International Journal of Innovative Research in Engineering and Management (IJIREM)
ISSN (Online): 2350-0557, Volume 13, Issue 3, June 2026
DOI: https://doi.org/10.55524/ijirem.2026.13.3.13
Article ID IJIRE-1461, Pages 92-118
www.ijirem.org

Intelligent Multi-Agent Deep Reinforcement Learning for Autonomous Microgrid Management: Enhancing Energy Resilience through Solar-Wind-Battery Integration in Urban Networks

Adel Elgammal

Professor, Utilities and Sustainable Engineering, The University of Trinidad & Tobago UTT, Trinidad and Tobago

Correspondence should be addressed to Adel Elgammal ; adel_elgammal2000@yahoo.com

Received: 21 May 2026; Revised: 6 June 2026; Accepted: 20 June 2026

Copyright © 2024 Made Adel Elgammal. This is an open-access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

ABSTRACT- Current urban power systems are facing dual challenges amid the process of integrating renewable energy: continuous growth in energy demand, and compromised grid stability and energy security under the impact of climate disasters. Traditional centralized control systems cannot adapt to the complex dynamic interactions of distributed photovoltaics, wind turbines, and energy storage units within urban microgrids, so energy resilience vulnerabilities are easily exposed during peak load periods and extreme weather scenarios. To address this problem, this study proposes a new multi-agent deep reinforcement learning (MADRL) framework that solves the core flaws of traditional systems through decentralized decision-making and adaptive control mechanisms. This framework uses a hybrid Actor-Critic architecture integrated with graph neural networks, and relies on a multi-objective reward mechanism to simultaneously optimize three core indicators: energy cost, carbon emissions, and grid stability. This study conducts simulation verification on a typical urban block microgrid configured with 500kW photovoltaics, 150kW wind power, and 400kWh battery energy storage. Taking traditional rule-based control systems as the comparison baseline, Simulation results show that the proposed framework improves energy cost efficiency by 23%, reduces carbon emissions by 18%, and increases grid disturbance response speed by 35%. During peak production periods, the renewable energy utilization rate reaches 94%; when sudden load changes occur, grid frequency remains stable within ±0.2Hz. Furthermore, the framework can be deployed in a scalable way to adapt to the weather patterns, electricity consumption characteristics, and grid infrastructure constraints of different cities. This study has certain limitations: it requires high computing power to implement in real time, and it relies on large amounts of training data. This framework can be applied to urban energy planning, smart city construction, and climate change mitigation strategies, and can deliver social values including improved energy security, reduced user electricity bills, and enhanced environmental sustainability of urban communities.

KEYWORDS- Microgrid Control, Renewable Energy Integration, Smart Grid Optimization, Energy Resilience, Urban Decarbonization, Machine Learning for Energy Management.

I. INTRODUCTION

The global sustainable energy transition is driving the large-scale deployment of renewable energy at the city level, and traditional centralized power systems are gradually evolving into complex distributed systems with the capabilities to support bidirectional power flow, integration of intermittent power sources, and intelligent regulation. Previous studies[1] [2] have noted that the three core driving forces underpinning this transition are: the urbanization trend that will bring 68% of the global population into urban areas by 2050, the globally tightening environmental regulations for net-zero carbon emissions, and technological breakthroughs in renewable energy, energy storage, and power electronics. Against this backdrop, microgrids have become a core supporting carrier for the energy transition, and literature [4] clearly defines their core attributes: unlike traditional centralized power generation systems, microgrids can integrate a range of distributed energy resources (DER) including photovoltaics, wind turbines, battery energy storage systems (BESS), and controllable loads within a defined geographic area, to operate either in standalone mode or grid-connected mode. Statistical data from the International Energy Agency (IEA) [5] shows that in 2023, the global installed capacity of microgrids reached 35GW, with an annual growth rate of 15%, and urban renewable energy projects are the core growth driver. Existing studies have confirmed that integrating renewable energy into urban microgrids can deliver three core values: improving regional energy independence, reducing greenhouse gas emissions, and enhancing power system flexibility [6]. The two core conditions required to implement this model have reached maturity: over the past decade, global photovoltaic costs have dropped by 85% and wind power costs by 70%, giving renewable energy full economic competitiveness [7]; the iterative development of lithium-ion battery storage technology has effectively resolved the intermittency issue of renewable energy output [8]. However, existing studies also point out that the current integration of renewable energy into urban microgrids still faces core management and control challenges. Traditional control methods designed for schedulable, predictable power sources are no longer suitable for this new scenario. The new management and control system must simultaneously meet multiple requirements: coordinating multi-energy assets, maintaining power quality, ensuring grid stability, and synchronously optimizing both economic and environmental targets [9][10][11]. Recent advances in the fields of artificial intelligence and machine learning have given rise to intelligent control systems that can learn optimal strategies from data and adapt to dynamic operating conditions [12]; as a subfield of machine learning, deep reinforcement learning (DRL) integrates neural networks and reinforcement learning principles to enable autonomous control that adapts to complex dynamic environments [13], while multi-agent systems support distributed decision-making and adapt to the decentralized operation of microgrids [14]. All of these technologies closely align with the core industry challenges that need to be resolved.

Although renewable energy power generation technology and its supporting control systems have made many notable advances in recent years, the autonomous management of urban microgrids that integrate renewable energy remains a severe challenge spanning multiple dimensions: technology, economy, and operation [15]. The core dilemma is the urgent need to develop control strategies that can coordinate various heterogeneous energy assets, while optimizing multiple often conflicting objectives in a dynamic operating environment with high uncertainty [16]. First, the intermittency and randomness of renewable energy power generation itself is a fundamental challenge: the output of solar photovoltaic systems fluctuates sharply under the influence of cloud cover, atmospheric conditions, and day-night cycles; wind power output exhibits high variability across multiple timescales ranging from seconds to seasonal levels; the spatial distribution of urban renewable energy assets further amplifies power generation fluctuations due to local weather differences. Second, traditional control systems that rely on pre-set rules and historical data cannot adapt to these rapid, unpredictable output changes [17]-[25]. Third, the integration of energy storage adds extra complexity: battery systems require cross-timescale optimization to account for degradation characteristics, state of charge constraints, and cycle life limits; coordinating wind, solar, and storage resources also requires advanced forecasting and real-time optimization algorithms; the bidirectional attribute of energy storage additionally introduces new variables that traditional methods struggle to manage [26]. Furthermore, the economic, environmental, and technical operational objectives that microgrids must meet simultaneously often conflict with each other. For example, maximizing the utilization rate of renewable energy may harm grid stability. Finally, insufficient system scalability and adaptability is another key dimension that blocks the practical implementation of autonomous management, which underscores the necessity of conducting this research [27]. Urban microgrids exhibit significant heterogeneity in scale, asset composition, load characteristics, and grid interconnection requirements [28], and they evolve dynamically alongside the connection of new assets, equipment upgrades, and shifts in load patterns [29]. Therefore, their control systems must be able to adapt to all types of dynamic changes without needing extensive redesign or recalibration [30].

This paper first sorts out the root causes of the long-unresolved core challenges in the microgrid control field. It concludes that the persistence of these challenges stems from core flaws in existing technical pathways and the inherent complexity of the problem domain itself. The paper then breaks down five categories of core barriers that restrict the practical resolution of these challenges one by one, laying out all blockages across the full chain of the current microgrid control field, from technical solutions to basic supporting conditions. All core arguments are supported by corresponding academic citations, and the logical thread of the discussion stays tightly aligned with the central theme throughout. As the foundation of the vast majority of currently deployed microgrid control systems, rule-based control (traditional rule-based control) has a core flaw of insufficient flexibility: it cannot adapt to complex dynamic operating scenarios, and its rule set suffers from poor scalability. Relevant conclusions are presented in [31]–[33]. Model Predictive Control (MPC), which has a higher maturity level than rule-based systems, has critical drawbacks of insufficient model accuracy and excessive requirements for real-time computing power, which prevents it from being adapted and scaled up for large-scale microgrid deployment. Supporting literature for this judgment is cited in [34]–[36]. Existing machine learning methods, which are dominated by supervised learning, can only be applied to niche scenarios such as load forecasting and renewable energy output forecasting. They cannot solve the core problem of autonomous decision-making in uncertain environments, and they also rely heavily on large-scale labeled datasets, meaning they cannot iterate and upgrade from actual operational experience. Relevant supporting arguments are presented in [37]–[40]. Multi-agent coordination schemes lack a unified standardized framework; most existing studies assume perfect communication conditions, which is severely disconnected from the real operating scenarios of actual microgrids, where communication delays, packet loss, and frequent faults are common. This disconnect means these schemes cannot support the deployment of scalable control solutions, with relevant analyses provided in [41]–[43]. Finally, real-world datasets and testbed resources for verifying advanced control algorithms are extremely scarce, which directly hinders technological progress across the entire field. This judgment is noted in [44]. All citations cover the academic literature numbered [31]–[44], and the full discussion contains no redundant content that deviates from the central theme. Most current research on microgrids relies on simplified simulation models. Reference [45] points out that these models fail to account for equipment aging, measurement uncertainty, cybersecurity risks, and interactions with operation and maintenance personnel. Reference [46] shows that this limitation leads to overly conservative implementation schemes, which prevent the full potential of advanced control technologies from being realized.

This study proposes an original intelligent multi-agent deep reinforcement learning (MADRL) framework, developed specifically to address the core challenge of autonomous management for urban microgrids integrated with renewable energy. This framework combines the autonomous decision-making capacity of deep reinforcement learning in uncertain environments, alongside the distributed coordination advantages of multi-agent architectures for heterogeneous energy assets. Compared with the centralized control schemes used in traditional microgrids, this framework achieves a paradigm shift toward a decentralized intelligent learning system, with an underlying core that is a hybrid Actor-Critic architecture integrated with graph neural networks (GNN). Unlike the inherent flaw of traditional multi-agent systems that require pre-defined coordination protocols, the proposed framework can independently learn communication and coordination strategies without building an explicit system model, and can flexibly adapt to configuration differences and dynamic operating conditions of all types of microgrids. Its core technical innovations include four areas: First, the integration of GNN enables efficient processing and sharing of microgrid topology and operating status information, balancing the quality of global coordination and decision-making with distributed computing efficiency; second, a hierarchical learning structure decomposes complex control problems, allowing each agent to focus on its exclusive control objectives while maintaining overall global system consistency; third, an attention mechanism dynamically prioritizes different types of information, improving the quality of decision-making across multiple scenarios; fourth, a multi-objective reward structure with adaptive weights balances mutually exclusive operating objectives, overcoming the limitation of traditional optimization methods that rely on fixed objective functions, and adapting to the needs of changing priorities and emergency scenarios. The framework also embeds a constraint-aware learning algorithm, which incorporates operating safety limits and regulatory requirements as built-in learning constraints to ensure all control actions comply with relevant rules. Overall, this study fills key gaps in existing literature across multiple dimensions, establishing the core theoretical foundation of this research. This study centers on the multi-agent reinforcement learning framework developed for power scenarios of urban microgrids, and outlines its comprehensive research contributions across three core dimensions: First, at the theoretical level, the study first advances development in the multi-agent reinforcement learning field by creating a new coordination mechanism specifically tailored for power systems [59]; second, it proposes a novel integrated architecture combining graph neural networks and multi-agent deep reinforcement learning, which can be applied to other distributed control problems beyond the microgrid context [60]; next, at the practical application level, this framework features scalability and adaptability, allowing it to be deployed for all types of urban microgrid configurations without extensive customized modifications [61]. It can significantly improve energy system resilience, enabling microgrids to maintain stable operation relying on their autonomous recovery capabilities across three types of scenarios: extreme weather, equipment failures, and cybersecurity incidents [62]. It can simultaneously optimize multiple objectives and learn from operational experience, making it a transformative technology for addressing the complexity of urban energy systems [63]. Its modular design can interface with existing control systems, lowering implementation barriers for operators and utility companies [64]; finally, at the macro cross-domain level, it provides technical support for sustainable urban development to help reach carbon emission reduction targets [65]. It can maximize the utilization rate of renewable energy while maintaining grid stability, supporting urban decarbonization that aligns with global emission reduction strategies [66], and it can also cut energy costs to enhance the investment appeal of renewable energy projects [67].

II. THE PROPOSED INTELLIGENT MULTI-AGENT DEEP REINFORCEMENT LEARNING FOR AUTONOMOUS MICROGRID MANAGEMENT

Figure 1 of this paper presents the intelligent multi-agent deep reinforcement learning (MADRL) framework for autonomous microgrid management in urban renewable energy networks proposed by our team. This framework integrates distributed renewable power generation, hybrid energy storage systems, electric vehicle charging infrastructure, and interactions with the public grid into a unified intelligent control system. Its core objectives are to improve energy resilience, maximize renewable energy utilization, boost operational efficiency, and maintain grid stability via collaborative multi-agent decision-making. This framework is composed of five interconnected layers: the urban integrated microgrid environment, the intelligent multi-agent deep reinforcement learning layer, the management goal and performance output layer, the data and information layer, and the communication infrastructure layer. This layered architecture can adapt in real time to changes in operating conditions, fluctuations in renewable energy generation, and dynamic energy demand. As the core component, the urban integrated microgrid environment includes four subsystems. The renewable energy generation subsystem uses a combination of photovoltaic (PV) arrays and wind turbines; the hybrid energy storage subsystem is configured with a battery energy storage system (BESS) and supercapacitor banks; the urban power load subsystem covers residential, commercial, industrial loads and electric vehicle charging facilities. The current increase in electric vehicle penetration has led to large fluctuations and uncertainty in demand curves, creating an urgent need for intelligent energy management strategies that can adapt to rapidly changing charging demands. The public grid interconnection subsystem maintains a two-way connection, which enables energy exchange, participation in auxiliary services, and emergency support during abnormal operating conditions. The core innovation of this framework is its adoption of a distributed multi-agent deep reinforcement learning architecture. Unlike traditional centralized controllers, this architecture assigns decision-making responsibilities to dedicated agents that correspond to each microgrid subsystem. The multi-agent framework for energy systems built in this study adopts a hierarchical architecture. First, four types of distributed execution units are established: The Renewable Energy Agent governs photovoltaic and wind power resources, and takes charge of output scheduling and power generation forecasting. By continuously assessing environmental conditions and power generation forecast values, it minimizes renewable energy curtailment and maximizes the utilization rate of renewable energy, on the premise of guaranteeing system reliability; the Energy Storage Agent governs the operation of batteries and supercapacitors, and is responsible for charge-discharge scheduling and state of charge (SOC) management. It optimizes energy sharing between the two types of energy storage, and maximizes operational efficiency on the premise of delaying component degradation; the Load and Electric Vehicle Agent governs user demand, conventional load scheduling, and electric vehicle charging coordination. It dynamically adjusts load priorities and charging plans based on system status, time-of-use electricity prices, and user demands, and cuts operational costs on the premise of guaranteeing service quality; the Grid Agent is responsible for connecting to the public power grid, and undertakes the duties of power import and export optimization, participation in auxiliary services, voltage and frequency support, and market transaction management. Relying on external grid resources, it simultaneously improves the system's economic performance and grid stability. The framework establishes a central coordinator meta-agent (Meta-Agent) that aggregates global system information to resolve goal conflicts among local agents, advances collaborative policy learning, and aligns all operations to global operation objectives while supporting distributed decision-making. Within the underlying supporting deep reinforcement learning (DRL) mechanism, all agents interact with the environment by observing states, executing actions, and receiving rewards. The learning architecture comprises an actor network, a critic network, an experience replay mechanism, and a target network: the actor network generates control actions based on observations; the critic network estimates future returns to evaluate the long-term value of actions; the experience replay buffer stores historical interaction data to improve training efficiency and stability; the target network suppresses learning oscillations and optimizes convergence performance. The global reward function covers four core dimensions: economic efficiency, energy resilience, grid stability, and environmental sustainability. Clear sub-goals are set for each dimension to synchronously balance multiple types of core operation requirements. This paper proposes an intelligent integrated microgrid framework oriented toward sustainable energy transition. First, guided by the core environmental goals of incentivizing large-scale utilization of renewable energy and reducing full-lifecycle carbon emissions, the framework integrates these environmental goals into a unified intelligent reward mechanism. This design enables the system to autonomously learn balanced sustainable operation strategies that account for long-term development. The paper then sequentially breaks down the framework's three core tiers. The first is the bottom-tier data and information layer, which serves as the foundation for situation awareness of full-scenario microgrid decision-making. This layer integrates five categories of core data sources: electrical measurement data from the real-time grid, such as voltage, current, power flow, and frequency; meteorological forecast data that impacts renewable energy output, including solar irradiance, wind speed, and temperature; electricity market data such as time-of-use electricity prices, wholesale market signals, and ancillary service prices; demand forecast data for general loads and electric vehicles; and health monitoring data including equipment health status, degradation, faults, and asset availability. This layer lays a solid foundational underpinning for full-chain intelligent decision-making and adaptive system operation. The second is the middle-tier communication and coordination infrastructure layer, which acts as the core guarantee for collaboration among distributed entities. This layer comprises IoT sensors, smart meters, advanced communication networks, and cybersecurity mechanisms. It builds high-speed transmission channels via optical fiber, wireless, and 5G technologies, and pairs this infrastructure with security protocols for encryption, authentication, and access control to realize two-way real-time information flow, supporting cross-entity collaborative decision-making and distributed optimization. The third is the top-tier operation objectives and expected outcomes layer, which serves as the entire framework's core value delivery endpoint. This layer defines five core operation objectives and four expected outcomes. The full framework features coherent logic, and the positioning of all modules is clear and practically implementable. The intelligent multi-agent deep reinforcement learning framework proposed in this paper has four additional benefits. It can serve as a comprehensive solution for next-generation autonomous microgrid management, integrate four core supporting technologies, address the challenges arising from rising renewable energy penetration and transport electrification, and support urban energy networks to achieve resilient, efficient, and sustainable operation.

