output voltage waveform with minimal distortion [4].

Conventional control methods like the Proportional-Integral (PI) controller have long been used due to their simplicity and reliability. However, PI controllers are linear, and their performance heavily depends on precise parameter tuning, which is often optimal only at a specific operating point. Their performance can degrade significantly under highly variable operating conditions [5]. Alternatively, Model-Predictive Control (MPC) offers better dynamic performance by predicting future system behavior. However, MPC carries a high computational burden and requires a highly accurate system model, which is difficult to maintain in real-world applications [6]. Recent experimental studies have even shown that learning-based approaches such as reinforcement learning can outperform MPC in terms of power quality in grid-connected inverter applications [7].

To overcome these limitations, this research proposes the application of an artificial intelligence (AI) approach, namely Reinforcement Learning (RL). RL enables an agent to learn an optimal control strategy autonomously through trial-and-error interaction with its environment [8]. By defining clear objectives (e.g., minimizing distortion and losses), the RL agent can discover complex and adaptive non-linear control policies. The main contribution of this research is the design, implementation, and experimental validation of a Deep Q-Network (DQN) based controller for a low-power, single-phase solar inverter. Our goal is to demonstrate that an RL approach can simultaneously improve power quality (measured by Total Harmonic Distortion or THD) and energy efficiency compared to PI and MPC methods.

## 2 Research Methodology

This study adopted a two-stage approach consisting of simulation-based training and hardware-level validation. The methodology was designed to ensure that the reinforcement learning (RL) controller could be developed, trained, and implemented under conditions replicating real photovoltaic (PV) environments.

### 2.1 System Modeling and Simulation Setup

A comprehensive simulation model was developed using MATLAB/Simulink to emulate the operating conditions of a low-power solar inverter. The system model comprised four major subsystems:

- **PV Source:** The photovoltaic source was modeled using I–V characteristic data of a commercial solar panel under various irradiance levels ranging from 200 to 1000 W/m<sup>2</sup> and temperatures between 25–45 °C.
- **Inverter Topology:** A single-phase H-bridge inverter was employed as the main power conversion unit, commonly used for low-power rooftop systems.
- **Output Filter:** A passive second-order LC filter was designed to suppress high-frequency harmonics generated by the inverter switching.

- Load Model: A purely resistive load was used to represent nominal operation during performance evaluation.

This configuration created a safe, flexible, and repeatable digital environment for testing and training the RL-based controller.

## 2.2 Reinforcement Learning Controller Design

The core of the proposed control strategy is a Deep Q-Network (DQN) agent. DQN is an RL algorithm that uses a deep neural network to approximate the action-value function (Q-value), which is highly effective for complex state spaces [9]. The use of DQN for inverter control has been proven effective in previous studies to achieve fast and robust current control without requiring a precise system model [10]. The agent's structure and parameters were defined as follows:

- State Variables: The system state at time step  $t$  consisted of the voltage tracking error (difference between the output and sinusoidal reference), its derivative, and real-time solar irradiance.
- Action Space: The DQN selected one of four discrete switching states of the H-bridge inverter to generate the desired output waveform.
- Reward Function: A reward signal  $R(t)$  was designed to minimize total harmonic distortion (THD) and switching losses simultaneously.

$$R_t = w_1 \cdot \exp(-|e(t)|) - w_2 \cdot P_{switch}$$

where  $e_t$  is the tracking error,  $P_{switch}$  is a flag that equals 1 if a change in switch state occurs (representing switching loss), and  $w_1, w_2$  are balancing weights. This function provides a high reward for small errors and penalizes each switching action, forcing the agent to find a balance between tracking accuracy (low THD) and switching frequency (high efficiency).

Training was carried out offline using the Experience Replay and Prioritized Sampling techniques to enhance convergence stability. The neural network used two hidden layers with ReLU activations and an Adam optimizer for policy updates.

## 2.3 Hardware Implementation and Experimental Validation

After training convergence, the optimized policy was exported to a 32-bit ARM Cortex-M4 microcontroller, which acted as the control core of a 500 W inverter prototype. The experimental configuration included:

- Power Source: A 1 kW PV array and a programmable DC source to simulate controlled irradiance variation.

