

Digital Twin-Enabled Deep Reinforcement Learning for PV-ESS Energy Management in Active Distribution Networks

Minh Phong Le

Faculty of Electrical and Electronics, Thu Duc College of Technology, Ho Chi Minh City, Vietnam
minhphongle232@gmail.com

Corresponding author, minhphongle232@gmail.com

Abstract— This paper proposes a Digital Twin-enabled Deep Reinforcement Learning (DRL) framework for real-time energy management of photovoltaic (PV) and battery energy storage systems (ESS) in active distribution networks under uncertainty. A cyber-physical Digital Twin architecture is developed to continuously synchronize operating measurements with a virtual network model, providing validated system states for a Proximal Policy Optimization (PPO)-based DRL controller. The energy management problem simultaneously considers operational cost, voltage regulation, battery degradation, and uncertainty robustness. Monte Carlo-generated PV and load scenarios are employed to improve policy generalization under stochastic operating conditions. The proposed framework is validated on the IEEE 33-bus distribution system with 60% PV penetration and a 1.5 MWh ESS. Simulation results show that the proposed DRL+DT approach achieves a 22.2% reduction in operational cost, reduces voltage deviation, and improves robustness compared with RBC, OPF, and standalone DRL methods. The results demonstrate that Digital Twin-assisted state synchronization significantly enhances the effectiveness and reliability of DRL-based energy management in renewable-rich active distribution networks.

Keywords: Active distribution network, Deep reinforcement learning, Digital twin, Energy storage system, Uncertainty-aware energy management.



This work is licensed under a [CC BY-NC](https://creativecommons.org/licenses/by-nc/4.0/). Copyright ©2026 by Author. Published by Universitas Riau.

INTRODUCTION

The ongoing transition toward low-carbon energy systems has significantly reshaped the operational paradigm of modern power grids. In particular, the increasing penetration of renewable energy sources, especially photovoltaic (PV) systems, has transformed traditional passive distribution networks into active and highly dynamic systems. While this transition contributes to decarbonization goals, it simultaneously introduces critical operational challenges, including voltage fluctuations, reverse power flow, and increased uncertainty in generation and demand [1], [2], [3].

To mitigate these issues, energy storage systems (ESSs) have emerged as a key enabling technology for balancing supply and demand in renewable-rich distribution systems [4], [5], [6]. The coordinated operation of PV and ESS has therefore attracted considerable research interest in microgrids and active distribution networks [7], [8]. However, integrating these resources fundamentally changes network dynamics and requires intelligent energy

management strategies capable of operating reliably under uncertain real-time conditions.

Conventional approaches for energy management in distribution systems are primarily based on deterministic optimization techniques, such as optimal power flow (OPF) and mixed-integer programming [3], [9]. Although these methods can achieve satisfactory performance under predefined operating scenarios, their effectiveness is highly dependent on forecast accuracy and static assumptions. Consequently, their adaptability becomes limited in practical systems with stochastic PV generation and rapidly varying load demand.

To overcome these limitations, artificial intelligence (AI)-based approaches, particularly deep reinforcement learning (DRL), have recently gained significant attention in power system operation and control [10], [11]. DRL enables data-driven and model-free decision-making by continuously learning optimal control policies through interaction with the environment. Existing studies have demonstrated the capability of DRL for real-time energy management, voltage regulation, and ESS scheduling in renewable-integrated systems [1], [8], [11].

Another emerging paradigm in modern power systems is the Digital Twin (DT), which provides a continuously synchronized virtual representation of physical assets for real-time monitoring, simulation, and control. Recent studies have highlighted its potential to improve operational awareness and decision-making in renewable-integrated energy systems [1], [6], [12]. DT technology has been increasingly explored in energy systems to enhance situational awareness and improve decision-making accuracy. However, its integration with AI-based control frameworks for distribution network operation remains relatively underexplored.

Despite the growing body of literature, several important gaps can be identified. First, many existing studies adopt deterministic or simplified stochastic models, which do not adequately capture the inherent uncertainty of PV generation and load demand [3]. Second, battery degradation, an essential factor affecting the lifecycle and economic viability of ESS is often neglected in energy management formulations [5]. Third, most conventional optimization frameworks lack scalability and real-time adaptability for practical large-scale distribution systems [9], [10].

Reliable operation of active distribution networks increasingly depends on energy management strategies capable of responding to renewable uncertainty in real time while maintaining economic and technical performance. In this context, this paper develops a Digital Twin-enabled Deep Reinforcement Learning (DRL) framework for coordinated photovoltaic (PV) and battery energy storage system (ESS) operation. The proposed framework combines a cyber-physical Digital Twin environment with a Proximal Policy Optimization (PPO)-based control agent to achieve continuous state synchronization, adaptive ESS scheduling, and real-time decision-making under uncertain operating conditions. In addition to minimizing operational costs, the framework explicitly considers voltage regulation, battery degradation, and uncertainty robustness within a unified optimization framework. The effectiveness of the proposed approach is evaluated on the IEEE 33-bus active distribution network using Monte Carlo-generated PV and load scenarios and is compared against conventional Rule-Based Control (RBC), Optimal Power Flow (OPF), and standalone DRL approaches.

The main contributions of this work are summarized as follows. First, a Digital twin-enabled DRL framework is developed to provide continuous state synchronization, power-flow-validated system awareness, and real-time adaptive control for coordinated PV-ESS operation. Second, a stochastic scenario-based uncertainty modeling approach is incorporated using Monte Carlo-generated PV and load variations to enhance the robustness and generalization capability of the DRL agent. Third, battery degradation is explicitly integrated into the control objective through a degradation-aware ESS scheduling mechanism, enabling a balanced trade-off between short-term economic benefits and long-term battery lifetime preservation.

Fourth, a PPO-based DRL controller is designed to optimize ESS charging and discharging actions under dynamically changing operating conditions while maintaining voltage stability and operational reliability. Finally, the proposed framework is comprehensively validated using benchmark comparisons with RBC, OPF, and standalone DRL approaches, together with convergence analysis, statistical validation, and sensitivity assessment under different uncertainty levels.

METHOD

A. System Architecture and Problem Description

This study considers an active distribution network based on the IEEE 33-bus test system integrated with photovoltaic (PV) generation and battery energy storage systems (ESSs). The increasing penetration of renewable energy resources introduces significant operational challenges, including voltage fluctuations, reverse power flow, and uncertainty in power generation. To address these challenges, a Digital Twin-enabled Deep Reinforcement Learning (DRL) framework is proposed for real-time energy management.