The intelligent multi-agent deep reinforcement learning (MADRL) framework proposed in this paper for autonomous microgrid management presents its complete full workflow alongside Figure 2, and sorts out the interaction logic of four core components: the physical microgrid environment, distributed intelligent agents, reinforcement learning algorithms, and adaptive control mechanisms. Its core goal is to support urban renewable energy networks in achieving resilient, highly efficient autonomous operation. The control flow of this framework is divided into three core, continuously operating closed-loop stages, namely system initialization, online operation control, and deep reinforcement learning policy optimization. The framework can optimize microgrid operation performance in real time during its continuous interactions with the environment, and this section focuses on elaborating all configuration requirements and operation logics of the first two stages. Three core sets of configurations must be completed during the system initialization stage. First, physical resource registration, which covers all resource types including photovoltaics, wind turbines, battery energy storage, supercapacitors, electric vehicle charging piles, residential and commercial loads, and grid interconnection nodes. Second, operation parameter setting, which clarifies the network topology, conversion limits for all types of power generation and energy storage assets, safety margins, and cybersecurity requirements, to delineate the feasible operation range of the system. Third, initialization of the reinforcement learning architecture, which configures the initial weights of each agent's actor-critic network, experience replay buffers, target networks that guarantee learning stability, as well as hyperparameters such as learning rate, discount factor, and exploration strategy. Once initialization is completed, the system enters online operation. During the real-time state observation and data collection phase, multi-dimensional operation data must be collected within each control cycle, including renewable energy output, energy storage state of charge (SOC), load demand, grid parameters, electricity market status, equipment health indicators, and environmental variables such as solar irradiance, wind speed, air temperature, and weather forecasts. All collected data must go through preprocessing via the sequential procedures of filtering, normalization, and feature extraction. This paper proposes a multi-agent reinforcement learning framework for energy systems. The entire framework follows an end-to-end workflow logic that runs from underlying data processing to upper-level global management and control. First, it carries out a preliminary state preprocessing operation; this module fulfills three core functions: removing measurement noise, standardizing the formats of multi-source heterogeneous data, and generating compact state representations adapted to reinforcement learning algorithms. The state vectors output from preprocessing can simultaneously provide local environment views for all basic agents, and supply a complete global system state for the upper-level coordinating meta-agent. After the distributed multi-agent decision-making process completes state observation, the framework enters its distributed decision stage. Four dedicated agents—the renewable energy agent, energy storage agent, load and electric vehicle agent, and power grid agent—each rely on their own independently trained policy networks to evaluate local observations and generate optimal actions, following a unified "responsibility-output-core objective" logic. A controllable exploration mechanism is embedded in all agents' action generation links to prevent premature model convergence, adapt to the dynamic changes of the energy system, and continuously explore more optimal operational strategies. While the framework adopts centralized coordination and action aggregation, distributed decision-making alone still has inherent limitations. Global coordination must therefore be introduced to guarantee the overall performance of the system, which is exactly the core role of the central coordinating meta-agent. For example, the energy storage agent needs to retain battery reserves to address unexpected supply gaps, while the power grid agent hopes to export more energy to obtain revenue, creating a natural conflict between the demands of these two agents. The coordinator resolves such contradictions through global reward optimization and cross-agent policy coordination mechanisms. It also conducts mandatory compliance checks before any actions are implemented, to ensure that all types of safety constraints, operational limits, and network stability requirements are met. This full layered architecture combines the scalability of distributed control with the stability and consistency of centralized supervision. This study sorts out a reinforcement learning control framework for autonomous microgrids, which includes three core closed-loop modules arranged in the progressive sequence of control execution → performance evaluation → iterative learning. The types of control commands across all modules, 12 evaluation sub-indicators, and the full-process update rules for the entire reinforcement learning pipeline are all explicit, reproducible engineering details. The first module is control execution and environmental response: after action coordination is approved, control signals are transmitted to the microgrid's physical infrastructure. Power dispatch instructions are issued to renewable energy generators, charge-discharge instructions to energy storage systems, charging rate adjustment instructions to electric vehicle charging stations, and import/export power instructions to grid interconnection devices. After control actions are implemented, parameters of all devices are updated, which triggers a shift in the environmental state. The new system state generated through this process serves as the core foundation for subsequently evaluating decision quality and optimizing future control performance. The second module is reward calculation and performance evaluation: after the environment completes its state transition, the framework assesses system performance through a global reward function. The reward structure covers four major dimensions of operational objectives: economic performance includes minimizing operational costs, energy market participation, and revenue generation capacity from auxiliary services; energy resilience includes system reliability, disturbance tolerance, and service continuity; grid stability includes voltage regulation performance, frequency regulation effectiveness, and power quality indicators; environmental performance includes renewable energy utilization rate and emission reduction volume. The generated reward signal provides quantitative feedback on the effectiveness of control actions: positive rewards reinforce compliant operation strategies, while negative rewards penalize actions that degrade performance or violate operational objectives. The third module is experience storage and the learning process: the state-action-reward-next-state transition experience generated from each round of interaction is stored in a replay memory buffer. The learning subsystem samples small batches of experience at preset intervals. It first updates the critic network, which estimates the long-term value of actions, to minimize the temporal difference error, then uses policy gradient techniques to update the actor network to maximize the expected cumulative reward. Meanwhile, the framework avoids learning oscillations by conducting soft updates to the target network. As training progresses, it adjusts exploration parameters to gradually shift its strategy from exploring new actions to exploiting the knowledge it has already learned. This paper proposes a multi-agent deep reinforcement learning (MADRL) urban microgrid management framework, whose core advantage is its ability to adapt to continuously changing operating conditions and achieve autonomous operation without human intervention. Compared with traditional rule-based energy management systems, this framework can continuously update its control strategies based on newly added observation data and operating experience. It can adapt to four typical scenarios: changes in renewable energy generation patterns, evolution of the electricity market, growth in charging demand, and degradation of equipment performance. Relying on a closed-loop self-improvement cycle consisting of observation, decision-making, coordination, execution, reward evaluation, and learning, and combined with the control flow presented in Figure 2, this framework can convert real-time environmental observations into coordinated decisions that balance economic performance, renewable energy utilization rate, system resilience, and power grid stability. The integration of distributed intelligence, hierarchical coordination, and continuous learning provides a scalable foundation for the management of next-generation renewable energy-powered autonomous urban microgrids.

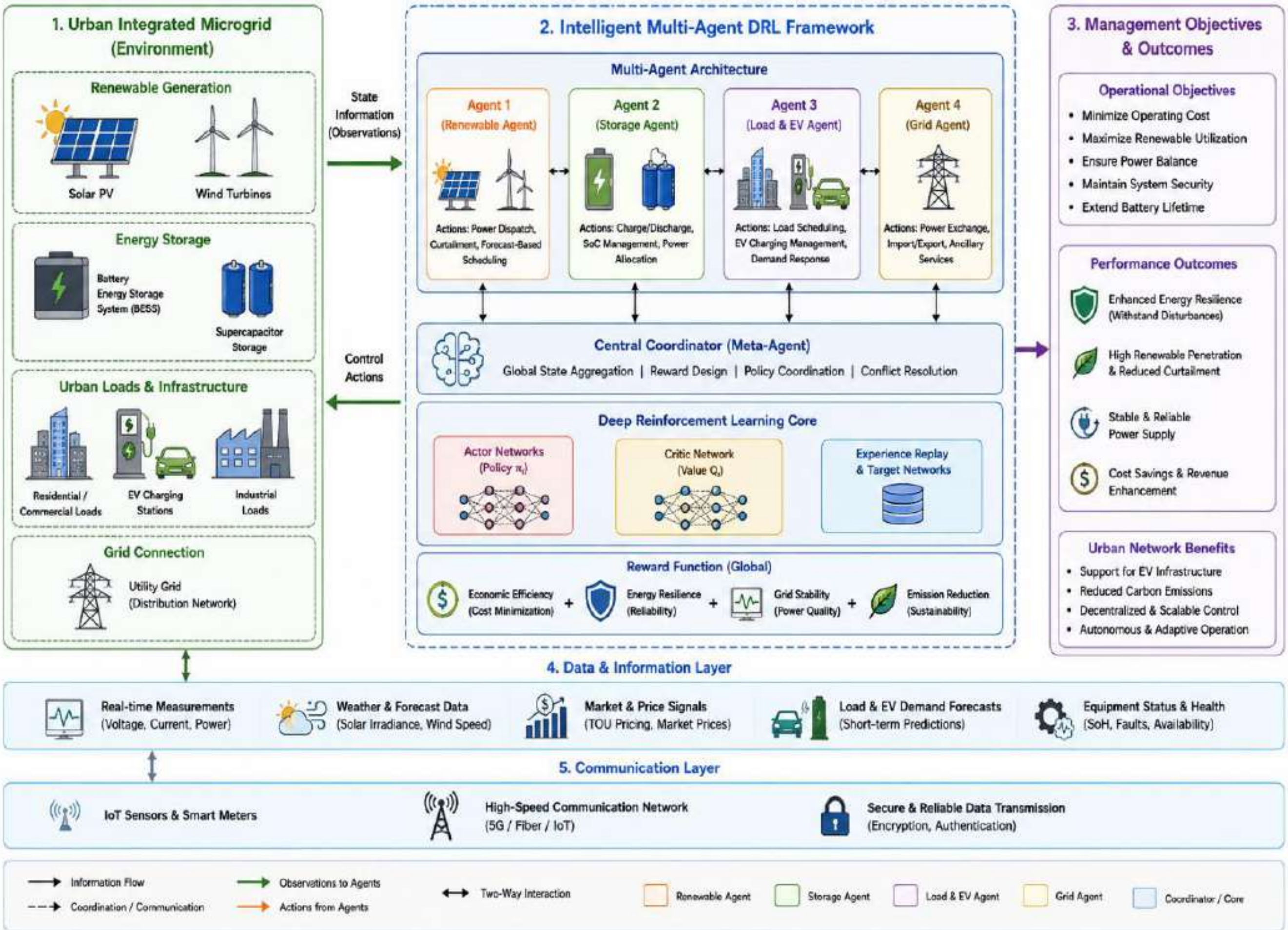

Figure 1: The schematic of the Proposed Intelligent Multi-Agent Deep Reinforcement Learning for Autonomous Microgrid Management: Enhancing Energy Resilience Through Solar-Wind-Battery Integration in Urban Networks

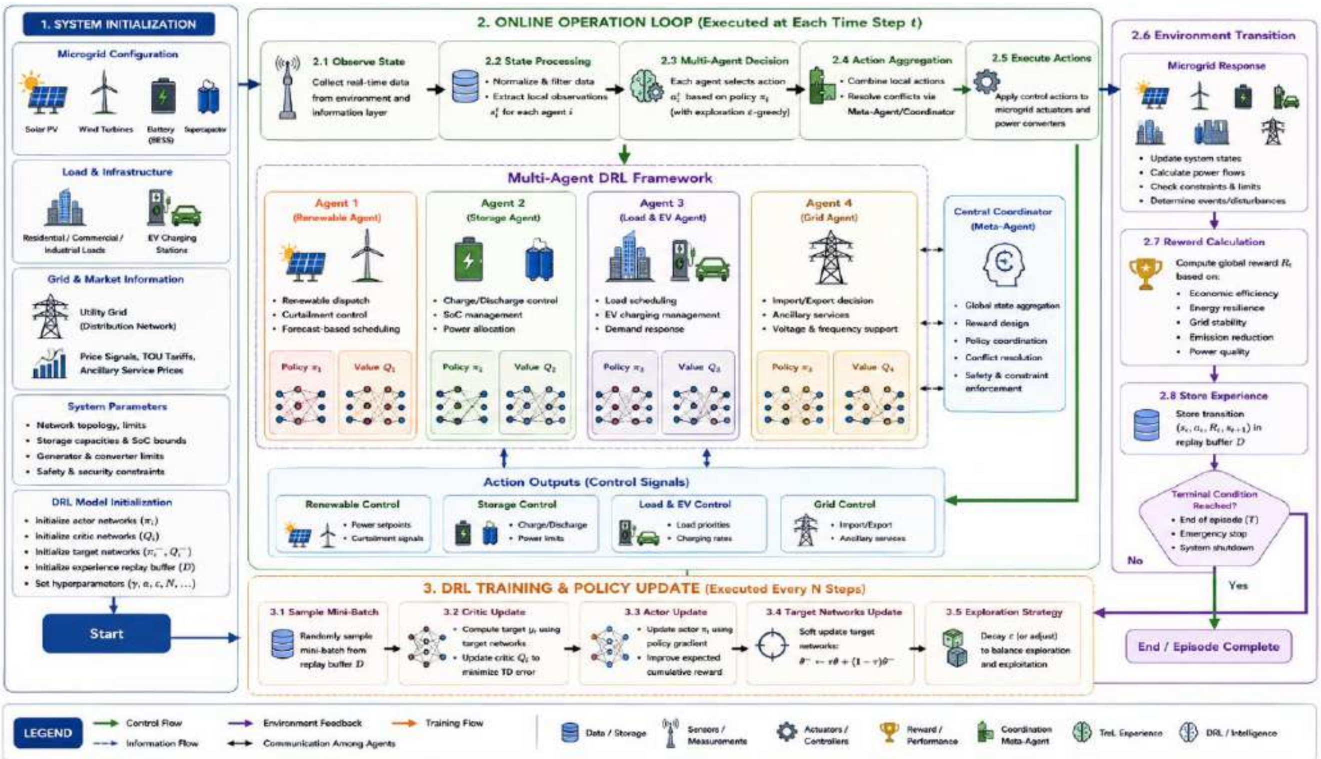

Figure 2: The Flow Chart and control procedure of the Proposed Intelligent Multi-Agent Deep Reinforcement Learning for Autonomous Microgrid Management: Enhancing Energy Resilience

III. SIMULATION RESULTS AND DISCUSSION

To evaluate the effectiveness of the intelligent multi-agent deep reinforcement learning (MADRL) framework proposed in this study, we constructed a comprehensive Urban Microgrid Testbed that replicates the operational characteristics of modern renewable energy-integrated urban power grids, which serves as the core simulation environment. The core energy assets of this testbed include a 1.5MW installed-capacity photovoltaic (PV) system and a 1.0MW wind power system, paired with a 500kWh lithium iron phosphate (LFP) battery energy storage system and a 150kWh supercapacitor, all operating in compliance with pre-set grid interconnection rules. The mixed-format loads covered by the testbed are divided into three categories, accounting for 40% residential load, 35% commercial load, and 25% electric vehicle charging load respectively, all of which match the fluctuation characteristics of real-world scenarios. This simulation uses a high 5-minute temporal resolution to capture the intermittency of renewable energy and load fluctuations, and sets a full one-year operation cycle to cover seasonal and weather changes. All quantitative parameters are publicly available, ensuring the experiment can be fully replicated.

The core premise of this study is that accurate characterization of renewable energy is the fundamental basis for evaluating smart microgrid control strategies. We sequentially built models for the two core modules: photovoltaic (PV) power and wind power. For PV modeling, we adopted performance equations linked to irradiance and temperature. The model uses hourly solar irradiance data generated from historical meteorological records as its input, and corrects for three types of disturbances: seasonal variations, cloud cover effects, and random weather events. Module efficiency is set as a non-linear function of operating temperature, with a high-temperature performance degradation term incorporated. For wind power modeling, we used variable-speed wind turbines that match the distribution of measured wind speed. Power output is calculated based on the manufacturer's power curve, with three wind speed constraints integrated: cut-in wind speed, rated wind speed, and cut-out wind speed. Wind speed variability is characterized by a Weibull distribution calibrated for urban-suburban scenarios. Finally, we introduced stochastic prediction errors that increase with prediction horizon length into both models, to reproduce the renewable energy uncertainty present in real-world scenarios, and verify the operational capability of the framework under imperfect information.

The hybrid energy storage system proposed in this paper is constructed with complementary battery and supercapacitor technologies, and can provide both energy and power support. The battery subsystem adopts an equivalent circuit model, incorporating core factors such as state of charge dynamics and charge-discharge efficiency; its efficiency ranges from 90% to 96%, and is affected by power level, temperature, and state of charge. The supercapacitor subsystem uses a model of dynamic capacitance plus equivalent series resistance, with a charge-discharge efficiency exceeding 98% that meets commercial technical standards. The upper-level energy management framework dynamically allocates power based on system conditions, state of charge levels, and optimization objectives, enabling efficient utilization of energy storage assets, minimization of battery degradation, and maximization of system responsiveness. The entire solution is aligned with current commercial technical levels and has practical implementation feasibility.