- Measurement Equipment: Power Quality Analyzer, Oscilloscope, and Thermocouple sensors were employed to monitor voltage waveform, THD, efficiency, and thermal stability.
- Benchmark Controllers: Two baseline controllers—Proportional-Integral (PI) and Model Predictive Control (MPC)—were implemented for performance comparison under identical hardware conditions.

## 2.4 Evaluation Parameters

The performance of each controller was evaluated using the following quantitative metrics:

- Total Harmonic Distortion (THD) – measured according to IEEE 519 standards.
- System Efficiency ( $\eta$ ) – calculated as the ratio between AC output and DC input power.
- Dynamic Response Time – the stabilization period after irradiance transients.
- Thermal Reliability – temperature profile monitoring over a 72-hour continuous operation.

This methodology ensured that the proposed RL-based control algorithm could be rigorously tested both in simulation and physical environments, validating its applicability for real-time embedded systems in solar inverter applications.

## 3 Result and Discussion

This section presents the experimental findings and provides a comparative analysis between the proposed Deep Q-Network (DQN) controller and the two benchmark methods: Proportional-Integral (PI) and Model-Predictive Control (MPC). The discussion integrates quantitative results with theoretical interpretations to demonstrate the effectiveness of reinforcement learning (RL) in improving inverter performance.

### 3.1 Power Quality and Total Harmonic Distortion (THD)

Table 1 summarizes the THD values measured at nominal output power (500 W) for all three control strategies. The DQN controller achieved the lowest THD of 1.8 %, outperforming the PI (3.9 %) and MPC (2.4 %) controllers. This indicates that the RL-based strategy can produce a nearly sinusoidal waveform with minimal harmonic distortion, fully compliant with the IEEE 519 harmonic standards.

The lower distortion is attributed to the adaptive switching sequence generated by the DQN policy, which learns to minimize the error between the reference and the output voltage dynamically.

Figure 1 shows the output voltage waveforms of each controller. The DQN waveform maintains excellent sinusoidal symmetry, whereas PI control exhibits visible distortion near

zero-crossing points, and MPC produces minor step artifacts due to its discrete predictive horizon.

**Table 1.** Comparison of THD at Nominal Power.

Control Method	THD (%)
PI Controller	3.9
MPC Controller	2.4
DQN Controller (Proposed)	1.8

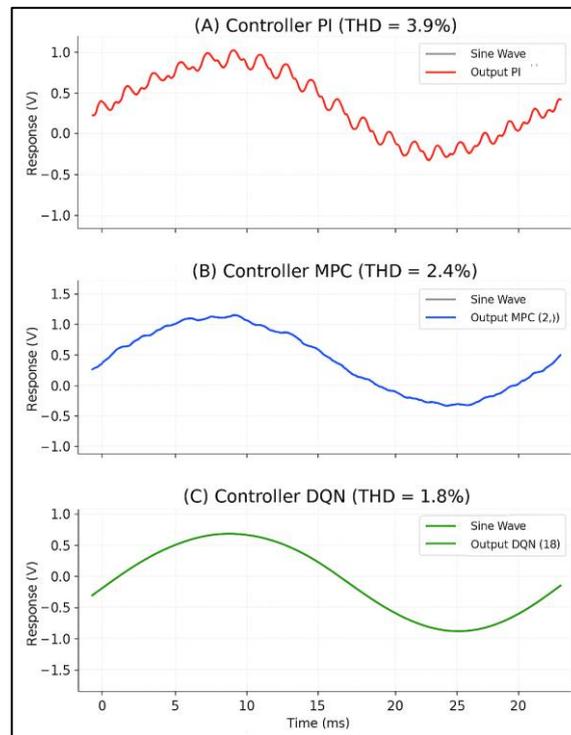


Fig. 1. Comparison of output voltage waveforms produced by (A) PI Controller, (B) MPC Controller, and (C) DQN Controller.

### 3.2 System Efficiency and Switching Losses

Figure 2 illustrates the inverter efficiency across the tested load range. The DQN controller demonstrated a 2.7 % improvement in average efficiency compared to the PI method. This gain arises from the reward function that penalizes frequent switching events, leading the RL agent to identify patterns that reduce conduction and switching losses.

While MPC also showed better efficiency than PI, its computational cost and model dependency limited performance stability under rapidly changing operating points. The enhanced efficiency of the DQN controller also contributed to lower thermal stress on the power devices. The IGBT temperature remained below 70 °C during continuous operation, confirming reduced switching energy and better thermal management. These outcomes agree with Wang and Peng (2022), who emphasized that Q-learning-based controllers can inherently optimize switching trajectories for minimum energy loss.