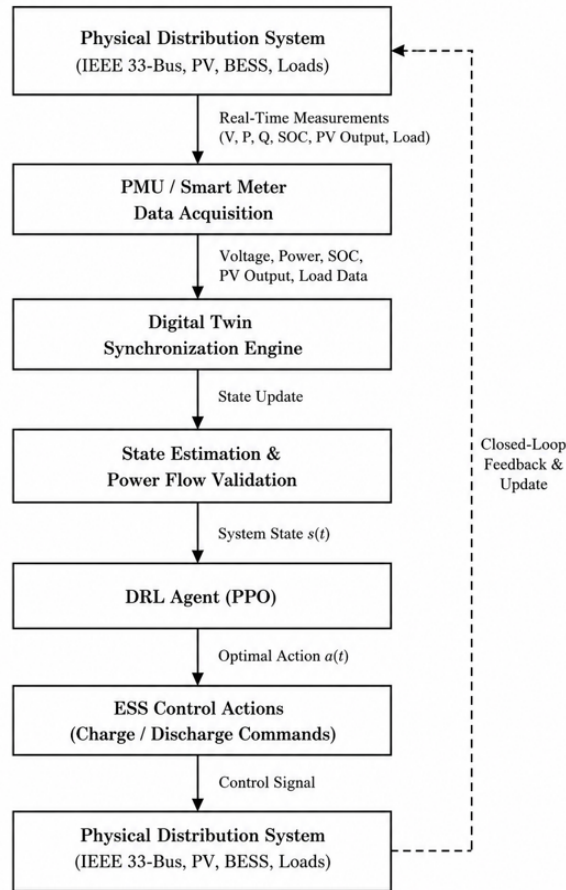


Figure 1. Digital Twin Architecture and Closed-Loop Synchronization Framework

The overall framework consists of three tightly coupled components: (i) the physical distribution network, (ii) the Digital Twin platform, and (iii) the DRL-based decision-making agent. Real-time measurements including load demand, PV generation, bus voltage magnitudes, and ESS operating conditions are continuously collected and synchronized with a virtual representation of the network. The Digital Twin provides a physically consistent operating environment through state estimation and AC power flow validation, while the DRL controller determines optimal ESS charging and discharging actions.

The proposed framework aims to simultaneously minimize operational cost, maintain voltage stability, and reduce battery degradation under uncertain operating conditions. Figure 1 illustrates the overall cyber–physical architecture adopted in this study.

B. Digital Twin-Based Cyber–Physical Modeling

The proposed Digital Twin is implemented as a cyber-physical platform that continuously synchronizes operating information between the physical distribution network and its virtual counterpart. As shown in Figure 1, the framework consists of a physical network layer, a data acquisition layer, a Digital Twin synchronization layer, and a PPO-based DRL control layer.

The physical system is represented by the IEEE 33-bus active distribution network integrated with PV generation and an ESS. Real-time measurements, including load demand, PV output, battery state-of-charge (SOC), and bus voltage magnitudes, are collected and transmitted to the Digital Twin environment. This closed-loop interaction enables continuous adaptation to system dynamics and uncertainty, which is not achievable using conventional deterministic optimization methods [6], [13]. The measurement vector is defined as:

$$s_t = [P_t^{Load}, P_t^{PV}, SOC_t, V_t] \quad (1)$$

where P_t^{Load} denotes the load demand, P_t^{PV} represents the photovoltaic generation output, SOC_t is the battery state-of-charge, and V_t denotes the network voltage profile at time t .

$$\hat{s}_t = f_{DT}(z_t; \theta_{DT}) \quad (2)$$

where z_t represents the measurement vector, $\hat{s}_t$ denotes the estimated virtual state, and θ_{DT} represents the Digital Twin model parameters.

The virtual model is developed in MATLAB/MATPOWER using AC power flow analysis of the IEEE 33-bus distribution network. Network topology, load demand, PV generation, and ESS operating constraints are continuously updated and validated to ensure physical consistency. The synchronization accuracy is evaluated by Equation (4). At each time step t , real-time measurements x_t (including load demand, PV output, and voltage profiles) are collected from the physical system. These measurements are processed by the Digital Twin, which reconstructs an estimated system state $\tilde{s}_t$ through a mapping function:

$$\tilde{s}_t = \mathcal{F}(x_t, z_t) \quad (3)$$

where z_t represents internal model parameters. Unlike purely data-driven approaches, the Digital Twin enforces physical consistency through embedded power flow constraints, ensuring that the estimated state remains feasible with respect to network operation [1], [7]. This step is critical, as inaccurate state representation can significantly degrade the performance and stability of DRL-based controllers.

The accuracy of the Digital Twin can be quantified by the estimation error:

$$e_t = \| s_t^{real} - \tilde{s}_t \| \quad (4)$$

Minimizing this error ensures that the DRL agent interacts with a high-fidelity representation of the system, thereby improving policy learning and generalization. This integration of physics-based modeling with learning algorithms has been identified as a key enabler for reliable AI deployment in power systems [14].

The validated operating states are subsequently provided to the PPO-based DRL controller for ESS scheduling. The resulting control actions are applied to the physical system, and updated measurements are fed back to the Digital Twin, forming a closed-loop cyber–physical control framework. This continuous interaction improves state awareness and enhances control performance under renewable generation uncertainty.

The interaction between the Digital Twin and the DRL agent forms a closed-loop control system:

$$x_t \rightarrow \tilde{s}_t \rightarrow a_t \rightarrow s_{t+1} \quad (5)$$

where x_t denotes the physical system state, s_t represents the Digital Twin-synchronized state provided to the DRL agent, a_t denotes the ESS control action, and s_{t+1} is the resulting system state at the next time step. Equation (5) summarizes the closed-loop interaction between the physical network, Digital Twin, and PPO controller. Although a laboratory-scale hardware platform is not available, the proposed Digital Twin emulates real-time cyber–physical interactions through synchronized measurement streams and power-flow-validated system states, providing a realistic environment for intelligent energy management studies. The proposed framework therefore represents a simulation-based Digital Twin environment designed for real-time energy management studies, where measurement synchronization, state validation, and closed-loop control interactions are continuously emulated within a cyber–physical setting.