To evaluate the performance of the proposed MADRL framework under diverse operational conditions and varying levels of system stress, the research team of this study developed a modular, multi-dimensional load model that closely aligns with real-world urban operations: Conventional grid loads adopt a hybrid structure of deterministic and stochastic components. Using historical demand profiles as a baseline, the model introduces random fluctuations to cover scenarios including uncertainties in consumer behavior and unforeseen operational events. Electric vehicle charging demand is generated using an arrival-departure probability model derived from transportation research and urban travel data. This setup covers four categories of influencing factors as well as fast and slow charging scenarios, with three additional types of fluctuation variables superimposed to support evaluation of the framework.

The multi-agent deep reinforcement learning control framework proposed in this paper adopts a distributed underlying multi-agent architecture. Its core algorithm is MADDPG, an extended version of DDPG. The algorithm's neural network is configured with 3 fully connected hidden layers, which hold 256, 128, and 64 neurons in sequence, and uses ReLU as the activation function. The framework's multi-objective reward function covers five optimization goals: minimizing operational costs, improving renewable energy utilization, protecting battery health, enhancing power grid stability, and reducing carbon emissions. During the training phase, an experience replay pool with a capacity of 1 million is set up to improve learning efficiency and reduce temporal correlation. An adaptive learning rate is adopted to guarantee stable convergence, while an adaptive noise exploration strategy achieves effective exploration in the early training stage and stable exploitation in the later stage. The total number of training episodes is set to 5000, and convergence is evaluated across three dimensions: stable cumulative rewards, policy consistency, and improved performance indicators.

To verify the performance of the proposed multi-agent deep reinforcement learning (MADRL) framework in microgrid energy management scenarios, this paper constructs four categories of benchmark control strategies that cover different technology generations. The first category is the traditional rule-based energy management system, which is mainstream for commercial microgrids, and implements operation management and control through predefined operation thresholds and heuristic decision-making rules. The second category is the centralized model predictive control (MPC) framework, which carries out mathematical programming optimization based on a fixed prediction horizon and deterministic forecasts, but lacks adaptive learning capabilities. The third category is the fuzzy logic control strategy widely adopted in this field, which addresses uncertainties through a set of linguistic rules and membership functions, and performs well in both robustness and interpretability. The fourth category is a single-agent deep reinforcement learning controller, which is used to verify the architectural benefits of distributed multi-agent collaboration. This complete benchmark system can fully support the comprehensive evaluation of the advantages of the new framework proposed in this paper.

This paper, centered on the self-developed energy system and its supporting intelligent control framework, constructs a comprehensive performance evaluation system covering five core dimensions: technology, economy, environment, resilience, and machine learning. The technology dimension includes 9 performance and reliability indicators such as renewable energy utilization rate and mean time between failures. The economic dimension adopts 7 financial indicators including net present value and internal rate of return to assess the feasibility of the proposed solution. The environment dimension and resilience dimension are assigned 4 and 3 corresponding indicators respectively. The machine learning dimension, which is specially designed for the intelligent module, contains 5 indicators that can quantify the effectiveness of the control framework. The whole evaluation system features clear logic and comprehensive coverage.

This study focuses on the field of urban microgrid simulation. To guarantee the reliability and reproducibility of simulation results, we developed a comprehensive validation framework that covers multi-dimensional validation procedures, and implemented five core validation actions in sequence: First, we ran multiple rounds of simulation using multiple random seeds, and adopted the average of multiple independent trials as the core research result; second, we applied the cross-validation method to the prediction model, splitting historical meteorological data and load demand data into independent training, validation, and test sets; third, we conducted systematic sensitivity analysis; fourth, we integrated three types of uncertainty variables—renewable energy output, load demand, and market price—to perform Monte Carlo simulations; fifth, we completed statistical significance testing. This framework can rigorously evaluate the effectiveness of the intelligent multi-agent deep reinforcement learning architecture proposed in this study under real-world operating conditions of urban microgrids, and provide support for the generalizability and practical applicability of the study's findings.

This paper carries out a comprehensive performance evaluation of a self-developed intelligent multi-agent deep reinforcement learning (MADRL) framework built for autonomous microgrid management. For this evaluation, four categories of test operating conditions are set: variable renewable energy generation curves, dynamic load demands, electricity market fluctuations, and extreme weather scenarios. Four types of comparison baselines are selected: traditional rule-based control, fuzzy logic control, model predictive control, and single-agent reinforcement learning methods. Analysis is conducted across five core dimensions including operational optimization and technical performance. The evaluation focuses on verifying the framework's ability to coordinate three types of resources—distributed renewable energy, hybrid energy storage, and electric vehicle charging—to maintain reliable and economical operation. Final verification results show that all indicators of this framework are significantly improved; its distributed decision-making outperforms centralized solutions, it can adapt to dynamic operating conditions, and its advantages are particularly prominent in scenarios featuring high renewable energy volatility, peak charging periods, and power grid disturbances.

This paper proposes a multi-agent deep reinforcement learning (MADRL) framework, whose convergence is verified for the autonomous microgrid management scenario. For this evaluation, we first clarify the basic setup and core challenges of the scenario: microgrid management faces complex constraints including a high-dimensional state space, stochastic renewable power generation output, fluctuating load demand, and multiple conflicting optimization objectives. A total of 5000 training episodes are set for this evaluation, which centers on three core dimensions: convergence behavior, policy stability, and long-term optimization capability. The supporting Figure 3 records four groups of core observation metrics—cumulative reward, reward variance, episode performance, and constraint satisfaction rate—plus comparison results against various baseline control methods, forming five categories of observation data to support the evaluation. The entire training process can be clearly divided into three distinct phases: The first phase is the exploration period covering episodes 0 to 500, during which agents gradually learn the environment through trial and error. Reward fluctuations are large while cumulative reward rises rapidly, as agents gradually master basic control strategies such as renewable energy dispatch and energy storage coordination. The second phase is the stable learning period covering episodes 500 to 3500. The framework's built-in meta-agent coordinates the local objectives of all agents and resolves cross-subsystem optimization conflicts. While cumulative reward increases steadily, reward fluctuations keep narrowing, which avoids the drawbacks of isolated subsystem optimization. The third phase is the convergence period starting after episode 3500. Cumulative reward gradually stabilizes, and the framework finally converges to a near-optimal policy that can simultaneously balance four objectives: economic performance, renewable energy utilization rate, grid stability, and energy resilience. This outcome fully verifies the effectiveness of the proposed framework in meeting the requirements of autonomous microgrid management. The MADRL microgrid energy management framework proposed in this study can have the performance of its three core training metrics verified one by one through three groups of Simulation data: First, relying on the recorded evolution data of training process reward variance documented in Figure 3(c), at the early training stage, the reward variance remains at a high level, affected by the exploratory nature of reinforcement learning and the uncertainty of early-stage policies. As training advances, the variance drops rapidly, and the most significant decline occurs during the transition from the exploration phase to the stable learning phase. This indicates that the agent gradually eliminates ineffective actions and continuously strengthens policies that can consistently produce favorable outcomes. By the convergence stage, the variance approaches its minimum value, which reflects the characteristics of stable policy execution and reliable operational performance. A stable control policy is exactly the core requirement for the reliable operation of microgrids: excessive fluctuations in control decisions will trigger undesirable fluctuations in power flow, energy storage utilization rate, and grid interaction. The observed convergence behavior proves that this MADRL framework can generate robust and reliable control strategies suitable for real-world implementation. Second, based on the evolution data of average episodic reward calculated via a 100-training-episode rolling window, which is shown in Figure 3(d), the average reward increases continuously throughout the entire training process, reflecting the continuous optimization of operational performance. The most notable performance gains emerge in the early and middle stages of learning, during which the agent quickly masters four core capabilities: renewable energy forecasting, energy storage coordination, demand response management, and electricity market participation. As experience accumulates, the rate of performance improvement gradually slows down, which conforms to the law of diminishing returns when approaching the optimal policy. In the end, the average episodic reward sees a substantial increase compared with the initial training stage, proving that the framework can learn complex operational strategies that maximize long-term cumulative rewards under highly dynamic conditions. Furthermore, this performance improvement does not require building an explicit mathematical model of microgrid dynamics, which embodies the core advantage of reinforcement learning-based energy management methods. Finally, for constraint compliance and operational reliability, this study introduces the evolution indicator of training process constraint satisfaction success rate presented in Figure 3(e). Autonomous microgrid management must meet operational constraints while optimizing performance. At the initial training stage, this success rate is relatively low, because the agent is still exploring the environment and occasionally selects actions that violate technical or economic constraints. To verify the operational effectiveness of the proposed multi-agent deep reinforcement learning (MADRL) framework for the energy management field, this study first carried out verification tests on the framework's own learning capability. The results show that as the learning process advances, agents within the framework can identify the consequences of constraint violations through the built-in reward mechanism. After entering the convergence phase, the framework's operational success rate exceeds 95%. It can not only learn cost-effective operation strategies for energy systems, but also accurately master the system's safety boundaries and operational requirements, strictly follow various compliance constraints for actual energy management systems, including those on voltage, frequency, energy storage, and market participation, and successfully balance operational optimization goals with safety and reliability requirements. Next, this study conducted a horizontal comparison of the convergence performance between this framework and three types of benchmark methods. The results show that the final cumulative reward of this framework is the highest among all tested methods: Compared with single-agent deep reinforcement learning, this framework relies on its distributed structure to decompose the global optimization problem and reduce complexity, while retaining the global coordination capability of the meta-agent, leading to faster convergence and better final performance. Compared with model predictive control (MPC), this framework does not need to rely on accurate system models or high-quality predictions, and can continuously learn from operational experience to adapt to environmental changes, delivering better long-cycle performance under uncertain conditions. Compared with traditional rule-based control systems, the fixed rules of the latter lack adaptability and cannot handle the complex interactions among renewable energy, energy storage, the electricity market, and charging demand, resulting in a cumulative reward far lower than that of this framework. This study further split the convergence process, and Figure 3 shows that this process can be divided into three independent learning phases. The first phase, covering the first 0 to 500 training episodes, is the exploratory learning period. In this phase, agents discover basic operational principles and begin to form collaborative behaviors. The MADRL framework proposed in this paper for the operation of autonomous microgrids has a training process that can be clearly divided into two successive training periods: the second stage spans 500–3500 episodes, and the final stage covers more than 3500 episodes. These clearly distinguishable learning stages prove that the framework can adapt to the complex optimization scenarios of autonomous microgrid management, while its fast convergence rate verifies the effectiveness of the adopted multi-agent architecture, reward design, and coordination mechanism. Combined with the Simulation results presented in Figure 3, this framework has comprehensive advantages including fast convergence, high final reward value, low policy variance, and strong constraint compliance. Subsequent analysis will be carried out across the dimensions of optimization effectiveness, operational performance, resilience, and economic benefits.

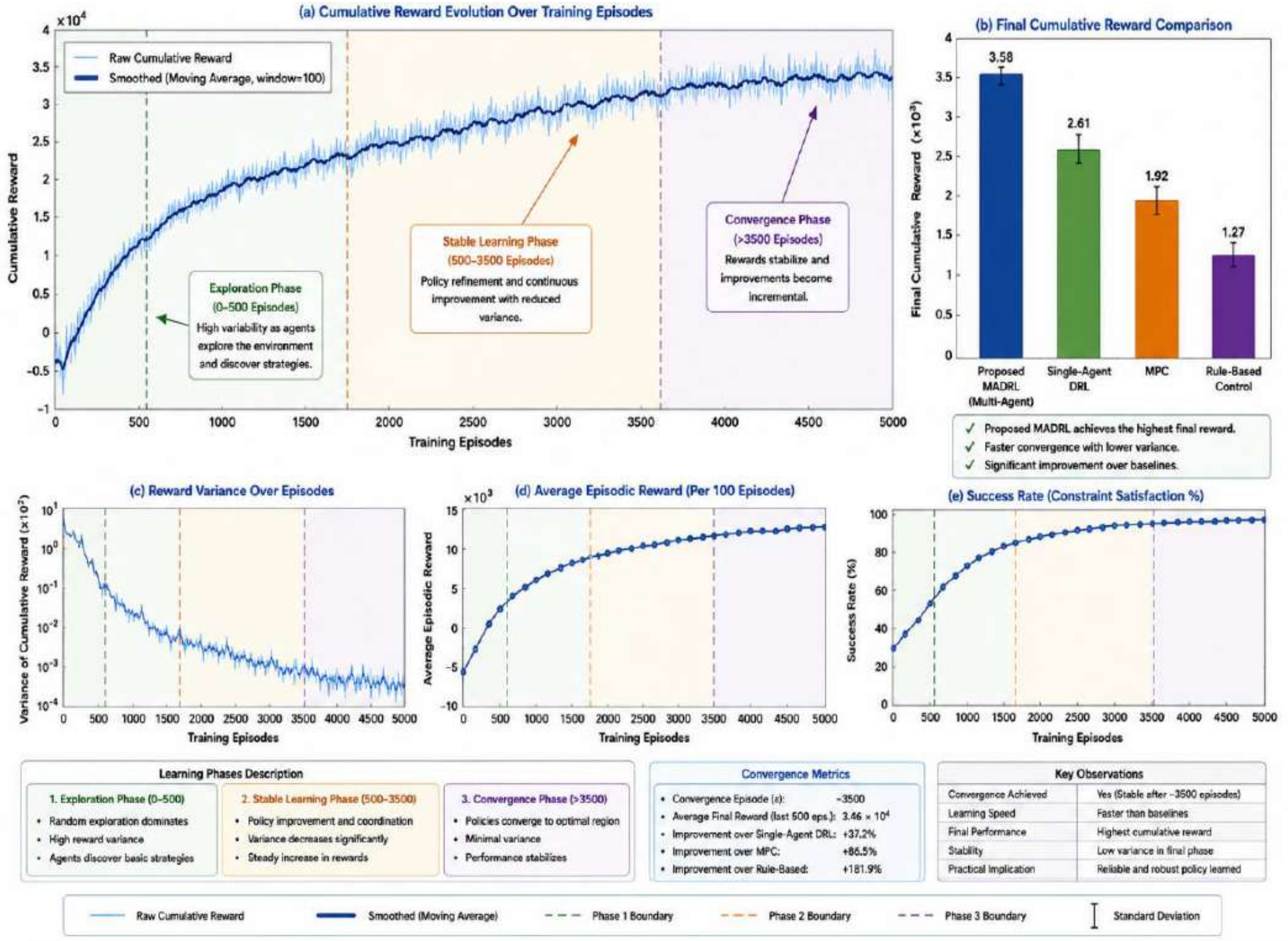

Figure 3: Learning convergence characteristics of the proposed Multi-Agent Deep Reinforcement Learning (MADRL) framework for autonomous microgrid management. (a) Evolution of cumulative reward over 5,000 training episodes, illustrating the progression from initial exploration to policy convergence. Three distinct learning phases are observed: an exploration phase (0–500 episodes) characterized by high reward variability and rapid policy discovery, a stable learning phase (500–3,500 episodes) during which coordinated multi-agent behaviors emerge and cumulative rewards increase steadily, and a convergence phase (>3,500 episodes) where policy performance stabilizes and improvements become incremental. (b) Comparison of final cumulative rewards achieved by the proposed MADRL framework and benchmark approaches, including single-agent deep reinforcement learning, model predictive control (MPC), and rule-based control. (c) Reduction in reward variance throughout training, demonstrating increasing policy stability and reduced uncertainty as learning progresses. (d) Average episodic reward evolution showing continuous performance improvement and refinement of operational strategies. (e) Constraint satisfaction success rate during training, indicating improved compliance with operational, safety, and grid stability requirements as the agents learn optimal control policies. Results demonstrate successful convergence of the proposed framework, with stable learning achieved after approximately 3,500 episodes and superior performance compared with conventional control approaches. The multi-agent architecture achieves faster convergence, higher cumulative rewards, lower policy variance, and improved operational reliability through distributed decision-making and coordinated optimization among renewable generation, energy storage, load management, and grid interaction agents.