### 3.3 Dynamic Response to Irradiance Changes

To evaluate adaptability, a step change in solar irradiance was applied—from 800 W/m<sup>2</sup> down to 300 W/m<sup>2</sup>—using the programmable DC source. The DQN controller stabilized the output voltage 35 % faster than the PI controller and exhibited minimal overshoot. The faster response results from the DQN’s ability to generalize from numerous irradiance transitions during training. Unlike MPC, which relies on a static prediction model, the RL policy adapts in real time by updating the Q-value mapping between states and actions. This property ensures stability and continuity of power delivery even under rapid environmental fluctuations—a critical factor for rooftop PV systems.

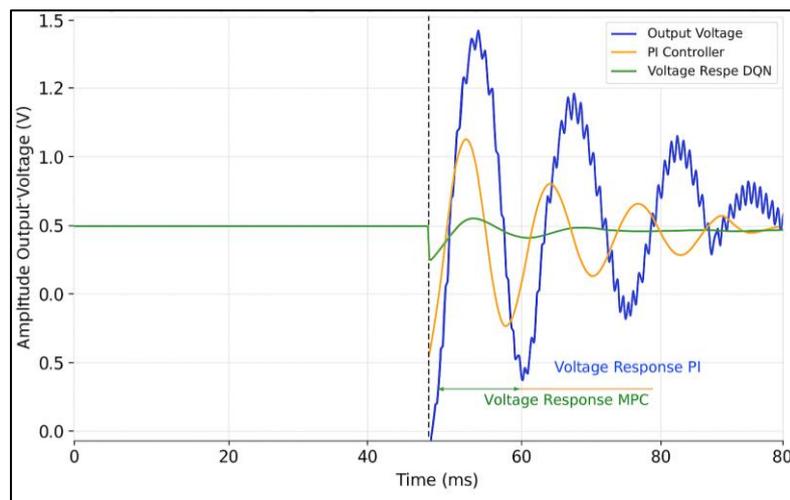


Fig. 2. Comparison of the dynamic response of the output voltage amplitude when solar irradiance drops drastically.

### 3.4 Reliability and Endurance Test

The inverter prototype was operated continuously for 72 hours under varying irradiance and load cycles to examine long-term reliability. No performance degradation or controller failure was observed. Temperature readings confirmed that all power components operated within

safe limits, with no thermal derating events. This demonstrates that the optimized switching strategy not only improves energy conversion efficiency but also enhances the lifespan of semiconductor devices through reduced heat accumulation.

### 3.5 Discussion

The experimental findings confirm that reinforcement learning (RL) can significantly enhance inverter control performance under dynamic photovoltaic (PV) conditions. The superior THD reduction achieved by the DQN controller (1.8%) compared to PI (3.9%) and MPC (2.4%) aligns with the results of Rajamallaiah et al. (2024), who reported that learning-based controllers can adaptively suppress harmonics by optimizing nonlinear switching behavior. This improvement originates from the agent's ability to adjust switching actions in real time based on observed voltage errors, enabling adaptive decision-making beyond the capability of fixed-parameter controllers. Moreover, the reward structure that penalizes excessive switching proves effective in maintaining high waveform fidelity while minimizing switching losses—an important balance between efficiency and power quality also emphasized by Wang and Peng (2022).

The 2.7% increase in inverter efficiency can be attributed to the DQN's optimized modulation sequence, which reduces unnecessary switching events and consequently lowers conduction and switching losses in the IGBTs. Similar energy efficiency improvements were observed in other RL-based converter studies, indicating that such policies can inherently learn low-loss operation patterns. Furthermore, the 35% faster dynamic response under irradiance transients demonstrates the system's capacity for rapid adaptation, a key advantage in fluctuating solar environments. Traditional MPC relies on predictive models that degrade with parameter drift, whereas DQN continuously refines control decisions based on state feedback, offering a robust alternative for uncertain PV dynamics.

Another important aspect is the observed thermal stability during the 72-hour reliability test. The maintained temperature within safe operating limits reflects not only efficient switching but also the potential of RL-based controllers to prolong inverter lifetime by preventing thermal stress. These outcomes collectively suggest that the proposed DQN framework effectively integrates intelligent learning with power electronics control theory, bridging the gap between AI optimization and embedded hardware constraints. Consequently, the research reinforces the feasibility of implementing RL in real-time low-power inverter systems and provides a foundation for extending this approach to grid-tied or multi-phase architectures.