C. Dataset Preparation and Uncertainty Modeling

The proposed framework is trained and evaluated using time-series operating data generated from the IEEE 33-bus active distribution network. The dataset contains measurements collected at 15-minute intervals over a 24-hour operating horizon, resulting in 96 samples per day. Each sample includes load demand, PV generation output, bus voltage magnitude, and ESS state-of-charge (SOC). Representative dataset samples are presented in Table 1.

Table 1. Representative Dataset Samples

Time	Load (MW)	PV Output (MW)	Voltage (p.u.)	ESS SOC (%)
00:00	2.15	0.00	0.984	80
06:00	2.48	0.32	0.978	72
12:00	3.12	1.85	1.012	65
18:00	3.45	0.42	0.971	58

To account for renewable intermittency, PV generation uncertainty is modeled as:

$$P_t^{PV} = \hat{P}_t^{PV} + \epsilon_t \quad (6)$$

where $\hat{P}_t^{PV}$ is the forecasted PV output and $\epsilon_t \sim \mathcal{N}(0, \sigma_{PV}^2)$ represents forecast uncertainty.

This probabilistic representation enables the proposed framework to evaluate system performance under realistic operating conditions where renewable generation uncertainty significantly impacts voltage stability and energy balance [4], [5].

To evaluate the robustness of the proposed framework under uncertain operating conditions, Monte Carlo simulation is employed to generate 100 stochastic scenarios. PV generation is assumed to vary within $\pm 20\%$ of forecast values, while load demand is varied within $\pm 15\%$ of nominal demand. The generated scenarios are used during both training and performance

evaluation to expose the DRL agent to diverse operating conditions and improve policy generalization under uncertainty.

D. Distribution Network and ESS Mathematical Model

The operation of a distribution network with integrated photovoltaic (PV) generation and energy storage systems (ESS) is inherently a multi-temporal and stochastic optimization problem. In this work, the energy management task is formulated as a sequential decision-making process, where control actions are determined based on evolving system states under uncertainty. This formulation aligns with recent advances in intelligent energy management using reinforcement learning and digital twin frameworks [9], [10].

The physical behavior of the distribution network is governed by nonlinear AC power flow equations. At each bus i , the active and reactive power balance can be expressed as:

$$P_i^{gen} - P_i^{load} = \sum_{j \in \mathcal{N}(i)} V_i V_j Y_{ij} \cos(\theta_i - \theta_j) \quad (7)$$

$$Q_i^{gen} - Q_i^{load} = \sum_{j \in \mathcal{N}(i)} V_i V_j Y_{ij} \sin(\theta_i - \theta_j) \quad (8)$$

where P_i^{gen} and Q_i^{gen} represent the net power injections from PV and ESS units, while P_i^{load} and Q_i^{load} denote the local demand. The voltage magnitude V_i and V_j phase angle θ_i , θ_j define the electrical state of the network, and Y_{ij} captures the admittance between buses. These equations reflect the fundamental coupling between nodal voltages and power injections, which becomes increasingly complex under high penetration of distributed energy resources [14].

To maintain secure operation, voltage magnitudes must satisfy operational limits:

$$V_i^{min} \leq V_i \leq V_i^{max}, \quad \forall i \quad (9)$$

In distribution systems with high PV penetration, voltage profiles are highly sensitive to local power imbalances. However, since ESS units mainly regulate active power rather than reactive power, the resulting voltage control mechanism becomes indirect and strongly nonlinear. This characteristic motivates the use of intelligent control approaches capable of learning complex system responses adaptively [8], [9].

The dynamics of the ESS are modeled through the state-of-charge (SOC), which ensures energy conservation throughout the scheduling horizon [5]. The SOC evolution is expressed as:

$$C_{t+1} = SOC_t + \frac{\eta_c P_t^{ch} - P_t^{dis} / \eta_d}{E_{max}} \cdot \Delta t \quad (10)$$

where: P_t^{ch} and P_t^{dis} denote charging and discharging power, η_c and η_d are charging and discharging efficiencies, E_{max} is the battery energy capacity, Δt is the simulation time interval.

This formulation captures both energy conservation and conversion losses during battery operation. Ignoring these physical characteristics may lead to unrealistic or infeasible operating strategies [6].

The ESS operational constraints are defined as:

$$0 \leq P_t^{ch} \leq P_{max}^{ch}; 0 \leq P_t^{dis} \leq P_{max}^{dis}; SOC^{min} \leq SOC_t \leq SOC^{max} \quad (11)$$

To improve long-term operational sustainability, battery degradation is explicitly incorporated into the optimization problem. A degradation-aware surrogate aging model is adopted as follows:

$$D_t^{ESS} = \alpha |P_t^{dis}| + \beta (SOC_t - SOC_{ref})^2 + k |\Delta SOC_t| \quad (12)$$

$$\text{With: } \Delta SOC_t = SOC_t - SOC_{t-1}$$

where α represents cycling-related degradation effects, β penalizes deviation from the optimal SOC operating range, and SOC_{ref} denotes the reference SOC value that minimizes battery stress. The additional term ΔSOC_t is introduced to capture the impact of frequent SOC fluctuations on battery aging. The weighting coefficients are selected as $\alpha = 1.0$, $\beta = 0.20$, and $k = 0.05$. These coefficients were selected through preliminary sensitivity analysis to ensure that degradation penalties influence policy learning without dominating the economic objective. The coefficient α represents degradation associated with battery energy throughput, β penalizes sustained operation away from the preferred SOC region, and k discourages excessive SOC fluctuations. Since k is significantly smaller than α and β , the third term mainly acts as a regularization component and does not substantially alter the overall optimization outcome.

The adopted degradation model is a surrogate aging representation that captures cycle-related and SOC-dependent degradation mechanisms. Compared with electrochemical and rainflow-counting-based aging models, the adopted formulation offers substantially lower computational complexity and is therefore more suitable for real-time DRL applications [15], [16], [17].

E. DRL Formulation and PPO Implementation

The ESS scheduling problem is formulated as a Markov Decision Process (MDP), which enables the DRL agent to learn an optimal control policy through continuous interaction with the Digital Twin environment. The MDP is defined as:

$$\mathcal{M} = (\mathcal{S}, \mathcal{A}, \mathcal{P}, \mathcal{R}) \quad (13)$$

where the state space $\mathcal{S}$ includes load demand, PV generation, voltage magnitudes, and ESS SOC, while the action space $\mathcal{A}$ consists of feasible ESS power outputs. The transition dynamics $\mathcal{P}$ capture the stochastic evolution of the system due to uncertainty in renewable generation and load demand.