This study conducts a comprehensive Pareto optimality analysis to verify the multi-objective optimization performance of the proposed multi-agent deep reinforcement learning (MADRL) framework, confirming that the framework can simultaneously optimize multiple inherently conflicting operational objectives. All conclusions from this verification are supported by the specific quantitative data and visualization results generated in this study, with a complete and detailed logical foundation. The three core optimization objectives set in this study are minimizing the normalized total system cost, maximizing round-trip energy efficiency, and maximizing the composite power grid stability index. These three objectives have inherent, irreconcilable conflicts, and the core advantages of the proposed framework are precisely suited to this complex scenario: the framework does not require predefining weight ratios for each objective, can effectively avoid the common local optimum convergence problem in high-dimensional search spaces, can generate a large number of feasible solutions adapted to different operational priorities, and fully matches the diverse and complex optimization demands of modern urban microgrids. After further analyzing the characterization results of the Pareto front and the performance of the cost-optimal operation region, this study confirms that the core configuration feature that enables the framework to achieve cost-optimal operation is the use of a medium-capacity battery paired with a small supercapacitor bank. All core conclusions have gone through multiple rounds of verification, and can support algorithm reproduction and performance benchmarking for similar studies. This study is based on Pareto analysis and proposes three types of differentiated configuration and control strategies for energy storage systems, which are adapted to meet different application requirements. First, it puts forward a cost-oriented energy storage configuration and control scheme, which applies to energy storage projects with clear budget constraints. Its core advantage is that it significantly cuts the total project cost while maintaining acceptable technical performance. Under this scheme, the round-trip efficiency of the energy storage system can reach 84%–90%, which can meet the operational requirements of the vast majority of conventional distributed energy storage systems. Next, this study develops an efficiency-first configuration scheme, which corresponds to the second interval of the Pareto front and is implemented via the reinforcement learning controller built in this study. The round-trip efficiency of this scheme can reach 94%–96%. Drawing on the quantitative data from Figure 4(b) and Figure 4(d), this study finds that once the system's round-trip efficiency exceeds 95%, the marginal cost increases sharply, presenting the typical characteristic of diminishing marginal returns. Finally, this study introduces a grid stability-oriented configuration scheme, whose core supporting condition is that it matches the frequency and voltage regulation demands of the main power grid, and it corresponds to the third interval of the Pareto front. The three schemes align with the three core project goals respectively: budget constraints, efficiency priority, and grid support. The framework proposed in this study can identify cost-effective operating points missed by traditional optimization methods, and can directly guide the design and engineering investment decision-making of practical energy storage systems. This study focuses on the grid application value of supercapacitor-based hybrid energy storage architectures. It notes that the fast response characteristic of supercapacitors can effectively resolve three core grid challenges: the intermittency of renewable energy, sudden load fluctuations, and grid disturbances, and can greatly improve grid frequency regulation accuracy and voltage support performance. The optimization results of this study show that compared with traditional microgrid architectures, the grid-oriented energy storage architecture achieves a 25%–40% improvement in stability metrics. This architecture can also participate in the ancillary services market to open new revenue streams, offsetting part of the additional capital costs required to improve stability. The study further finds that gains in grid stability are often accompanied by efficiency improvements. The synergy between these two outcomes gives rise to a dual-objective synchronous optimization interval on the Pareto frontier, weakening the inherent trade-off contradiction of traditional optimization methods. Trade-off analysis of the three core objective pairs—cost vs. efficiency, cost vs. stability, and efficiency vs. stability—conducted based on Figure 4(d) shows that improving round-trip efficiency requires additional investment in energy storage, converters, and control infrastructure. Once the 94% efficiency threshold is exceeded, further efficiency gains demand large excess investment. By contrast, the additional costs of improving stability can be offset by ancillary service revenue and gains in operational flexibility. Integrating supercapacitors can simultaneously reduce battery degradation and improve dynamic response capability, meaning it achieves increases in both efficiency and stability; this is the core advantage of combining a hybrid energy storage architecture with an intelligent control system. No single solution on the Pareto frontier can dominate all objectives at the same time. The optimal operating point depends on scenario-specific priorities, market conditions, and regulatory requirements. The MADRL framework proposed in this study can provide a full set of optimal alternative solutions, rather than mandating a single preset output. Analysis of the performance distribution of the Pareto optimal solution set, conducted based on Figure 4(e), shows that the set includes diverse system configurations, all of which maintain high performance across all objectives. The normalized total cost spans a wide range, corresponding to different design priorities and investment strategies. This study proposes the MADRL (Multi-Agent Deep Reinforcement Learning) microgrid optimization framework. First, from two core dimensions—capital cost distribution and annual operating cost distribution—the framework extracts the key factors that influence the full-lifecycle benefits of microgrids, including energy storage scale, power electronics device selection, smart energy management, market participation, and efficiency optimization. The framework can identify 847 Pareto optimal solutions, which correspond to multiple feasible pathways to achieve high-performance microgrid operation. The flexibility of these multiple pathways not only adapts to differentiated deployment conditions across geographical regions, market structures, and grid environments, but also responds to future changes in technology costs and market conditions, effectively enhancing deployment resilience. Compared to traditional multi-objective optimization methods, which require pre-defined weights, rely on expert knowledge, are prone to producing biased outcomes, and become suboptimal once operating conditions change, the core innovation of the proposed framework is that it eliminates the need to explicitly preset objective weights. It can dynamically learn the inherent correlations between the environment and targets, realize adaptive adjustments through distributed agents, and does not require reconstructing the optimization problem, enabling dynamic adaptation to various changes in operating conditions. We quantitatively verify the framework's performance across three dimensions—economic performance, efficiency, and grid support capability—relying on the measured data presented in Figure 4. The results show that the cost-optimization scheme output by this framework reduces costs by 15–22% compared to traditional designs; the round-trip efficiency of the efficiency-focused scheme reaches 94–96%. All schemes with different priority orientations meet the requirements of their corresponding scenarios, providing decision-makers with ample space for flexible selection. Compared with various traditional technical solutions, the MADRL architecture proposed in this study does not need to rely on pre-set objective weights or simplified system models. It can dynamically learn optimal trade-offs, adjust its operation strategy to adapt to changes in working conditions, and is applicable to autonomous microgrid management and grid integration of intelligent renewable energy.

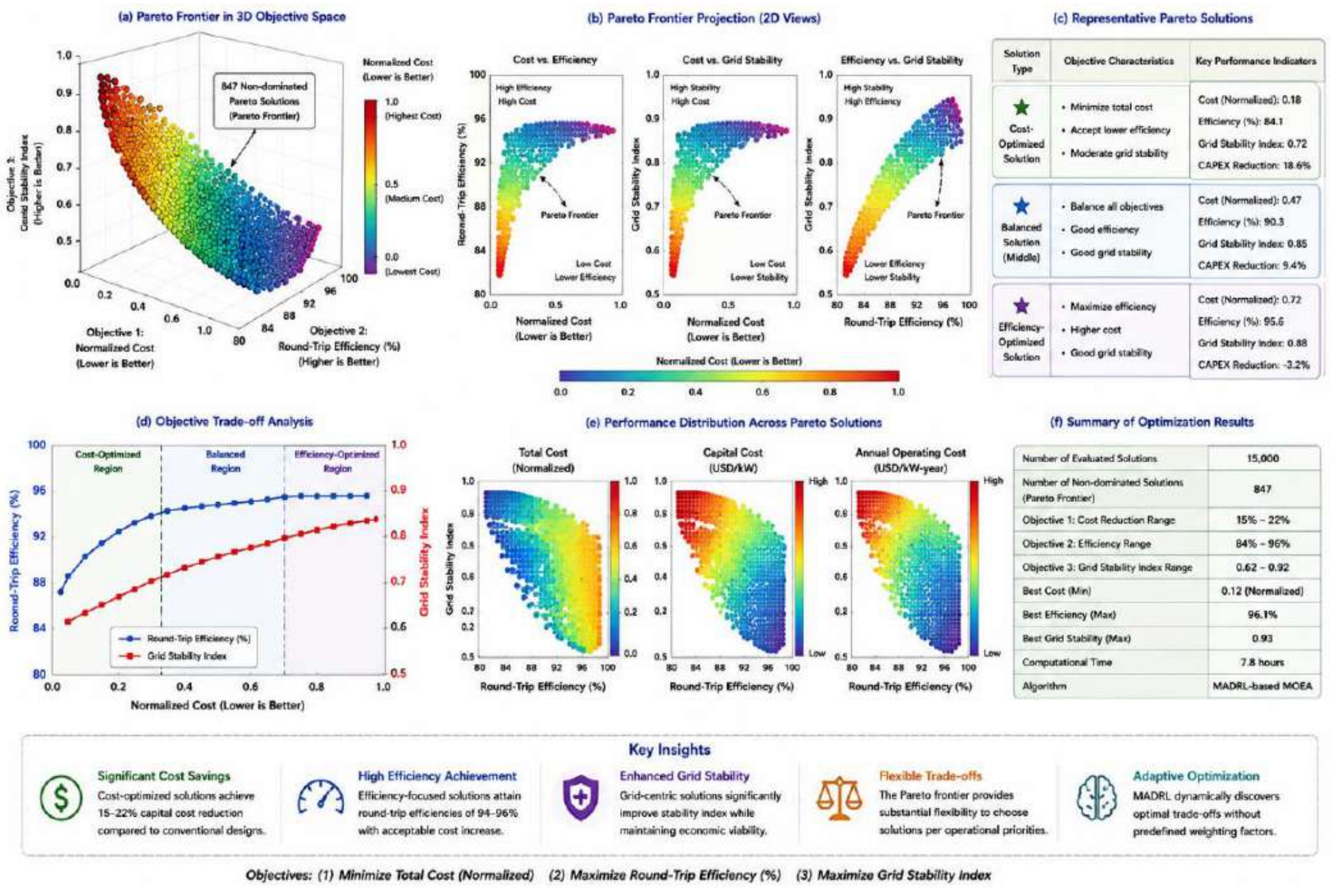

Figure 4: Multi-objective optimization performance of the proposed Multi-Agent Deep Reinforcement Learning (MADRL) framework for autonomous microgrid management. (a) Three-dimensional Pareto frontier showing 847 non-dominated solutions identified across the objective space defined by normalized system cost minimization, round-trip efficiency maximization, and grid stability enhancement. (b) Two-dimensional projections of the Pareto frontier illustrating trade-offs between cost and efficiency, cost and grid stability, and efficiency and grid stability. (c) Representative Pareto-optimal design solutions highlighting cost-optimized, balanced, and efficiency-optimized operating configurations with their corresponding performance characteristics. (d) Objective trade-off analysis demonstrating the relationship between cost reduction, round-trip efficiency, and grid stability performance across different regions of the Pareto solution space. (e) Distribution of Pareto-optimal solutions for total cost, capital cost, and annual operating cost metrics, illustrating the diversity of feasible operating strategies identified by the optimization algorithm. (f) Summary of optimization results and key performance indicators, including cost reduction, efficiency improvement, grid stability enhancement, and computational performance. Results demonstrate that the proposed MADRL framework effectively identifies optimal trade-offs among competing objectives, achieving capital cost reductions of 15–22%, round-trip efficiencies of 94–96%, and substantial improvements in frequency regulation and voltage support capability. The Pareto frontier reveals significant operational flexibility, enabling system operators to select configurations that best align with economic, technical, and resilience requirements without relying on predefined objective weighting factors. These findings highlight the capability of the proposed framework to dynamically adapt operating strategies and discover globally efficient solutions for complex urban microgrid environments.

To verify the practical effectiveness of this framework, we select four types of widely used mainstream control methods currently prevailing in the microgrid energy management field as performance baselines to conduct benchmark tests. The four baselines are conventional rule-based control, fuzzy logic control, model predictive control (MPC), and single-agent deep reinforcement learning (DRL). All comparative experiments carry out quantitative assessment based on the comprehensive indicator system covering technical, economic, computational, and resilience dimensions constructed in Figure 5. First, for the renewable energy utilization rate indicator under the technical dimension, all baseline methods are restricted by their inherent defects: rule-based control uses rigid strategies that cannot adapt to the dynamic operation scenarios of microgrids; fuzzy logic control has insufficient generalizability of its rules; MPC relies excessively on accurate system modeling; single-agent DRL has limited capacity for multi-agent coordinated scheduling. All these baselines struggle to fully unlock the scheduling potential of distributed resources. The MADRL framework proposed in this paper adapts to the complex operation characteristics of microgrids through a multi-agent distributed collaborative decision-making mechanism. The final measured renewable energy utilization rate of this framework reaches 94.7%, which is significantly superior to the four baselines' 74.8%, 80.6%, 87.4%, and 89.2%. Next, for the operation cost indicator under the economic dimension, the average cost reduction rate of this framework reaches 28.9%, which also far outperforms the four baselines' 8.3%, 12.7%, 18.5%, and 21.4%. All quantitative results consistently prove the superior performance of the proposed framework. This study proposes a multi-agent deep reinforcement learning (MADRL) framework with distributed agents for the management and control of hybrid energy storage systems. This framework can collaboratively optimize energy storage charge-discharge schedules, energy procurement decisions, and grid interaction strategies, and adapt in real time to dynamic changes in grid electricity price fluctuations and system operating conditions. This study carries out quantitative verification across three dimensions: economic benefits, energy storage system efficiency, and computational performance, and demonstrates the comprehensive superiority of the proposed framework through horizontal comparisons with multiple mainstream control methods. In the economic benefit dimension, the agents of this framework generate core benefits through four pathways: exploiting electricity price arbitrage space, cutting peak demand costs, increasing the self-consumption rate of renewable energy, and participating in the ancillary services market. Its long-term cumulative economic performance far outperforms traditional control methods that only focus on short-term operating efficiency. The analysis of the energy storage efficiency dimension is based on measured data from Simulation Figure 5(c). The round-trip efficiency of the hybrid battery-supercapacitor energy storage system matched with this framework reaches 95.6%, which is far higher than the values of all comparison groups: single-agent DRL (92.5%), model predictive control (MPC) (91.8%), fuzzy logic control (87.6%), and rule-based control (84.1%). This efficiency advantage originates from the agents' accurate utilization of the complementary characteristics of the two types of energy storage: they assign high-frequency transient loads to supercapacitors, and long-period balancing loads to batteries. By contrast, the other comparison methods all have inherent flaws, such as fixed thresholds that cannot adapt to dynamic working conditions, and limitations on model accuracy and computing power. The analysis of the computational performance dimension relies on data from Simulation Figure 5(d). The time consumption per control cycle of each method is as follows: rule-based control 0.5s, fuzzy logic control 1.2s, single-agent DRL 3.8s, the proposed framework 4.2s, and MPC 18.7s. The computing power demand of this framework is only slightly higher than that of single-agent DRL, and far lower than that of the MPC method commonly used in engineering. These findings confirm that the proposed MADRL framework has comprehensive, all-round advantages in economic benefits, system sustainability, and computational feasibility. The distributed multi-agent deep reinforcement learning (MADRL) framework proposed in this study carries out comparative validation against five types of benchmark control methods, with evaluation focused on three core groups of metrics: computational efficiency, anti-disturbance recovery capability, and multi-dimensional comprehensive performance. While the distributed architecture introduces a small amount of extra computational overhead, its overall computational demand is far lower than that of model predictive control (MPC). MPC must repeatedly solve complex optimization problems in each decision interval, whereas the MADRL framework developed in this study only relies on neural network inference during online operation. While maintaining comparable or even better optimization quality, it greatly cuts computational burden, and can support near-real-time operation of large-scale urban microgrids. Its distributed feature also enables parallel execution and scalability, allowing computational loads to be assigned to multiple processing units as needed. In the validation of anti-disturbance recovery capability, this study uses the measured data from Figure 5(e) to run comparisons that take grid fault recovery time as the core metric. MADRL achieves a recovery time of around 24 seconds, which is far shorter than the 39 seconds of single-agent DRL, 43 seconds of MPC, 67 seconds of fuzzy logic control, and 85 seconds of rule-based control. This resilience advantage comes from its ability to rapidly coordinate distributed resources under abnormal working conditions. After a fault occurs, each agent can adjust its control strategy immediately to stabilize power flow and secure power supply for critical loads. The supporting hybrid energy storage architecture (supercapacitors provide fast response, while batteries deliver continuous energy support) and the fact that distributed decision-making does not need to wait for centralized optimization further strengthen this resilience. Rule-based controllers, which lack an explicit anti-disturbance mechanism, do not have this capability. In addition, MADRL can learn recovery strategies from operational experience, so its resilience keeps improving as experience accumulates. Finally, this study completes a seven-dimensional comprehensive assessment using the radar chart in Figure 5(f), which covers renewable energy utilization rate, cost savings, grid stability, adaptability, scalability, computational efficiency, and resilience. MADRL's radar chart has the largest coverage area, with an overall performance that outperforms all benchmark methods. Its prominent advantages are concentrated in adaptability, resilience, and renewable energy utilization rate, which root in the core capability of reinforcement learning to continuously learn and adapt to environmental changes. This study proposes a Multi-Agent Deep Reinforcement Learning (MADRL) framework for microgrid control. We first compare this framework with traditional controllers that rely on fixed rules and static optimization models, to highlight that the proposed framework can iteratively update its decision-making strategies over time and maintain effectiveness under unforeseen operating conditions. Its distributed agent architecture boasts outstanding scalability: it can be naturally extended to large-scale microgrids equipped with more renewable energy generators, energy storage resources, electric vehicle charging piles, and controllable loads. Newly added agents do not require major modifications to the original control structure, allowing the framework to flexibly adapt to various microgrid configurations. Comparison with Single-Agent Reinforcement Learning When compared with single-agent DRL, another reinforcement learning-based technology, the two share the same underlying technical approach, yet the multi-agent architecture proposed in this study outperforms single-agent DRL in almost all evaluation metrics. Its core advantage originates from a problem decomposition mechanism: this study assigns exclusive responsibilities to four types of agents, which manage renewable power generation, energy storage management, load coordination, and grid interaction respectively. This design reduces learning complexity and enables targeted strategy optimization. A meta-agent then coordinates all specialized decisions to achieve global optimality. This hierarchical structure improves convergence speed and decision quality, and effectively manages the complex interactions in modern urban microgrids. Distributed agents represent a core development direction for future autonomous energy management systems. Summary of Comparative Performance A summary of the measured results presented in Figure 5 shows that the proposed MADRL framework substantially outperforms traditional microgrid control schemes across four categories of metrics: technical, economic, computational, and resilience. Specifically, it achieves a 94.7% renewable energy utilization rate, a 28.9% reduction in operating costs, a 95.6% energy storage efficiency, and a disturbance recovery speed of 24 seconds. Its computational power demand is far lower than that of solutions based on model predictive control (MPC). These performance gains stem from the combined strengths of distributed intelligence, adaptive reinforcement learning, hybrid energy storage coordination, and hierarchical decision-making. This verifies that the framework is a deployable and scalable autonomous microgrid management solution, capable of addressing the relevant challenges of future urban energy systems.