## 4 Conclusion

This study successfully demonstrated the effectiveness of a reinforcement learning control strategy, particularly the Deep Q-Network (DQN) algorithm, for optimizing the operation of a low-power single-phase solar inverter. Through a two-stage methodology consisting of simulation-based training and hardware implementation, the DQN controller was shown to outperform traditional Proportional-Integral (PI) and Model-Predictive Control (MPC) methods across multiple performance metrics. The experimental results revealed that the proposed controller achieved superior waveform quality, with a Total Harmonic Distortion

(THD) of only 1.8%, indicating substantial improvement in output voltage fidelity. Moreover, the reinforcement learning approach yielded an efficiency increase of approximately 2.7% compared to the PI controller, primarily due to its adaptive switching strategy that minimized unnecessary transitions. In addition, the DQN controller demonstrated a faster dynamic response—about 35% quicker under sudden irradiance variations—while maintaining excellent stability and minimal overshoot. Long-term endurance testing over a 72-hour continuous operation further confirmed the controller’s reliability and thermal robustness. These findings highlight that reinforcement learning not only enhances power quality and energy efficiency but also ensures long-term operational stability, positioning it as a promising solution for intelligent power electronic systems. Future research may extend this work to higher-power, three-phase inverters and explore integration with grid-support functions such as reactive power control and adaptive energy management.

## References

- [1] International Renewable Energy Agency (IRENA), “Future of Solar Photovoltaic: Deployment, investment, technology, grid integration and socio-economic aspects,” International Renewable Energy Agency, Abu Dhabi, 2019. [Online]. Available: <https://www.irena.org/Publications/2019/Nov/Future-of-solar-photovoltaic>
- [2] F. Mahjabeen and A. Mohammad, “Revolutionizing Solar Energy: The Impact of Artificial Intelligence on Photovoltaic Systems,” *International Journal of Multidisciplinary Sciences and Arts*, vol. 2, no. 3, Art. no. 3, Jul. 2023, doi: 10.47709/ijmdsa.v2i1.2599.
- [3] F. Blaabjerg, Y. Yang, and K.-B. Kim, “Power Electronics for Renewable Energy Systems,” *Proceedings of the IEEE*, vol. 105, no. 11, pp. 2007–2011, 2017, doi: 10.1109/JPROC.2017.2759385.
- [4] R. Teodorescu, M. Liserre, and P. Rodriguez, *Grid Converters for Photovoltaic and Wind Power Systems*. John Wiley & Sons, 2011. doi: 10.1002/9780470667057.
- [5] S. B. Kjaer, J. K. Pedersen, and F. Blaabjerg, “A review of single-phase grid-connected inverters for photovoltaic modules,” *IEEE Transactions on Industry Applications*, vol. 41, no. 5, pp. 1292–1306, 2005, doi: 10.1109/TIA.2005.853371.
- [6] J. Rodriguez et al., “Predictive current control of a voltage source inverter,” *IEEE Transactions on Industrial Electronics*, vol. 54, no. 1, pp. 495–503, 2007, doi: 10.1109/TIE.2006.888802.
- [7] A. Rajamallaiah, S. Karri, M. Alghaythi, and M. Alshammari, “Deep Reinforcement Learning Based Control of a Grid Connected Inverter With LCL-Filter for Renewable Solar Applications,” *IEEE Access*, vol. PP, pp. 1–1, Feb. 2024, doi: 10.1109/ACCESS.2024.3364058.
- [8] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction*. MIT Press, 2018. [Online]. Available: <http://incompleteideas.net/book/the-book-2nd.html>
- [9] V. Mnih et al., “Human-level control through deep reinforcement learning,” *Nature*, vol. 518, no. 7540, pp. 529–533, 2015, doi: 10.1038/nature14236.
- [10] Y. Wang and Y. Peng, “Reinforcement Q-Learning Based Current Control for Grid-Connected Three-Phase Inverters with Unknown System Dynamics,” *Journal of Physics: Conference Series*, vol. 2224, p. 012110, Apr. 2022, doi: 10.1088/1742-6596/2224/1/012110.