The state space is designed to capture the most relevant operating conditions of the active distribution network and is defined as

$$s_t = [P_t^{Load}, P_t^{PV}, SOC_t, V_t] \quad (14)$$

where P_t^{Load} denotes the load demand, P_t^{PV} represents the photovoltaic generation output, SOC_t is the battery state-of-charge, and V_t denotes the network voltage profile at time t .

The action space corresponds to the ESS charging and discharging power command:

$$a_t = P_t^{BEES} \quad (15)$$

where P_t^{BEES} represents the charging/discharging power command of the ESS. Positive values correspond to charging actions, while negative values indicate discharging actions.

To simultaneously achieve economic operation, voltage regulation, and battery lifetime preservation, the reward function is formulated as:

$$R_t = -(C_{op} + w_1 V_{dev} + w_2 D_{ESS}) \quad (16)$$

where C_{op} denotes the operational cost, V_{dev} represents voltage deviation, D_{ESS} is the battery degradation penalty, and w_1 and w_2 are weighting coefficients. The reward formulation encourages the DRL agent to simultaneously minimize operating cost, maintain voltage stability, and reduce battery degradation, thereby achieving a balanced energy management strategy under uncertainty. The training and control procedure of the proposed Digital Twin-enabled PPO agent is illustrated in Figure 2.

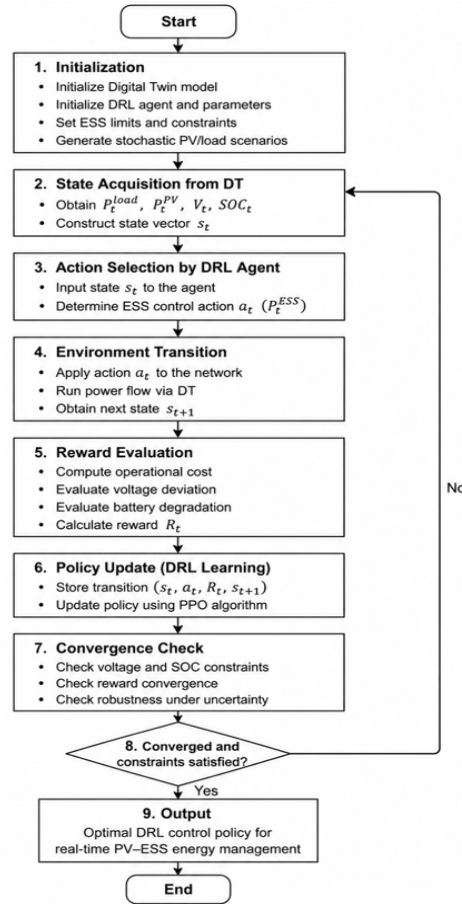


Figure 2. Flowchart of the Proposed Digital Twin-Enabled PPO Training and Control Procedure

To solve the MDP, a policy-based DRL algorithm is employed. In this work, the Proximal Policy Optimization (PPO) algorithm is adopted due to its stability and sample efficiency in continuous control problems [6], [7].

Neural Network Architecture

The PPO controller employs separate actor and critic neural networks with identical architectures. The input layer consists of four neurons corresponding to load demand, PV generation, ESS state-of-charge (SOC), and network voltage magnitude. Two fully connected hidden layers with 128 neurons each are utilized, with Rectified Linear Unit (ReLU) activation functions to improve nonlinear feature extraction and learning efficiency. The output layer contains a single neuron representing the ESS charging/discharging power command.

The neural network architecture can be summarized as follows:

- i) Input Layer: 4 neurons (Load, PV output, SOC, Voltage)
- ii) Hidden Layer 1: 128 neurons (ReLU)
- iii) Hidden Layer 2: 128 neurons (ReLU)
- iv) Output Layer: 1 neuron (ESS action)

Both actor and critic networks share the same architecture but maintain independent parameter sets during training. The actor network generates control actions, whereas the critic network estimates the state-value function to guide policy updates and improve learning stability.

The policy is updated by maximizing the clipped surrogate objective:

$$L^{CLIP}(\theta) = \mathbb{E}[\min(r_t(\theta)A_t, \text{clip}(r_t(\theta), 1 - \epsilon, 1 + \epsilon)A_t)] \quad (17)$$

where $r_t(\theta)$ is the probability ratio and A_t is the advantage function. This formulation prevents excessively large policy updates, ensuring stable convergence during training.

PPO Hyperparameter Configuration

The DRL controller is implemented using the Proximal Policy Optimization (PPO) algorithm. The policy and value networks each consist of two hidden layers with 128 neurons per layer and ReLU activation functions. The Adam optimizer is employed for parameter updates. All parameter settings are presented in Table 2.

Table 2. PPO Hyperparameter Settings

Parameter	Value
Learning rate	1×10^{-4}
Discount factor γ	0.99
Clipping factor ϵ	0.2
Batch size	64
Epochs	10
Hidden neurons	128
Training episodes	2000
Optimizer	Adam

Training Procedure

To enhance robustness under renewable uncertainty, the PPO agent is trained using 100 Monte Carlo-generated PV and load scenarios. Policy optimization is performed using the

Adam optimizer with a batch size of 64 and 10 training epochs per update. The training process runs for 2,000 episodes, and convergence is evaluated using the average cumulative reward. As shown in Figure 3, the reward progressively increases and eventually stabilizes, indicating that the agent has learned a stable and effective ESS scheduling policy.

The Digital Twin environment is implemented in MATLAB/MATPOWER, where AC power flow calculations are continuously executed to validate network feasibility. The DRL controller is developed using the MATLAB Reinforcement Learning Toolbox. Communication between the Digital Twin and the DRL agent is performed at each 15-minute control interval, enabling real-time state synchronization and adaptive decision-making. Although a laboratory-scale physical testbed is not available, the Digital Twin continuously synchronizes measurement streams and power-flow-validated states within the simulation environment, thereby emulating practical cyber–physical interactions.