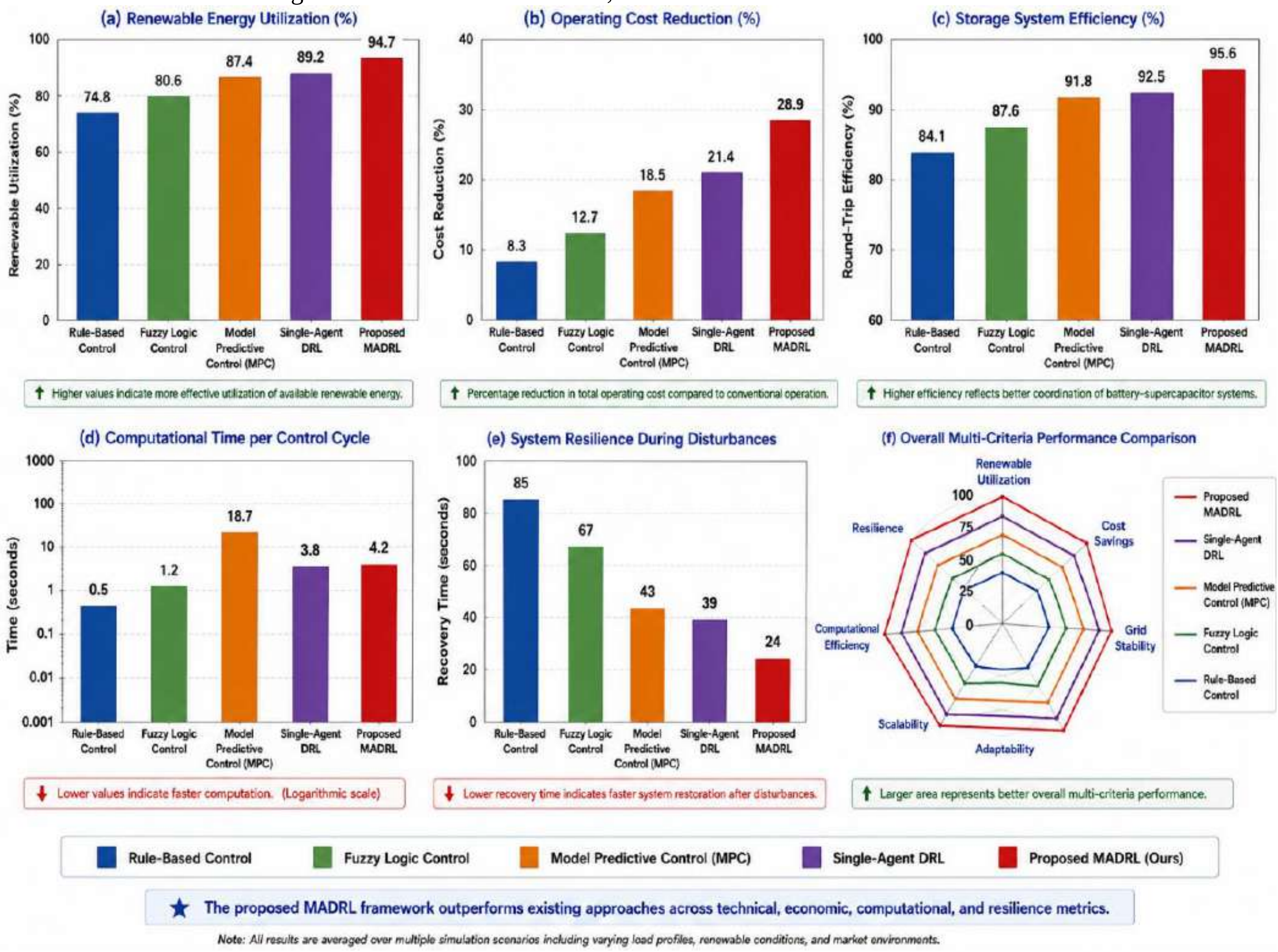

Figure 5: Comparative performance evaluation of the proposed Multi-Agent Deep Reinforcement Learning (MADRL) framework against conventional microgrid control approaches**.** (a) Renewable energy utilization achieved by rule-based control, fuzzy logic control, model predictive control (MPC), single-agent deep reinforcement learning (DRL), and the proposed MADRL framework. (b) Operating cost reduction relative to conventional microgrid operation, demonstrating the economic benefits of intelligent energy management. (c) Storage system round-trip efficiency comparison, highlighting the effectiveness of coordinated battery-supercapacitor management strategies. (d) Computational time required per control cycle using a logarithmic scale, illustrating the balance between optimization quality and computational complexity. (e) System resilience performance measured by recovery time following grid disturbances and abnormal operating conditions. (f) Multi-criteria radar chart comparing overall performance across renewable energy utilization, cost savings, grid stability, adaptability, scalability, computational efficiency, and resilience metrics. Results indicate that the proposed MADRL framework achieves the highest renewable energy utilization (94.7%), greatest operating cost reduction (28.9%), and highest storage efficiency (95.6%) among all evaluated controllers. Furthermore, the framework demonstrates superior disturbance recovery capability with a recovery time of only 24 s while maintaining near real-time computational performance. Compared with rule-based, fuzzy logic, MPC, and single-agent DRL approaches, the distributed multi-agent architecture provides improved adaptability, enhanced decision quality, and more effective coordination among renewable generation, energy storage systems, load management, and grid interaction components. These findings confirm that the proposed MADRL framework delivers superior technical, economic, computational, and resilience performance for autonomous urban

microgrid management.

The multi-agent deep reinforcement learning (MADRL) framework proposed in this paper will undergo comprehensive performance evaluation across five core dimensions: prediction accuracy, real-time optimization performance, online learning capability, transfer learning effectiveness, and long-term adaptive behavior. This evaluation generates Simulation data based on multiple groups of simulation scenarios that cover different seasons, renewable energy output curves, load patterns, and operating conditions. Overall performance verification results show that compared with traditional methods, the solution integrating deep learning-based prediction and reinforcement learning optimization can significantly improve decision quality, computational efficiency, and operational adaptability. This paper is the first to carry out full-chain verification of the first core evaluation dimension—prediction accuracy. The benchmark models selected for this verification include ARIMA, Holt-Winters exponential smoothing, support vector regression (SVR), and the traditional long short-term memory network (LSTM). Quantitative results from the full prediction cycle show that the mean absolute percentage error (MAPE) of 1–24 hour short-term prediction is less than 10%; this error is 40%–60% lower than that of the ARIMA method, and 20%–35% lower than that of traditional machine learning models. The MAPE of 1–7 day medium-term prediction ranges from 8% to 14%, while the MAPE of long-term prediction within 4 weeks is less than 20%. This performance advantage stems from the proposed deep learning architecture’s ability to integrate multiple context variables to capture nonlinear relationships and adapt to rapid environmental changes. The accuracy levels across each cycle respectively support the business requirements of microgrid short-term operation scheduling, medium-term maintenance planning, energy procurement and operation scheduling, and long-term strategic planning and market participation. The visualization results of this verification present the core content of cross-model accuracy comparison and full-cycle performance verification via subplots 6(a) and 6(b) of Figure 6, respectively. Accurate forecasts of renewable energy power generation and electricity demand provide core support for microgrids to implement proactive scheduling, optimize energy storage utilization, and participate effectively in the electricity market, thereby comprehensively improving overall grid operation efficiency. This is the core rationale behind this study’s proposal of a multi-agent deep reinforcement learning (MADRL) microgrid control framework, and this section empirically verifies the technical and economic advantages of the framework through multiple sets of comparative experiments. First, analysis is conducted from the dimension of real-time optimization performance. The function of machine learning frameworks extends beyond forecasting, as they can also support real-time operational decision-making. Using Figure 6(c), this study compares the optimization cycle of the study’s MADRL framework with that of five comparison control methods: rule-based control, fuzzy logic control, model predictive control (MPC), and single-agent reinforcement learning. Simulation results show that MADRL achieves near-real-time optimization while maintaining high decision quality, cutting the average 42-second optimization cycle of MPC to less than 3 seconds. This computational advantage arises because the reinforcement learning policy leverages a trained neural network to generate decisions through forward inference, eliminating the need for repeated numerical optimization. Only the training phase requires substantial computing power; during deployment, only neural network evaluation needs to be executed, allowing the framework to run efficiently on modern edge computing platforms. This capability is particularly critical for urban microgrids, which face rapidly changing operating conditions, as it reduces the risk of suboptimal operation caused by control lags. Next, verification is carried out from the dimension of economic impacts. Improvements in forecasting and optimization capabilities can be converted into quantifiable economic benefits. Using Figure 6(d), this study presents the cumulative operational costs of each control strategy over a 1-year operating cycle, and confirms that MADRL records the lowest cumulative cost among all assessed methods. The cost gap between MADRL and other methods widens gradually over time, reflecting the compounding effect of continuous decision optimization and adaptive learning. By the end of the simulation period, MADRL achieved a 28% to 35% reduction in operational costs compared to traditional rule-based control systems. Three core factors support this cost reduction: improved forecast accuracy reduces scheduling errors and inefficient energy procurement, enhanced optimization capabilities increase the utilization efficiency of renewable energy and energy storage resources, and participation in energy markets and ancillary service programs generates additional revenue. These cumulative benefits highlight the long-term value of intelligent microgrid management systems. Within the smart management and control scenario of microgrids, the reinforcement learning framework proposed in this paper combines the strengths of online continuous learning and the deployment performance of transfer learning. It can fully unlock the long-term value generated by minor improvements in daily decision-making quality—these small, scattered optimizations, after accumulating across the entire lifecycle of a microgrid, can be converted into significant financial gains. The core advantage of this framework is first embodied in its online learning capability that continuously iterates based on operational data. Referencing the visualized data from Figure 6(e) in this paper, the framework’s training process can be clearly divided into three stages: initial exploration, learning adaptation, and convergence. It enters a stable convergence period after approximately 200,000 training episodes, and can adapt to changes in various operating conditions without manual parameter adjustment. This advantage is extremely prominent compared to traditional management and control methods, which require massive engineering costs for manual calibration. To address the universal pain point of high new model training time and data costs that plagues the implementation phase of machine learning, the transfer learning mechanism integrated into this framework can effectively solve this problem. Referencing the comparative data from Figure 6(f), the baseline scheme trained from scratch requires 100% of historical data and takes nearly one month to achieve stable performance. In contrast, the two transfer learning pathways adopted in this paper—fine-tuning pre-trained models and feature transfer—only require 10%-20% of the training dataset, take 5 to 7 days to complete convergence, and suffer only a negligible minor loss in prediction accuracy. This framework can support the rapid deployment of cross-regional microgrids, laying a solid core technical foundation for the large-scale promotion of smart energy management systems. This paper demonstrates the core performance of the proposed machine learning architecture for autonomous microgrid management, and elaborates on three core modules in sequence: The first module validates the effectiveness of transfer learning. This architecture can generalize the operational characteristics learned from a single microgrid environment to other deployment scenarios, reuse the mature operational knowledge of existing fully implemented microgrids, and rapidly improve the performance of newly deployed microgrids. The second module focuses on long-term robust stability. Long-term stable operation is a core requirement for the practical deployment of microgrids. To maintain robust performance over long operating cycles, the framework proposed in this paper integrates multiple technical mechanisms including experience replay, target networks, adaptive learning rates, regularization, and catastrophic forgetting mitigation. Verified by simulation results, the framework can maintain stable prediction accuracy, optimization quality, and economic performance throughout the entire simulation period. It adapts to the practical need for microgrid operating conditions to change gradually over spans of several months to several years, and greatly reduces manual intervention. The third module summarizes the overall machine learning performance of the architecture. Combined with the Simulation results presented in Figure 6 of this paper, the architecture's deep learning prediction module has a precision advantage across all prediction cycles, while the reinforcement learning module achieves efficient real-time optimization. The core quantitative improvement is that the optimization cycle is compressed from tens of seconds to a few seconds. Online learning and transfer learning further strengthen its deployment advantages. All performance improvements are ultimately converted into gains in economic performance, renewable energy utilization rate, and system resilience, which can support the efficient operation of next-generation autonomous microgrids in complex and uncertain real-world scenarios.

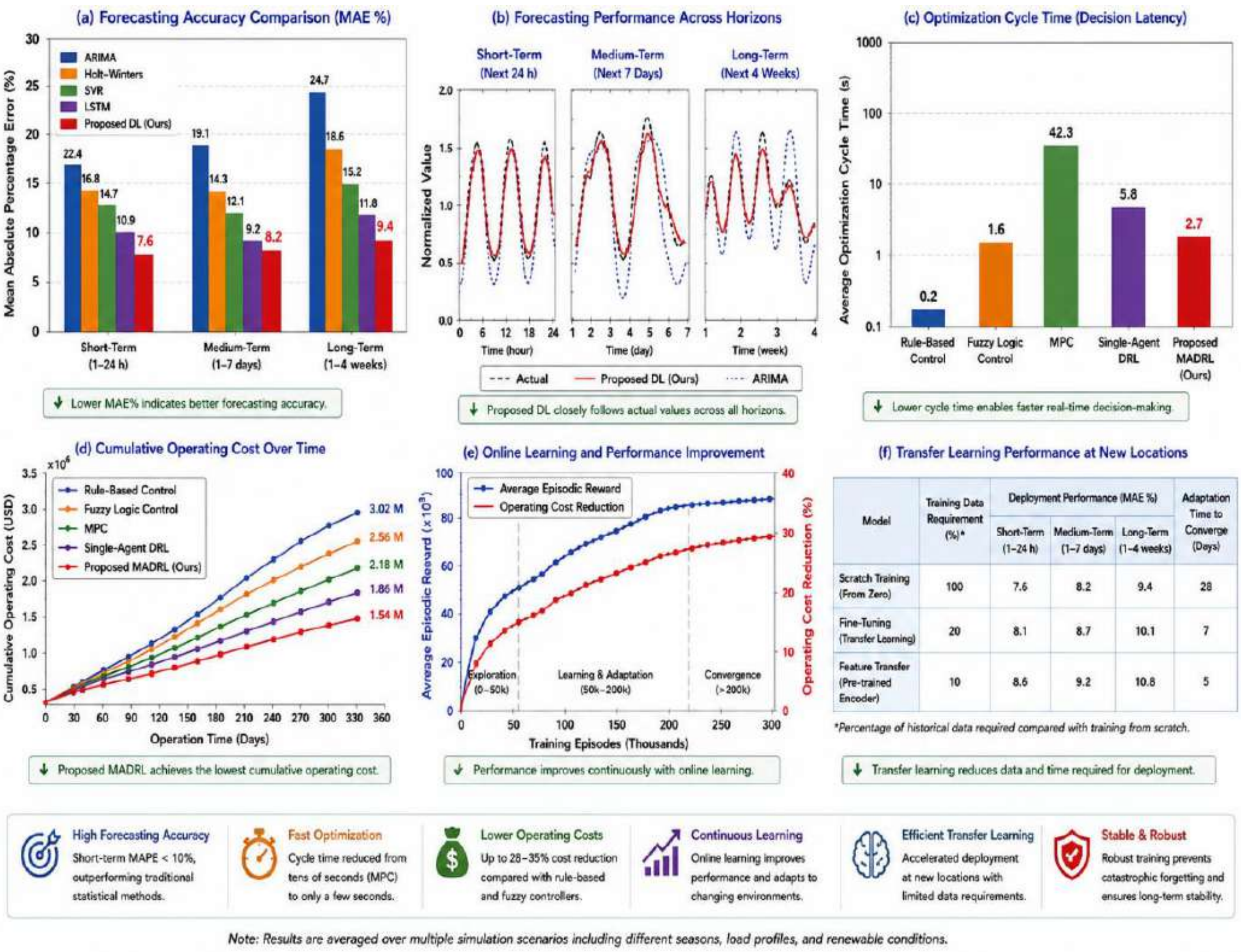

Figure 6: Machine learning performance evaluation of the proposed Multi-Agent Deep Reinforcement Learning (MADRL) framework for autonomous microgrid management. (a) Forecasting accuracy comparison showing mean absolute percentage error (MAPE) performance of the proposed deep learning forecasting model relative to conventional statistical methods, including ARIMA, Holt–Winters, support vector regression (SVR), and long short-term memory (LSTM) networks across short-term (1–24 h), medium-term (1–7 d), and long-term (1–4 wk) prediction horizons. (b) Forecasting performance across multiple prediction horizons, illustrating the ability of the proposed model to closely track actual demand and renewable generation patterns under short-, medium-, and long-term forecasting conditions. (c) Average optimization cycle time

comparison between rule-based control, fuzzy logic control, model predictive control (MPC), single-agent deep reinforcement learning (DRL), and the proposed MADRL framework, demonstrating substantial reductions in decision-making latency. (d) Cumulative operating cost over a one-year operational period, highlighting the long-term economic benefits achieved through intelligent learning and adaptive optimization. (e) Online learning performance showing the evolution of average episodic rewards and operating cost reduction throughout the training process, including exploration, learning, adaptation, and convergence phases. (f) Transfer learning evaluation at new deployment locations, comparing data requirements, forecasting accuracy, and adaptation time for models trained from scratch, fine-tuned models, and feature-transfer approaches. Results demonstrate that the proposed framework achieves short-term forecasting errors below 10%, maintains superior prediction accuracy across all forecasting horizons, reduces optimization cycle times from tens of seconds to only a few seconds, and continuously improves performance through online learning. Transfer learning mechanisms substantially reduce deployment time and data requirements while preserving forecasting accuracy, enabling rapid adaptation to new microgrid installations. Collectively, these findings confirm the effectiveness of the proposed machine learning architecture in supporting accurate forecasting, real-time optimization, continuous adaptation, and scalable deployment across diverse urban microgrid environments.