F. Multi-Objective Optimization Framework

The proposed energy management framework is formulated as a multi-objective optimization problem that simultaneously considers economic efficiency, voltage regulation, battery lifetime preservation, and operational robustness under uncertainty. These objectives are integrated into a unified decision-making framework to ensure balanced system performance under varying operating conditions.

Let $t \in T$ denote discrete time intervals over a finite horizon. The system state at time t , denoted by s_t , includes load demand, PV generation, nodal voltages, and ESS states. The control action a_t corresponds to the charging or discharging power of the ESS units. The objective is to determine an optimal control policy π that minimizes the expected cumulative operational cost over time:

$$\min_{\pi} \mathbb{E} \left[\sum_{t=1}^T \gamma^t f(s_t, a_t) \right] \quad (18)$$

where $\gamma \in (0,1)$ is the discount factor reflecting the relative importance of future costs. This formulation enables the explicit consideration of long-term operational impacts, including battery degradation, which is often neglected in conventional optimization approaches.

The overall operational objective combines economic efficiency, voltage regulation, and battery health preservation into a unified multi-objective function [18]:

$$f(s_t, a_t) = w_1 C_t + w_2 \sum_i |V_i - V_{ref}| + w_3 D_t^{ESS} \quad (19)$$

where w_1 , w_2 , and w_3 are weighting coefficients representing the relative importance of each objective term. The operational cost component is defined as:

$$C_t = c_{grid} P_t^{grid} + c_{ESS} |P_t^{ESS}| \quad (20)$$

where c_{grid} denotes the electricity purchase price from the utility grid and c_{ESS} represents the ESS operating cost coefficient.

To evaluate robustness under uncertainty, a scenario-based Monte Carlo framework is employed. A set of stochastic scenarios k is generated to represent variations in PV output and load demand. The expected system performance is calculated as:

$$\bar{J} = \frac{1}{N} \sum_{k=1}^N J(\omega_k) \quad (21)$$

while the associated variance is given by:

$$\sigma^2 = \frac{1}{N} \sum_{k=1}^N (J(\omega_k) - \bar{J})^2 \quad (22)$$

where $J(\omega_k)$ denotes the operational performance under scenario k . The expected value reflects the average system performance, whereas the variance quantifies robustness against uncertainty.

G. Validation Framework

The proposed Digital Twin-enabled DRL framework is validated using the IEEE 33-bus active distribution network with a PV penetration level of 60% and a 1.5 MWh battery energy storage system. Simulations are performed in MATLAB/MATPOWER using a 24-hour scheduling horizon with a 15-minute time resolution. To account for renewable generation and load uncertainty, 100 Monte Carlo scenarios are generated and used throughout the evaluation process.

To demonstrate the effectiveness of the proposed approach, three benchmark methods are considered for comparison:

- i) Rule-Based Control (RBC): a conventional fixed-rule energy management strategy.
- ii) Optimal Power Flow (OPF): a deterministic optimization-based scheduling approach.
- iii) Deep Reinforcement Learning (DRL): a PPO-based controller without Digital Twin integration.

The performance of each method is evaluated using multiple technical and economic indicators, including operational cost, voltage deviation, voltage violation rate, ESS cycling frequency, SOC variance, battery degradation, and cost variance. These metrics provide a comprehensive assessment of economic efficiency, voltage stability, battery lifetime preservation, and operational robustness. In addition to deterministic performance evaluation, statistical validation is conducted using the mean, standard deviation, and best- and worst-case operating costs obtained from Monte Carlo simulations. Furthermore, a sensitivity analysis is performed by gradually increasing PV uncertainty levels from 10% to 40% to examine the generalization capability of the proposed framework under increasingly uncertain operating conditions. The adopted validation framework enables a comprehensive evaluation of both control performance and robustness, ensuring that the proposed Digital Twin-enabled DRL strategy is assessed under realistic and uncertain operating environments.

RESULT AND DISCUSSION

A. Simulation Setup and Evaluation Framework

To validate the effectiveness of the proposed Digital Twin-assisted DRL framework, simulations are conducted on the IEEE 33-bus distribution system [19]. The network operates at a nominal voltage level of 12.66 kV with a total load demand of 3.72 MW and 2.3 MVar. The test system is selected due to its widespread adoption in active distribution network research and its suitability for evaluating voltage-sensitive operating conditions under high renewable penetration. The Digital Twin environment is implemented using MATLAB/MATPOWER for

real-time power flow validation, whereas the DRL controller is developed using the PPO algorithm in the MATLAB Reinforcement Learning Toolbox. All benchmark methods are evaluated under identical operating conditions to ensure fair comparison. The virtual model is developed based on AC power flow analysis of the IEEE 33-bus distribution network. Network topology, line parameters, load profiles, PV generation outputs, and ESS operating constraints are continuously updated and validated within the Digital Twin environment to ensure physical consistency. The key simulation parameters, including PV penetration level, BESS capacity and power rating, time resolution, and voltage limits, are summarized in Table 3.

Table 3. Simulation Parameters

Parameter	Value
PV penetration	60% of peak load
BESS capacity	1.5 MWh
BESS power rating	± 0.5 MW
Time resolution	15 min
Voltage limits	0.95–1.05 p.u.

B. PPO Training Convergence Analysis

To evaluate the learning effectiveness and stability of the proposed DRL controller, the training convergence behavior of the PPO algorithm is analyzed over 2,000 training episodes. Figure 3 illustrates the evolution of the average episode reward during the learning process. The reward value gradually increases as the agent gains experience through continuous interaction with the Digital Twin environment and learns more effective ESS scheduling policies. The convergence trend provides important evidence of training stability and demonstrates the capability of the proposed DRL framework to identify near-optimal control strategies under uncertain operating conditions.