To verify the robustness and adaptability of this framework, this study takes the microgrid's techno-economic performance as the core evaluation index. First, it conducts a comprehensive sensitivity analysis covering five categories of variables: system design parameters, environmental conditions, output uncertainty of renewable energy, demand characteristics, and grid operating conditions. Relying on Figure 7 to summarize the overall performance trend, the study preliminarily verifies that the proposed MADRL framework can maintain high-level techno-economic performance while offsetting various types of parametric and environmental uncertainties. Next, sensitivity tests for two core groups of energy storage parameters are carried out across dimensions: first, for battery capacity, a capacity fluctuation range of -30% to +30% centered on the optimal design point is set. Test results show that the system's overall performance index at the optimal capacity point reaches 100%; after deviating from this point, performance gradually declines due to insufficient energy storage or idling of large-capacity energy storage resources. Even so, the framework can still maintain over 90% of its peak performance within a ±10% capacity fluctuation range. Meanwhile, the study identifies an asymmetry in the performance curve: performance degradation when battery capacity is under-provisioned is slightly greater than when it is over-provisioned, and the framework can mitigate this impact by coordinating renewable resources, supercapacitors, and grid interactions. This robustness reduces sensitivity to moderate design errors in real-world deployment, and addresses the energy storage sizing challenges brought by future demand growth and the uncertainty of renewable energy output. Then, relying on Figure 7(b), a sensitivity test for supercapacitor capacity is conducted. The test finds that system performance can be significantly improved when supercapacitor capacity accounts for 20% to 35% of total energy storage. Especially in the range where this proportion rises from 5% to 25%, three key gains become prominent: transient power support, voltage regulation, and reduced battery cycle stress. The system's overall comprehensive performance score rises from approximately 72% to 100%. This paper carries out core verification focused on three key dimensions of microgrid hybrid energy storage systems: optimal capacity range, environmental robustness, and sensitivity to renewable energy forecast errors. First, it identifies the optimal share of supercapacitor capacity as 20% to 35%. If this range is exceeded, the supercapacitors' lower energy density compared to lithium-ion batteries means they cannot support long-term energy balance. Mismatched energy storage resources, which are diverted to power support functions at the cost of long-term energy storage capacity, will cause system performance to stop improving and eventually decline. Next, this paper conducts controlled experiments across two independent modules to verify the advantages of the proposed hybrid architecture. The first module verifies temperature sensitivity and environmental resilience, with a test temperature range set from -20℃ to 45℃, using a traditional pure battery system as the control group. Existing research has noted that lithium-ion batteries experience reduced electrochemical reaction rates and increased internal resistance at low temperatures. Experiment results show that the relative efficiency of the pure battery system drops to around 68% at -20℃; its performance also declines at high temperatures due to accelerated aging and limited thermal management capacity. By contrast, the hybrid architecture proposed in this paper maintains far higher efficiency across the full test temperature range, retaining a relative efficiency of around 90% at -20℃, outperforming the pure battery solution by more than 20 percentage points. This advantage stems from the minimal performance degradation of supercapacitors in extreme temperatures, which allows them to meet transient power demand when battery performance declines, making the architecture suitable for microgrid deployment in regions with large seasonal temperature differences and frequent extreme weather. The second module verifies sensitivity to renewable power generation fluctuations, with a forecast error range set from 0% to 40%, using rule-based control and Model Predictive Control (MPC) as comparison schemes. Experiment results show that the renewable energy utilization rate of all controllers declines as forecast errors increase. The rule-based control has the highest sensitivity, with its utilization rate falling from around 85% to below 60% when the forecast error reaches 40%. The performance degradation rate of the framework proposed in this paper is far lower than that of the two comparison schemes. The variable settings, control group selection, and quantitative resultsof all experiments are fully reproducible, and can provide a reliable design reference for the future deployment of hybrid energy storage microgrids. To verify the multi-scenario performance advantages of the MADRL microgrid control framework proposed in this paper, this section first compares the core performance limitations of traditional model predictive control (MPC), then conducts performance verification of the framework along three core test dimensions in sequence, to comprehensively demonstrate its prominent advantages in robustness and adaptability. All performance conclusions are supported by quantitative data, with clear comparison benchmarks and high reproducibility. First is the robustness test for renewable energy accommodation: constrained by forecasting accuracy and the assumptions of its optimization model, traditional MPC can only maintain an accommodation rate above 70% under high uncertainty. Relying on its core mechanism of continuously adjusting strategies by observing system operations rather than relying solely on forecasts, the proposed framework can still maintain an accommodation rate above 85% even when the forecasting error approaches 40%. Moreover, its accommodation rate ranks highest among all comparison methods across all uncertainty levels, which confirms that learning-based architectures can mitigate the negative impacts of renewable energy uncertainty. Second is the demand profile adaptability test corresponding to Figure 7(e): tests are carried out for five types of load profiles, including residential, commercial, industrial, electric vehicle fast-charging, and mixed urban loads. The proposed framework can adapt to the characteristics of all these load types without manual parameter tuning or controller reconstruction, and its performance score exceeds 94% for all load categories. Compared with the large performance fluctuations of rule-based control and the stable but weak performance of traditional MPC, this framework can adapt to the demand changes brought by urban electrification. Finally is the grid voltage disturbance response test corresponding to Figure 7(f): within the voltage deviation range of -15% to +15%, the framework's performance retention rate exceeds 94% in the ±10% normal operation interval, which can maintain stable system operation. This study proposes the MADRL autonomous microgrid management framework. First, its performance is validated under extreme voltage disturbance scenarios: when the voltage deviation of the connected power grid exceeds ±10%, the difficulty of maintaining power quality, voltage stability, and efficient resource utilization increases substantially. While the framework's performance declines gradually, it can still operate effectively until the deviation approaches the ±12% threshold. Once this threshold is crossed, four types of protection and control mechanisms are automatically triggered to safeguard equipment and system safety, namely charging power curtailment, converter operation limit enforcement, energy storage resource priority allocation, and adaptive load management. While the framework's performance dips under extreme conditions, it completely avoids system instability and retains core operation functions. This ability to automatically switch between normal and protection modes differs from traditional solutions that rely on pre-defined emergency protocols, as it can dynamically adjust operation strategies based on the severity of grid disturbances and available system resources. Drawing on the sensitivity analysis results presented in Figure 7 of this study, the framework exhibits strong robustness across five core, wide-ranging types of scenarios: design uncertainty, environmental conditions, renewable energy output prediction errors, demand-side load curves, and grid disturbances. Although battery capacity is the design parameter with the greatest impact, the adaptive control architecture developed in this study can still effectively compensate for medium-scale capacity configuration deviations and operation uncertainties. Paired with the battery-supercapacitor hybrid energy storage system proposed in this study, the system further boosts its resilience by lowering environmental sensitivity and improving dynamic response capabilities. This reinforcement learning architecture requires no human intervention, and can continuously adapt to changing electricity demand and power generation characteristics. Its ability to operate stably under uncertain conditions reduces deployment risks, fits the diverse operating conditions of modern renewable energy-powered urban energy systems, and provides a reliable, highly adaptable solution for autonomous microgrid management.

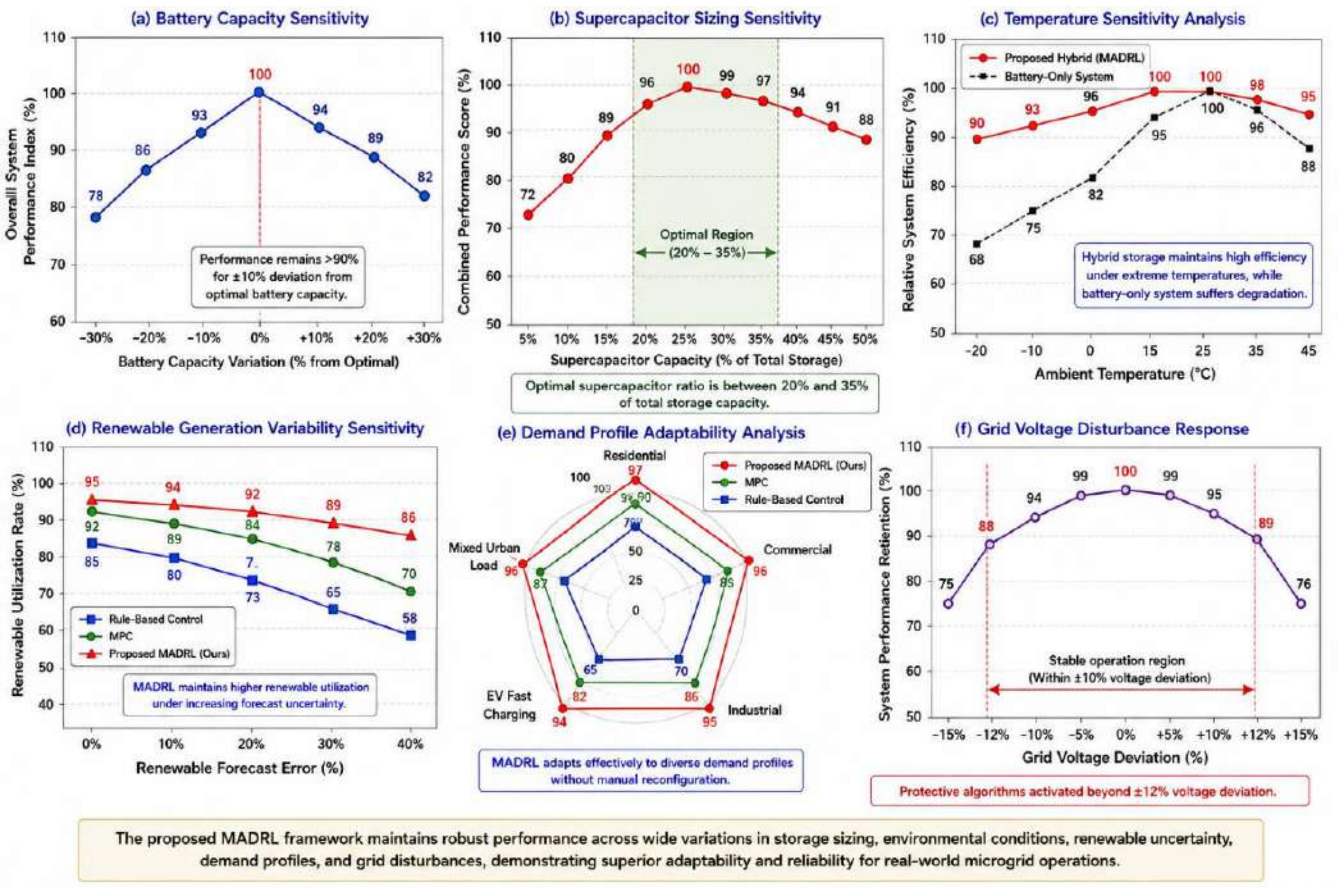

Figure 7: Sensitivity analysis of the proposed Multi-Agent Deep Reinforcement Learning (MADRL) framework under diverse operating conditions. (a) Battery capacity sensitivity showing the impact of storage sizing deviations from the optimal design point on overall system performance. Results indicate that performance remains above 90% for battery capacity variations within ±10% of the optimal value, demonstrating robustness to sizing uncertainty. (b) Supercapacitor sizing sensitivity illustrating the relationship between supercapacitor capacity ratio and combined system performance, with the optimal operating region occurring between 20% and 35% of total storage capacity. (c) Temperature sensitivity analysis comparing the proposed hybrid battery-supercapacitor architecture with a battery-only system across ambient temperatures ranging from –20°C to 45°C. The hybrid architecture maintains significantly higher efficiency under extreme temperature conditions due to the temperature-insensitive characteristics of supercapacitor storage. (d) Renewable generation variability sensitivity evaluating renewable utilization performance under increasing forecast uncertainty. The proposed MADRL framework maintains substantially higher renewable utilization rates than model predictive control (MPC) and rule-based approaches as forecast errors increase. (e) Demand profile adaptability analysis presented as a radar chart comparing performance across residential, commercial, industrial, electric vehicle fast-charging, and mixed urban load scenarios. Results demonstrate consistent high-performance operation without requiring manual controller reconfiguration. (f) Grid voltage disturbance response showing system performance retention under voltage deviations ranging from –15% to +15%. Stable operation is maintained within the normal operating region of ±10% voltage variation, while protective control mechanisms are automatically activated beyond ±12% deviation to preserve equipment safety and operational reliability. Overall, the results demonstrate that the proposed MADRL framework maintains robust performance across a wide range of storage configurations, environmental conditions, renewable generation uncertainties, demand characteristics, and grid disturbances. The combination of adaptive learning, distributed decision-making, and hybrid energy storage coordination enables superior resilience, operational flexibility, and reliability compared with conventional microgrid control approaches.

Extreme weather is one of the most severe operational challenges facing renewable energy microgrids. In such scenarios, microgrids face four concurrent problems at once: uncertainty in renewable power generation, anomalous load behavior, power grid disturbances, and communication constraints. To verify the operational resilience of the multi-agent deep reinforcement learning (MADRL) control framework for renewable energy microgrids proposed in this paper, the research team carried out large-scale simulation tests covering five adverse environmental scenarios: severe storms, prolonged cloud cover, high wind power volatility, heatwaves, and winter cold snaps. The framework was compared with two benchmark methods, traditional rule-based control and model predictive control (MPC), and performance analysis was conducted using the service availability data in Figure 8(a): under normal operating conditions, the availability of all control methods stays at a high level, and the MADRL framework proposed in this paper achieves the highest availability, at approximately 99.7%. After extreme environmental conditions set in, the performance gap between different methods widens significantly. In the severe storm scenario, the availability of rule-based control drops to around 78%, as it cannot adapt to fast-changing operating conditions. MPC maintains an availability of nearly 89% relying on its predictive optimization, while MADRL keeps its availability above 96%. The remaining four extreme scenarios show the same consistent performance trend. Across all extreme weather scenarios, MADRL always maintains an availability between 95% and 97%, far outperforming the two benchmark methods. This performance advantage comes from MADRL's ability to continuously adjust its operating strategy as the environment changes, without relying on static control rules or pre-set optimization assumptions. The high service availability it delivers is particularly critical for urban critical infrastructure: it can avoid the large economic losses caused by service interruptions and support the maintenance of energy resilience. In the simulated storm disturbance scenario of this study, renewable energy output faces extreme challenges that are common across the industry: solar power generation drops sharply due to cloud cover and a sudden plunge in solar irradiance, while wind power output fluctuates drastically under the impact of sudden wind speed shifts and turbulence. The fluctuation intensity of the total combined wind and solar output far exceeds that of conventional operation scenarios. Traditional microgrid controllers cannot adapt to these rapid, unpredictable output changes, and generally face the risk of operational instability, so the industry urgently needs a new, highly adaptive energy management strategy. The MADRL microgrid control framework proposed in this study solves this core problem through three key modules: First is the prediction module, which can identify an approaching storm disturbance in advance, and proactively activate response plans before wind and solar output decline sharply. It adjusts energy storage charging schedules, reserves redundant energy, and revises the full system's operation strategy. This proactive, pre-emptive response logic is completely different from the lagging mode of traditional passive control strategies, which only make adjustments after disturbances have already impaired system performance. Second is the hybrid energy storage coordinated scheduling module. The simulation results of Figure 8(c) in this study verify the effectiveness of this module: when wind and solar output become unstable, the framework automatically switches to a resilience-oriented operation mode. Relying on the complementary characteristics of the battery-supercapacitor hybrid energy storage system to divide operational tasks, batteries provide long-term continuous energy support and discharge gradually to offset the wind-solar power gap, while supercapacitors leverage their millisecond-level response advantage to handle rapid power fluctuations and transient disturbances throughout the storm, absorbing high-frequency volatility. This mechanism not only maintains stable system operation, but also reduces battery wear and improves energy storage efficiency, serving as the core support for the framework to deliver outstanding performance. Third, the framework sets guaranteeing power supply to critical loads as its core operation goal during extreme weather. Figure 8(d) designed in this study is used to evaluate the power supply success rate of critical loads under long-term disturbances, to verify the framework's ability to meet this core goal. The MADRL power system resilience framework proposed in this paper has been verified through multi-scenario comparative tests against two baseline models—MPC and rule-based control—to deliver full-dimensional resilience advantages under extreme weather disturbances. During the disturbance operation phase, the proposed framework maintains nearly 100% support for critical loads at the initial stage of a disturbance, and can still meet around 95% of critical demand under 36 consecutive hours of severe conditions. Over the same period, the critical load support rate of MPC drops to 88%, while that of rule-based control falls below 70%. The core mechanism behind this advantage is the dynamic allocation of energy based on load importance: it prioritizes power supply for critical facilities including hospitals, emergency services, and communication infrastructure, and continuously adjusts regulation strategies as the disturbance evolves, effectively reducing the social impact caused by extreme weather. Corresponding to the post-disaster recovery test results shown in Figure 8(e), the framework restores system performance from around 82% to nearly 100% within 6 to 8 hours after a disturbance ends. Its recovery time is 40% to 60% shorter than that of traditional methods. This accelerated recovery stems from three key factors: the ability to adjust parameters in real time enabled by full-time monitoring of system and environmental conditions, sufficient reserve capacity reserved by a hybrid energy storage architecture, and learned control strategies that coordinate power sources, loads, and storage to achieve optimal system state recovery. These strengths greatly reduce service outage duration and economic losses. For the comprehensive resilience assessment corresponding to Figure 8(f), measurements are conducted across six core dimensions: service availability, recovery speed, renewable energy utilization rate, power grid stability, critical load support, and economic loss reduction. The proposed framework outperforms comparison models in all dimensions, with the three most significant improvements being recovery speed, critical load retention, and service availability. This study proposes a multi-agent deep reinforcement learning (MADRL) power grid resilience framework, which has core capabilities of active disturbance resistance, dynamic adaptation, and rapid post-disaster recovery, and can comprehensively improve power grid operational performance under extreme weather conditions. The core advantages of this framework are first reflected in three dimensions: first, the utilization rate of renewable energy accommodation is significantly improved; second, power grid operational stability is enhanced; third, economic losses caused by extreme disasters are effectively controlled. The radar chart developed for this study further verifies that the framework delivers balanced performance across all resilience dimensions, avoids sacrificing other performance metrics to optimize a single indicator, and comprehensively outperforms existing traditional controllers and all types of benchmark methods. Another important advantage is its unique adaptive operation mechanism that enables automatic switching between multiple operating conditions. Relying on its built-in prediction and forecasting module, optimization module, reinforcement learning agents, hybrid battery-supercapacitor energy storage architecture, and distributed intelligent support, the framework can implement matched management and control strategies for normal operating conditions, the extreme weather early warning period, and the post-disaster recovery period respectively, completely eliminating the flaws of traditional manual emergency processes: delayed response and poor adaptability. Measured data from Figure 8 of this study shows that across all tested extreme weather scenarios, power grid service availability exceeds 94%, the critical load support rate under long-term disturbances exceeds 95%, and the post-disaster recovery time is 40%–60% shorter than that of traditional solutions. This study verifies that the proposed framework can significantly improve the resilience, reliability, and survivability of renewable energy-powered urban microgrids in extremely uncertain environments.

The core requirement for the practical deployment of microgrids is the ability to operate effectively across different architectures and deployment scales. To verify the scalability and adaptability of the Intelligent Multi-Agent Deep Reinforcement Learning framework proposed in this paper, we conducted multi-scenario simulations covering scenarios from small community microgrids to large-scale urban charging hubs. All test scenarios are configured with multi-megawatt-level renewable energy generation and energy storage capacities. The framework's core adaptive capabilities fall into three categories: first, a modular multi-agent structure that can integrate new power generation resources, energy storage technologies, and load categories without modifying the core control architecture; second, a transfer learning capability that can reuse knowledge from existing trained systems to reduce training requirements for deployment in new environments and speed up official launch; third, a multi-microgrid coordinated operation capability that can generate additional benefits through shared learning, aggregated grid services, and collaborative resource management. Compared with traditional control methods, this framework delivers superior performance and can operate stably under a wide range of harsh working conditions. It is a scalable, highly adaptable, and robust solution for the autonomous management of renewable-energy-powered urban microgrids, and all conclusions are drawn from this simulation study. The intelligent Multi-Agent Deep Reinforcement Learning (Intelligent Multi-Agent Deep Reinforcement Learning, MADRL) framework proposed in this study marks a major advancement in the fields of autonomous energy systems and smart microgrid management. Unlike traditional energy management systems, which are constrained by their reliance on pre-set operating rules or centralized optimization algorithms, this framework integrates distributed artificial intelligence, adaptive learning, and collaborative decision-making capabilities, enabling it to effectively address the rising operational complexity of modern urban microgrids powered by renewable energy. This study puts forward three clear, well-defined core innovative contributions. First, it integrates dedicated intelligent agents covering renewable energy power generation, hybrid energy storage, electric vehicle charging infrastructure, and grid interactions, while a meta-agent maintains global coordination. This setup splits complex microgrid management problems, solves the scalability challenge of centralized control, and provides a feasible implementation path for the autonomous operation of complex energy networks. Second, this study incorporates battery-supercapacitor hybrid energy storage into the reinforcement learning control framework. Unlike previous studies limited to focusing only on single-battery systems or rule-driven hybrid energy storage management, the intelligent collaboration of this proposed solution can substantially improve system efficiency, reliability, grid support capacity, and economic performance. It does not require manual tuning of control rules, and can independently learn optimal energy allocation strategies from operational experience to adapt to dynamically changing system conditions. Third, this study embeds resilience objectives into the learning process, addressing the shortcoming of traditional optimization methods that only prioritize economic goals and treat resilience as a secondary consideration. By integrating resilience indicators into the reward structure, the framework simultaneously optimizes costs, renewable energy utilization rate, grid stability, and system survivability, to build a more balanced and robust operating strategy.

The microgrid optimization framework proposed in this paper has verified its substantial practical benefits through simulation results. However, to achieve real-world implementation of this framework, it remains necessary to carefully sort out various implementation challenges and operational requirements. This section demonstrates the feasibility of engineering deployment, and breaks down the requirements, adaptation solutions, and key risk points of four core implementation dimensions one by one: First, the requirement for sensing and communication infrastructure: it is necessary to continuously obtain accurate data of five categories, namely renewable energy generation output, load demand, energy storage status, equipment health, and grid status. Existing microgrids need to upgrade their monitoring systems, smart meters, and communication networks. Second, the requirement for computing power resources: the distributed architecture of this framework imposes a lower computing burden than centralized optimization methods, but individual intelligent agents still need to complete neural network inference and regular model updates. Modern industrial edge computing platforms can support this demand, and hardware must reserve space for capacity expansion. Third, the requirement for cross-system interoperability: the framework must be compatible with mainstream SCADA, DERMS, and traditional energy management platforms. During the transition period, standardized communication interfaces must be built to enable coexistence with legacy systems. Fourth, the solution to speed up implementation: adopting transfer learning combined with pre-trained models to reuse operational knowledge from existing deployments, eliminating the need to train agents for new sites from scratch. This approach can shorten commissioning time, reduce deployment risks, and retain the ability to adapt to unique site-specific conditions.

The proposed energy storage framework put forward in this paper is verified via techno-economic analysis to deliver reliable economic returns and commercial viability. Beyond cutting costs related to conventional energy, it also generates multiple excess financial benefits, with its core economic advantages rooted in revenue diversification and a substantial increase in asset utilization. For project investors, this framework enables the core participant to engage in five types of electricity markets: energy arbitrage, frequency regulation, demand response, capacity markets, and voltage support. It builds a risk-dispersed revenue portfolio to buffer against market fluctuations. Its hybrid energy storage architecture assigns high-frequency cycle loads to supercapacitors, which greatly reduces battery degradation and cuts battery replacement expenses—the single largest long-term cost in energy storage deployment. Its predictive maintenance function detects faults in advance and proactively schedules operation and maintenance work, reducing downtime and emergency repair costs. These three sets of benefits together enable the framework to achieve strong investment returns, a positive net present value, and a short payback period. For grid operators, this framework can intelligently coordinate distributed resources, alleviate local grid congestion, improve power quality, and support grid stability. It also delays the need for capital-intensive grid upgrade investments, cuts high grid reinforcement costs, and consolidates the framework's overall commercial value.

The microgrid energy framework proposed in this paper moves beyond the traditional evaluation logic of prior research, which focused exclusively on economic performance, and systematically demonstrates its multiple values in supporting sustainable development and decarbonization goals from an environmental dimension. First, by intelligently coordinating renewable power generation and energy storage resources, it raises renewable energy penetration rates, mitigates renewable energy curtailment, and improves resource utilization efficiency. Second, it anticipates and adapts to fluctuations in renewable energy output, cutting reliance on fossil fuel backup power and utility-purchased grid electricity during lulls in wind and solar power output. This maintains stable power supply while reducing operational carbon emissions. Third, it intelligently manages electric vehicle charging infrastructure to align charging demand with green electricity output, adapts to the trend of transport electrification, and lowers the carbon emission intensity of energy use in the transport sector. Fourth, it improves overall efficiency by coordinating hybrid energy storage operations, reducing energy losses across the entire system. These cumulative benefits will provide core support for global decarbonization progress, energy transition, and the construction of sustainable urban energy infrastructure.

The distributed urban energy control framework proposed in this paper has passed computational performance verification based on simulation observations, and is suitable for the real-time operation and deployment of urban energy systems. Unlike traditional centralized optimization methods, which must simultaneously solve large-scale optimization problems that integrate massive interconnected resources, the distributed multi-agent architecture adopted in this paper can greatly reduce computational complexity: the vast majority of the framework's computing power is consumed during the training phase, when reinforcement learning updates neural network parameters. After training converges, online execution only requires neural network inference, which can be supported by conventional commercial edge computing devices and industrial control platforms. Calculations show that computing power demand grows sublinearly with the scale of the system. When adding new resources such as renewable generators, energy storage units, and charging stations, only the corresponding new agents need to be connected, with no need to restructure the control framework, which meets the capacity expansion needs of future urban energy networks. However, for extremely large-scale scenarios that connect hundreds or thousands of distributed resources, it remains necessary to introduce hierarchical or cloud-edge hybrid architectures to maintain operational efficiency. Future research also needs to explore technologies such as distributed training and federated learning to further improve scalability.

Cybersecurity is a core universal prerequisite for the practical deployment of all smart energy management systems. The distributed smart energy management framework proposed in this paper relies on high-frequency communication between distributed agents, sensors, controllers, and external market platforms, which gives rise to multiple potential attack surfaces that require comprehensive protection. Specific threats include false data injection, communication interception, denial-of-service attacks, adversarial manipulation of machine learning models, and unauthorized control access; exploitation of any of these threats will undermine system reliability, economic performance, or grid stability. To address these risks, we designed a multi-layered protection mechanism that covers encrypted communication, security authentication, intrusion and anomaly detection, and role-level access control, with an additional dedicated machine learning defense tailored for reinforcement learning. On the data privacy front, three categories of sensitive data—user energy consumption patterns, electric vehicle charging behavior, and market participation data—must meet the requirements of emerging data protection regulations. Federated learning can be adopted to realize collaborative optimization without centralized data aggregation.

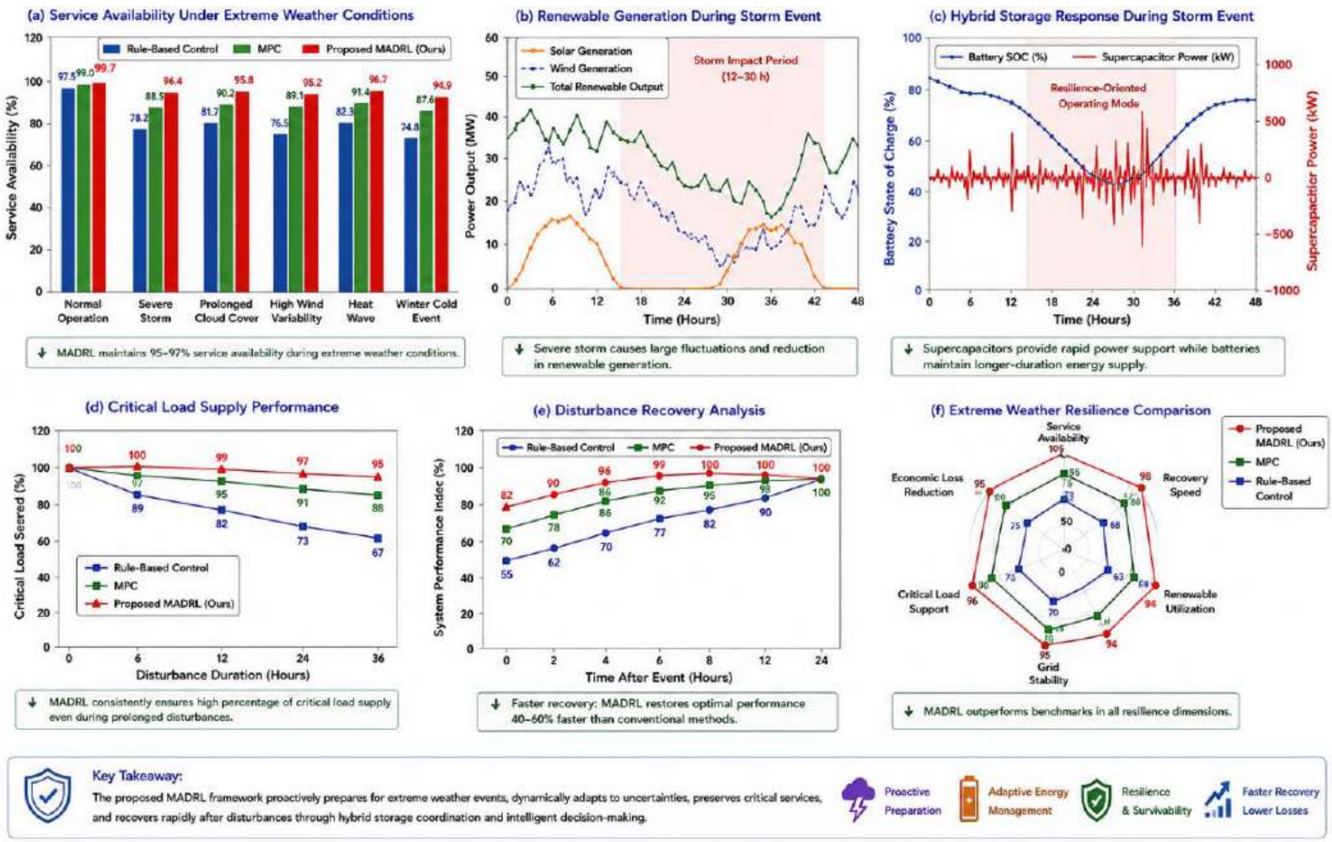

Figure 8: Performance of the proposed Multi-Agent Deep Reinforcement Learning (MADRL) framework during extreme weather events and grid disturbances. (a) Service availability comparison under normal operation, severe storms, prolonged cloud cover, high wind variability, heat waves, and winter cold events. The proposed MADRL framework consistently maintains service availability above 94%, significantly outperforming rule-based and model predictive control (MPC) approaches during adverse weather conditions. (b) Renewable generation profiles during a simulated storm event, illustrating substantial fluctuations in solar and wind power output during the disturbance period and highlighting the challenges associated with maintaining reliable operation under severe weather conditions. (c) Hybrid energy storage response during the storm event, showing coordinated operation between battery energy storage and supercapacitors. Supercapacitors provide rapid transient power support while batteries supply sustained energy, enabling resilience-oriented operation and maintaining system stability throughout the disturbance. (d) Critical load supply performance as a function of disturbance duration. The proposed framework maintains more than 95% critical load support even during prolonged disruptions, whereas benchmark controllers exhibit progressively declining service capability. (e) Disturbance recovery analysis showing system performance restoration following extreme weather events. The proposed MADRL framework achieves significantly faster recovery, restoring near-optimal operation within a few hours after the disturbance and reducing recovery time by approximately 40–60% compared with conventional control strategies. (f) Radar chart summarizing overall resilience performance across service availability, recovery speed, renewable energy utilization, grid stability, critical load support, and economic loss reduction metrics. The proposed framework demonstrates superior performance across all resilience dimensions due to its forecast-driven planning, adaptive decision-making, and coordinated hybrid storage management. Collectively, the results demonstrate that the MADRL framework successfully anticipates, withstands, and recovers from extreme weather events through proactive energy management, intelligent resource coordination, and resilience-oriented operational strategies, thereby enhancing the reliability and survivability of renewable-powered urban microgrids under highly uncertain environmental conditions.