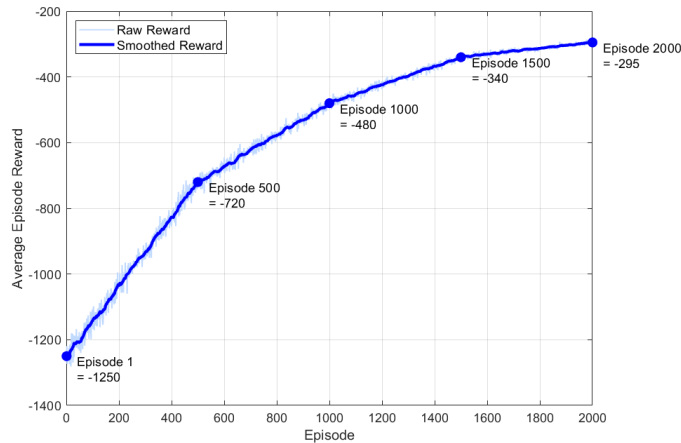


Figure 3. PPO Training Convergence Curve

As shown in Figure 3, the average episode reward improves significantly from approximately $-1,250$ at the beginning of training to about -295 after 2,000 episodes. A rapid performance improvement is observed during the early training stage, indicating efficient exploration and policy learning. The reward curve gradually stabilizes after approximately 1,800 episodes, confirming the convergence of the PPO algorithm. These results demonstrate that the proposed DRL agent successfully learns a stable and effective energy management policy for real-time PV–BESS coordination under renewable uncertainty.

C. Voltage Regulation and Operational Cost Performance

The effectiveness of the proposed DRL + DT framework is validated through comprehensive simulations on the IEEE 33-bus distribution system. The results clearly demonstrate that integrating Digital Twin technology with Deep Reinforcement Learning substantially enhances both operational efficiency and system reliability under uncertain operating conditions.

Table 4. Voltage Regulation and Operational Cost Performance

Method	Cost (USD/day)	Cost Reduction (%)	Maximum Voltage Deviation (p.u.)	Average Voltage Deviation (p.u.)
RBC [8]	1,245	–	0.081	0.034
OPF [9]	1,082	13.1%	0.056	0.021
DRL [2]	1,035	16.9%	0.049	0.018
DRL + DT	968	22.2%	0.039	0.015

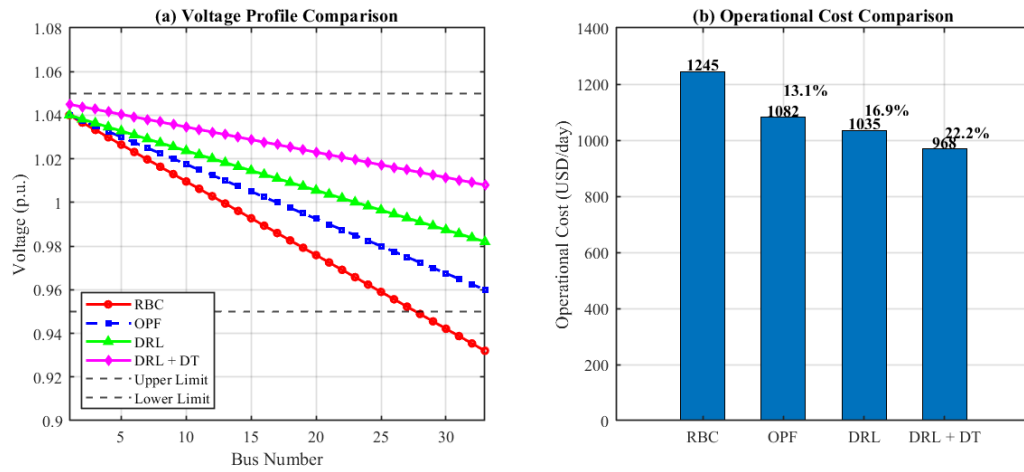


Figure 4. Performance Comparison of Different Energy Management Methods: (a) Voltage Profile Comparison, (b) Operational Cost Comparison

As presented in Table 4 and Figure 4, the voltage regulation and operational cost performance of all evaluated energy management strategies are compared. Among all methods, the proposed DRL + DT framework achieves the best overall performance, reducing the operational cost to 968 USD/day, which corresponds to a 22.2% reduction compared with RBC. In addition, it maintains the lowest maximum and average voltage deviations of 0.039 p.u. and 0.015 p.u., respectively. While OPF and standalone DRL improve system performance relative to RBC, their effectiveness is limited by deterministic scheduling and state estimation uncertainty. By combining adaptive DRL decision-making with Digital Twin-assisted state synchronization, the proposed framework provides more accurate ESS control actions, resulting in improved economic efficiency, voltage regulation, and operational reliability under uncertain operating conditions.

D. ESS Operation and Robustness Under Uncertainty

To further assess operational robustness, ESS behavior and system performance are evaluated under stochastic PV and load scenarios. The results presented in Table 5 represent the average operational performance obtained under 100 Monte Carlo uncertainty scenarios.

Table 5. ESS Operational Characteristics and Robustness under Uncertainty

Method	Avg. Cycles/Day	SOC Variance	Mean Cost (USD)	Std Dev (USD)	Voltage Violations (%)
RBC [8]	2.8	0.19	1,310	182	12.4
OPF [9]	2.1	0.14	1,150	135	6.7
DRL [2]	1.9	0.11	1,085	118	4.1
DRL + DT	1.6	0.09	1,020	88	2.3

As presented in Table 5 and Figure 5, the proposed DRL + DT framework achieves the best overall robustness among all evaluated methods. The proposed approach reduces ESS cycling frequency, SOC variance, operating cost variability, and voltage violations while maintaining the lowest average operating cost. These improvements are attributed to the combination of uncertainty-aware DRL training and Digital Twin-assisted state synchronization, which enables more reliable ESS scheduling under stochastic PV and load conditions. As a result, the proposed framework effectively balances operational efficiency, voltage stability, and battery lifetime preservation.

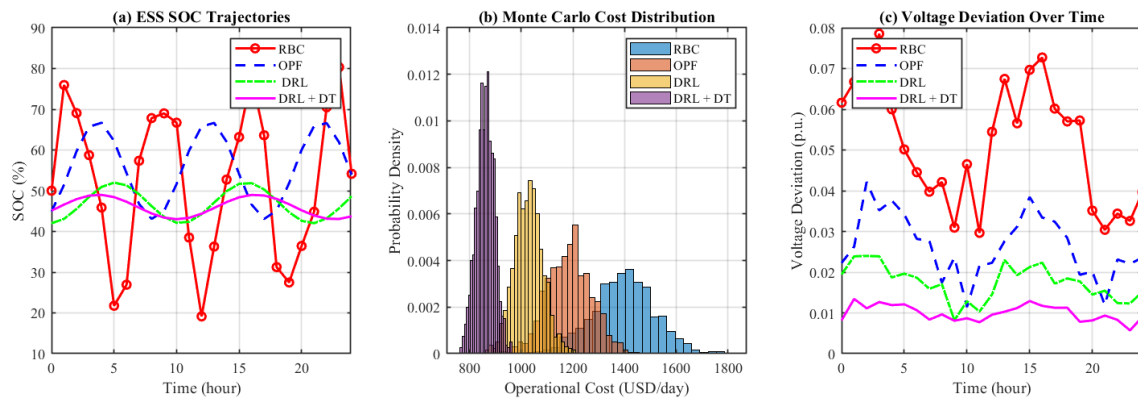


Figure 5. ESS Performance and System Robustness Analysis: (a) ESS SOC Trajectories, (b) Monte Carlo Cost Distribution, (c) Voltage Deviation Over Time

E. Statistical Validation

To further investigate the robustness and consistency of the proposed framework under uncertainty, Monte Carlo simulations with 100 stochastic operating scenarios are conducted. Table 6 presents the statistical characteristics of the operating cost obtained from repeated simulation runs under the nominal operating scenario.

Table 6. Statistical Validation of Operational Cost Under Stochastic Conditions

Method	Mean (USD/day)	Std (USD/day)	Best (USD/day)	Worst (USD/day)
RBC [8]	1,245	76	1,162	1,368
OPF [9]	1,082	52	1,018	1,163
DRL [2]	1,035	38	984	1,107
DRL + DT	968	21	935	1,002

As shown in Table 6, the proposed DRL + DT framework achieves the lowest mean operational cost and exhibits the smallest standard deviation among all compared methods. The standard deviation of the proposed framework is significantly lower than those of RBC, OPF, and standalone DRL, indicating superior robustness and reduced sensitivity to renewable intermittency and load uncertainty. Moreover, the proposed framework exhibits the narrowest gap between the best- and worst-case operating scenarios, demonstrating highly consistent

performance across a wide range of uncertainty realizations. These results provide quantitative evidence that the observed performance improvements are not limited to a single operating condition but remain reliable under stochastic environments.

F. Sensitivity Analysis

To evaluate the generalization capability of the proposed framework under different levels of renewable uncertainty, a sensitivity analysis is conducted by gradually increasing the PV forecast uncertainty from 10% to 40%. The corresponding operational costs are summarized in Table 7 and illustrated in Figure 6.

Table 7. Impact of PV Uncertainty on Operational Cost

PV Uncertainty (%)	Operational Cost (USD/day)
10	942
20	968
30	1014
40	1087

As shown in Figure 6 and Table 7, the operational cost increases with higher PV uncertainty due to the growing difficulty of balancing renewable variability and load demand. Nevertheless, the increase remains moderate across all tested scenarios. Even at the severe uncertainty level of 40%, the proposed DRL + DT framework maintains acceptable voltage profiles and stable operating costs. This behavior demonstrates the strong generalization capability of the learned control policy and confirms the effectiveness of the Digital Twin-assisted state synchronization mechanism under highly uncertain operating conditions.

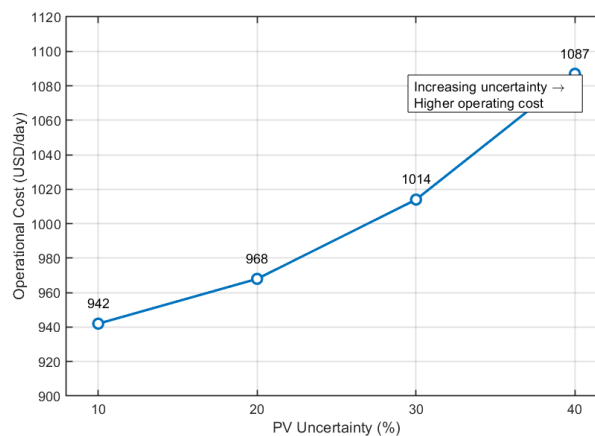


Figure 6. Sensitivity Analysis Under Different PV Uncertainty Levels

G. Comparison with Existing Approaches

To position the proposed framework within the current state of the art, a quantitative comparison with representative existing methods is provided in Table 8.

The reported performance ranges for existing approaches are summarized from the corresponding references and are provided for qualitative benchmarking purposes. As shown in Table 8, the comparative analysis reveals that most existing approaches only partially address practical operational requirements. Conventional DRL methods generally lack degradation awareness, while stochastic OPF methods suffer from limited real-time capability. In contrast, the proposed DRL + DT framework simultaneously integrates uncertainty handling,

degradation-aware optimization, real-time adaptive control, and accurate state estimation within a unified architecture. The results confirm that the proposed framework provides a robust, intelligent, and practically deployable solution for modern renewable-integrated distribution microgrids, offering significant advantages in operational efficiency, voltage stability, and long-term system sustainability.

Table 8. Quantitative comparison with existing approaches

Study	Uncertainty Handling	Battery Degradation	Real-Time Capability	Cost Reduction
DRL-based EMS [11]	Partial	No	Yes	12–15%
Stochastic OPF [9]	Yes	No	No	15–18%
Heuristic Control [8]	No	No	Yes	Limited
DRL + DT	Yes	Yes	Yes	22.2%

CONCLUSION

This paper proposes a Digital Twin-enabled Deep Reinforcement Learning framework for real-time PV–ESS energy management in active distribution networks under renewable uncertainty. By integrating a cyber-physical Digital Twin with a PPO-based controller, the proposed framework achieves continuous state synchronization, adaptive ESS scheduling, and uncertainty-aware decision-making. Simulation results on the IEEE 33-bus system demonstrate superior performance over conventional RBC, OPF, and standalone DRL approaches, achieving a 22.2% reduction in operational cost while improving voltage regulation, reducing ESS cycling frequency, and enhancing battery lifetime preservation. Statistical and sensitivity analyses further verify the robustness and generalization capability of the proposed framework under stochastic PV and load conditions. These findings indicate that Digital Twin-assisted DRL provides an effective and computationally efficient solution for reliable energy management in renewable-rich active distribution networks. Future work will extend the proposed framework to larger-scale distribution systems incorporating electric vehicles, demand response, and multiple distributed energy resources.

REFERENCES

- [1] Z. Liu, Y. Zhang, and H. Li, “Digital twin for power systems: State-of-the-art and future trends,” *Renew. Sustain. Energy Rev.*, vol. 152, p. 111675, 2021, doi: 10.1016/j.rser.2021.111675.
- [2] K. Zhou, S. Yang, and Z. Shao, “Digital twin-based optimization for energy systems,” *Appl. Energy*, vol. 285, p. 116419, 2021, doi: 10.1016/j.apenergy.2020.116419.
- [3] U. I. Musa and S. Ghosh, “Advancing Digital Twin through the Integration of new AI Algorithms,” *Int. Res. J. Eng. Technol.*, pp. 922–932, 2023, [Online]. Available: <https://www.researchgate.net/publication/369693333>
- [4] N. B. Mohamed, G. Kaddoum, and M. Z. Hassan, “Digital Twin-Driven Continual Deep Reinforcement Learning for Coexistence of Multiple Radio Access Technology IoT Links with Nonlinear Receivers,” *IEEE Trans. Mach. Learn. Commun. Netw.*, vol. 4, pp. 591–611, 2026, doi: 10.1109/TMLCN.2026.3669837.
- [5] M. R. Kabir, D. Halder, and S. Ray, “Digital Twins for IoT-Driven Energy Systems: A Survey,” *IEEE Access*, vol. 12, pp. 177123–177143, 2024, doi: 10.1109/ACCESS.2024.3506660.
- [6] C. Nneoma Ugwu, F. Chukwudi Ogenyi, J. Nnenna Ugwu, P. C. Okechukwu Ugwu, and

- M. Ben Okon, "Integration of AI-driven digital twins for real-time optimization of renewable energy grids," *Front. Energy Res*, vol. 14, pp. 1–13, 2026, doi: 10.3389/fenrg.2026.1748233.
- [7] N. Kumari, A. Sharma, B. Tran, N. Chilamkurti, and D. Alahakoon, "A comprehensive review of digital twin technology for grid-connected microgrid systems," *Energies*, vol. 16, no. 14, p. 5525, 2023, doi: 10.3390/en16145525.
- [8] J. M. Guerrero, J. C. Vasquez, J. Matas, L. G. Vicuña, and M. Castilla, "Hierarchical control of droop-controlled AC and DC microgrids—A general approach toward standardization," *IEEE Trans. Ind. Electron*, vol. 58, no. 1, pp. 158–172, 2011, doi: 10.1109/TIE.2010.2066534.
- [9] A. Parisio, E. Rikos, and L. Glielmo, "A model predictive control approach to microgrid operation optimization," *IEEE Trans. Control Syst. Technol*, vol. 22, no. 5, pp. 1813–1827, 2014, doi: 10.1109/TCST.2013.2295737.
- [10] M. Glavic, R. Moreno, T. Bi, and M. Shahidehpour, "Reinforcement learning for electric power system decision and control: Past considerations and perspectives," *IEEE Trans. Smart Grid*, vol. 12, no. 6, pp. 5128–5142, 2021, doi: 10.1109/TSG.2021.3070934.
- [11] Z. Zhang, J. Wang, and B. Chen, "Deep reinforcement learning for energy management in microgrids," *Appl. Energy*, vol. 288, p. 116618, 2021, doi: 10.1016/j.apenergy.2021.116618.
- [12] Y. Wang, X. Liu, J. Zhao, and M. Shahidehpour, "Digital Twin-Driven Resilient Operation of Active Distribution Networks With High Renewable Penetration," *IEEE Trans. Smart Grid*, vol. 15, no. 2, pp. 1648–1661, 2024, doi: 10.1109/TSG.2023.3334869.
- [13] A. A. Ardebili, M. Zappatore, A. I. H. A. Ramadan, A. Longo, and A. Ficarella, "Digital Twins of smart energy systems: a systematic literature review on enablers, design, management and computational challenges," *Energy Informatics*, vol. 7, no. 1, Art. no. 94, 2024, doi: 10.1186/s42162-024-00385-5.
- [14] W. F. Mbasso, A. Harrison, I. Dagal, P. Jangir, M. Khishe, H. Kotb, M. S. Shaikh, A. Smerat, E. F. Donfack, "Digital twins in renewable energy systems: A comprehensive review of concepts, applications, and future directions," *Int. J. Electr. Power Energy Syst*, vol. 61, Art. no. 101814, 2025, doi: 10.1016/j.esr.2025.101814.
- [15] B. Xu, A. Oudalov, A. Ulbig, G. Andersson, and D. S. Kirschen, "Modeling of Lithium-Ion Battery Degradation for Cell Life Assessment," *IEEE Trans. Smart Grid*, vol. 9, no. 2, pp. 1131–1140, 2018, doi: 10.1109/TSG.2016.2578950.
- [16] B. Xu, Y. Shi, D. S. Kirschen, and B. Zhang, "Optimal Battery Participation in Frequency Regulation Markets Considering Degradation Cost," *IEEE Trans. Power Syst*, vol. 33, no. 5, pp. 4908–4918, 2018, doi: 10.1109/TPWRS.2018.2799437.
- [17] J. Chen, X. Wang, M. Zhou, and Y. Fu, "Battery Degradation-Aware Deep Reinforcement Learning for Energy Storage Scheduling in Distribution Networks," *IEEE Trans. Sustain. Energy*, vol. 16, no. 1, pp. 215–227, 2025, doi: 10.1109/TSTE.2024.3478912.
- [18] T. N. Ton, P. M. Le, and T. M. Le, "Multi-objective optimization of distributed generation placement and sizing in active distribution networks considering harmonic distortion," *Int. J. Electr. Comput. Eng*, vol. 16, no. 2, pp. 598–607, 2026, doi: 10.11591/ijece.v16i2.pp598-607.
- [19] T.N.Ton, P.M. Le, L.H.Pham, and T. M. Le, "Multi-Objective Optimization of Electric Distribution Systems with integrated distributed Generation using Deep Reinforcement Learning," *Eng. Technol. Appl. Sci. Res*, vol. 15, no. 2, pp. 22166–22171, 2025.

BIOGRAPHIES OF AUTHORS



Phong Minh Le was born in Tien Giang, Vietnam, in 1982. He received his B.Eng. (2005) and M.Eng. (2008) degrees in Electrical Engineering from Ho Chi Minh City University of Technology and Education, Vietnam and completed his Ph.D. degree in 2017. He is currently a lecturer at the Department of Electrical and Electronics Engineering, Thu Duc College of Technology, Ho Chi Minh City, Vietnam. His teaching and research interests focus on power systems, electrical machines, and the application of optimization techniques in electrical engineering.