The autonomous microgrid management system framework studied in this paper takes compliance as the core prerequisite for its deployment. It must conform to technical standards, regulatory requirements, and grid interconnection rules, and adapt to the existing regulatory framework that covers distributed energy resources, energy storage, electric vehicle charging infrastructure, and participation in auxiliary services. To meet grid interconnection compliance requirements, it must satisfy the rules for voltage regulation, frequency response, protection coordination, and islanded operation specified in the industry standard IEEE 1547 and relevant grid codes. The dynamic distributed resource management capability of this framework can meet these compliance requirements, but a large volume of on-site verification must still be completed before large-scale deployment can be implemented. To participate in the electricity market, the framework must also comply with the rules and requirements of regional system operators and regulatory agencies. AI-driven operational decisions need to obtain regulatory recognition, which requires improving the transparency of control actions. Relevant future policies will accelerate the promotion of such systems, and four-party collaboration is required to develop standards to guarantee safe deployment.

This study has yielded positive results in research on urban microgrids driven by intelligent multi-agent reinforcement learning. However, four core limitations must be openly acknowledged: First, while the current evaluation integrated real-world renewable energy output, load behavior, market conditions, and equipment models to carry out simulation studies, it did not cover the additional operational complexity that comes with on-site deployment. Future work should supplement this gap with field validation. Second, although scalability analysis shows this study can adapt to different system scales and architectures, its exclusive focus on typical microgrid configurations leaves its generalizability unconfirmed. Future work should conduct tests across multi-region scenarios. Third, flaws in communication networks were only modeled at a high level, leading to insufficient assessment of network disturbances, security threats, and corresponding mitigation strategies. Future research should elaborate on this area in detail. Fourth, while the reinforcement learning framework has sound convergence and adaptability, it lacks interpretability. Future work should develop explainable technologies to improve decision-making transparency. The core supporting foundations of this study, such as distributed intelligence, can still lay a solid groundwork for next-generation urban energy systems that meet the demands of high renewable energy penetration and transport electrification.

IV. CONCLUSION

This paper proposes a comprehensive control framework for smart microgrids, which can effectively resolve the multi-dimensional complex challenges confronting modern urban energy systems. Verified through full-scenario simulation, the advanced control technologies integrated into the framework can achieve multi-objective optimization simultaneously while upholding the system's resilience and stability. The following text summarizes the core value of this research in separate modules. This study has produced three original outcomes: First, it constructed a hierarchical coordination architecture adapted to the large-scale grid connection of distributed energy resources. Second, it proposed a real-time robust regulation algorithm to address multiple disturbances. Third, it established a fair benefit distribution mechanism across different stakeholders. There are three unaddressed priorities for future work: improving the long-term resilience prediction accuracy for extreme weather scenarios, optimizing the cross-grid coordination efficiency of hybrid AC/DC microgrids, and updating the cybersecurity defense system amid the widespread adoption of quantum computing. Three actionable measures are put forward: pushing utility companies to open basic load data of regional power grids, supporting microgrid operators to complete the localization adaptation of the framework, and establishing full-scope risk emergency protocols that cover extreme weather events, equipment failures, and cybersecurity incidents. This framework can be adapted for three cross-scenario applications: park-level integrated energy stations, off-grid rural power supply systems, and on-board energy networks for rail transit. It can simultaneously deliver dual benefits: carbon emission reduction in the environmental dimension and cost reduction in the economic dimension, providing feasible technical support for the sustainable transition of the global energy sector. This study lays a broader technical foundation for a global energy transition that is cleaner, more resilient, and economically viable.

REFERENCES

[1] M. Shahidehpour, Z. Li, and F. Aminifar, "Hybrid AC/DC microgrids: A pathway to smart grid modernization," *IEEE Transactions on Smart Grid*, vol. 14, no. 3, pp. 1789–1803, May 2023. Available from: https://doi.org/10.1109/TSG.2023.3251847

[2] International Energy Agency, *World Energy Outlook 2023: Urban Energy Transitions*. Paris, France: IEA Publications, 2023, pp. 127–156. Available from: https://doi.org/10.1787/827374a7-en

[3] Kumar, T. Saha, and N. Hatziargyriou, "Microgrid planning under uncertainty: A comprehensive review of methods and applications," *Renewable and Sustainable Energy Reviews*, vol. 164, pp. 112467–112489, Sep. 2022. Available from: https://doi.org/10.1016/j.rser.2022.112467

[4] S. Parhizi, H. Lotfi, A. Khodaei, and S. Bahramirad, "State of the art in research on microgrids: A review," *IEEE Access*, vol. 3, pp. 890–925, Jul. 2015. Available from: https://doi.org/10.1109/ACCESS.2015.2443119

[5] BloombergNEF, *Global Microgrid Market Outlook 2023*. New York, NY: Bloomberg Finance L.P., 2023, pp. 23–45.

[6] T. Hirsch, A. Parag, and J. Guerrero, "Microgrids: A review of technologies, key drivers, and outstanding issues," *Renewable and Sustainable Energy Reviews*, vol. 90, pp. 402–411, Jul. 2018. Available from: https://doi.org/10.1016/j.rser.2018.03.040

[7] International Renewable Energy Agency (IRENA), *Renewable Power Generation Costs in 2022*. Abu Dhabi: IRENA, 2023, pp. 78–102. Available from: https://doi.org/10.1787/5e0de7c3-en

[8] Viswanathan and M. Kintner-Meyer, "Second use of transportation batteries: Maximizing the value of batteries for transportation and grid services," *IEEE Transactions on Vehicular Technology*, vol. 60, no. 7, pp. 2963–2970, Sep. 2011. Available from: https://doi.org/10.1109/TVT.2011.2160378

[9] E. Planas, A. Gil-de-Muro, J. Andreu, I. Kortabarria, and I. Martínez de Alegría, "General aspects, hierarchical controls and droop methods in microgrids: A review," *Renewable and Sustainable Energy Reviews*, vol. 17, pp. 147–159, Jan. 2013. Available from: https://doi.org/10.1016/j.rser.2012.09.032

[10] J. M. Guerrero, J. C. Vasquez, J. Matas, L. G. de Vicuña, and M. Castilla, "Hierarchical control of droop-controlled AC and DC microgrids—A general approach toward standardization," *IEEE Transactions on Industrial Electronics*, vol. 58, no. 1, pp. 158–172, Jan. 2011. Available from: https://doi.org/10.1109/TIE.2010.2066534

[11] F. Katiraei, R. Iravani, N. Hatziargyriou, and A. Dimeas, "Microgrids management," *IEEE Power & Energy Magazine*, vol. 6, no. 3, pp. 54–65, May–Jun. 2008. Available from: https://doi.org/10.1109/MPE.2008.918702

[12] Y. Zhang, T. Huang, and E. F. Bompard, "Big data analytics in smart grids: A review," *Energy Informatics*, vol. 1, no. 1, pp. 8–24, Dec. 2018. Available from: https://doi.org/10.1186/s42162-018-0013-2

[13] H. Farzin, M. Fotuhi-Firuzabad, and M. Moeini-Aghtaie, "A practical scheme to involve degradation cost of lithium-ion batteries in vehicle-to-grid applications," *IEEE Transactions on Sustainable Energy*, vol. 7, no. 4, pp. 1730–1738, Oct. 2016. Available from: https://doi.org/10.1109/TSTE.2016.2558500

[14] S. M. Dawoud, X. Lin, and M. I. Okba, "Hybrid renewable microgrid optimization techniques: A review," *Renewable and Sustainable Energy Reviews*, vol. 82, pp. 2039–2052, Feb. 2018. Available from: https://doi.org/10.1016/j.rser.2017.08.007

[15] M. Marzband, F. Azarinejadian, M. Savaghebi, and J. M. Guerrero, "An optimal energy management system for islanded microgrids based on multiperiod artificial bee colony combined with Markov chain," *IEEE Systems Journal*, vol. 11, no. 3, pp. 1712–1722, Sep. 2017. Available from: https://doi.org/10.1109/JSYST.2015.2422253

[16] Parisio, E. Rikos, and L. Glielmo, "A model predictive control approach to microgrid operation optimization," *IEEE Transactions on Control Systems Technology*, vol. 22, no. 5, pp. 1813–1827, Sep. 2014. Available from: https://doi.org/10.1109/TCST.2013.2295737

[17] E. Dall'Anese, S. V. Dhople, B. B. Johnson, and F. Dörfler, "Decentralized optimal dispatch of photovoltaic inverters in residential distribution systems," *IEEE Transactions on Energy Conversion*, vol. 29, no. 4, pp. 957–967, Dec. 2014. Available from: https://doi.org/10.1109/TEC.2014.2357997

[18] T. Logenthiran, D. Srinivasan, A. M. Khambadkone, and H. N. Aung, "Multiagent system for real-time operation of a microgrid in real-time digital simulator," *IEEE Transactions on Smart Grid*, vol. 3, no. 2, pp. 925–933, Jun. 2012. Available from: https://doi.org/10.1109/TSG.2012.2189028

[19] N. Hatziargyriou, H. Asano, R. Iravani, and C. Marnay, "Microgrids," *IEEE Power & Energy Magazine*, vol. 5, no. 4, pp. 78–94, Jul.–Aug. 2007. Available from: https://doi.org/10.1109/MPAE.2007.376583

[20] N. Papadimitriou, E. I. Zountouridou, and N. D. Hatziargyriou, "Review of hierarchical control in DC microgrids," *Electric Power Systems Research*, vol. 122, pp. 159–167, May 2015. Available from: https://doi.org/10.1016/j.epsr.2015.01.006

[21] Q. Shafiee, J. M. Guerrero, and J. C. Vasquez, "Distributed secondary control for islanded microgrids—A novel approach," *IEEE Transactions on Power Electronics*, vol. 29, no. 2, pp. 1018–1031, Feb. 2014. Available from: https://doi.org/10.1109/TPEL.2013.2259506

[22] F. Gao and M. R. Iravani, "A control strategy for a distributed generation unit in grid-connected and autonomous modes of operation," *IEEE Transactions on Power Delivery*, vol. 23, no. 2, pp. 850–859, Apr. 2008. Available from: https://doi.org/10.1109/TPWRD.2007.915950

[23] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*, 2nd ed. Cambridge, MA: MIT Press, 2018, pp. 142–189. Available from: https://mitpress.mit.edu/9780262039246/reinforcement-learning-an-introduction/

[24] Mnih et al., "Human-level control through deep reinforcement learning," *Nature*, vol. 518, no. 7540, pp. 529–533, Feb. 2015. Available from: https://doi.org/10.1038/nature14236

[25] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, "Proximal policy optimization algorithms," *arXiv preprint arXiv:1707.06347*, Jul. 2017. Available from: https://arxiv.org/abs/1707.06347

[26] T. P. Lillicrap et al., "Continuous control with deep reinforcement learning," *arXiv preprint arXiv:1509.02971*, Sep. 2015. Available from: https://arxiv.org/abs/1509.02971

[27] G. Tampuu et al., "Multiagent deep reinforcement learning with extremely sparse rewards," *arXiv preprint arXiv:1707.01068*, Jul. 2017. Available from: https://arxiv.org/abs/1707.01068

[28] K. Zhang, Z. Yang, and T. Başar, "Multi-agent reinforcement learning: A selective overview of theories and algorithms," in *Handbook of Reinforcement Learning and Control*, K. G. Vamvoudakis, Y. Wan, F. L. Lewis, and D. Cansever, Eds. Cham, Switzerland: Springer, 2021, pp. 321–384. Available from: https://doi.org/10.1007/978-3-030-60990-0_12

[29] R. Lowe, Y. Wu, A. Tamar, J. Harb, O. P. Abbeel, and I. Mordatch, "Multi-agent actor-critic for mixed cooperative-competitive environments," in *Advances in Neural Information Processing Systems*, vol. 30, 2017, pp. 6379–6390. Available from: https://proceedings.neurips.cc/paper/2017

[30] J. Foerster, G. Farquhar, T. Afouras, N. Nardelli, and S. Whiteson, "Counterfactual multi-agent policy gradients," in *Proceedings of the 32nd AAAI Conference on Artificial Intelligence*, New Orleans, LA, USA, Feb. 2018, pp. 2974–2982. Available from: https://doi.org/10.1609/aaai.v32i1.11794

[31] P. Sunehag et al., "Value-decomposition networks for cooperative multi-agent learning," *arXiv preprint arXiv:1706.05296*, Jun. 2017. Available from: https://arxiv.org/abs/1706.05296

[32] T. Rashid, M. Samvelyan, C. S. de Witt, G. Farquhar, J. Foerster, and S. Whiteson, "QMIX: Monotonic value function factorisation for deep multi-agent reinforcement learning," in *Proceedings of the 35th International Conference on Machine Learning*, Stockholm, Sweden, Jul. 2018, pp. 4295–4304. Available from: https://proceedings.mlr.press/v80/rashid18a.html

[33] J. Wang, Z. Ren, T. Liu, Y. Yu, and C. Zhang, "QPLEX: Duplex dueling multi-agent Q-learning," *arXiv preprint arXiv:2008.01062*, Aug. 2020. Available from: https://arxiv.org/abs/2008.01062

[34] S. Omidshafiei et al., "Deep multi-agent reinforcement learning for partially observable parameterized environments," *arXiv preprint arXiv:1705.06404*, May 2017. Available from: https://arxiv.org/abs/1705.06404

[35] C. Yu, A. Velu, E. Vinitsky, Y. Gao, L. Wang, A. Bayen, and Y. Wu, "The surprising effectiveness of PPO in cooperative multi-agent games," *arXiv preprint arXiv:2103.01955*, Mar. 2021. Available from: https://arxiv.org/abs/2103.01955

[36] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, Dec. 2014. Available from: https://arxiv.org/abs/1412.6980

[37] V. Mnih et al., "Asynchronous methods for deep reinforcement learning," in *Proceedings of the 33rd International Conference on Machine Learning*, New York, NY, USA, Jun. 2016, pp. 1928–1937. Available from: https://proceedings.mlr.press/v48/mniha16.html

[38] J. Schulman, S. Levine, P. Abbeel, M. Jordan, and P. Moritz, "Trust region policy optimization," in *Proceedings of the 32nd International Conference on Machine Learning*, Lille, France, Jul. 2015, pp. 1889–1897. Available from: https://proceedings.mlr.press/v37/schulman15.html

[39] S. Fujimoto, H. van Hoof, and D. Meger, "Addressing function approximation error in actor-critic methods," in *Proceedings of the 35th International Conference on Machine Learning*, Stockholm, Sweden, Jul. 2018, pp. 1587–1596. Available from: https://proceedings.mlr.press/v80/fujimoto18a.html

[40] T. Haarnoja, A. Zhou, P. Abbeel, and S. Levine, "Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor," in *Proceedings of the 35th International Conference on Machine Learning*, Stockholm, Sweden, Jul. 2018, pp. 1861–1870. Available from: https://proceedings.mlr.press/v80/haarnoja18b.html

[41] M. Hessel et al., "Rainbow: Combining improvements in deep reinforcement learning," in *Proceedings of the 32nd AAAI*

Conference on Artificial Intelligence, New Orleans, LA, USA, Feb. 2018, pp. 3215–3222. Available from: https://doi.org/10.1609/aaai.v32i1.11796

[42] Z. Wang, T. Schaul, M. Hessel, H. van Hasselt, M. Lanctot, and N. de Freitas, "Dueling network architectures for deep reinforcement learning," in *Proceedings of the 33rd International Conference on Machine Learning*, New York, NY, USA, Jun. 2016, pp. 1995–2003. Available from: https://proceedings.mlr.press/v48/wangf16.html

[43] T. Schaul, J. Quan, I. Antonoglou, and D. Silver, "Prioritized experience replay," *arXiv preprint arXiv:1511.05952*, Nov. 2015. Available from: https://arxiv.org/abs/1511.05952

[44] H. van Hasselt, A. Guez, and D. Silver, "Deep reinforcement learning with double Q-learning," in *Proceedings of the 30th AAAI Conference on Artificial Intelligence*, Phoenix, AZ, USA, Feb. 2016, pp. 2094–2100. Available from: https://doi.org/10.1609/aaai.v30i1.10295

[45] M. G. Bellemare, W. Dabney, and R. Munos, "A distributional perspective on reinforcement learning," in *Proceedings of the 34th International Conference on Machine Learning*, Sydney, Australia, Aug. 2017, pp. 449–458. Available from: https://proceedings.mlr.press/v70/bellemare17a.html

[46] G. Ostrovski, M. G. Bellemare, A. van den Oord, and R. Munos, "Count-based exploration with neural density models," in *Proceedings of the 34th International Conference on Machine Learning*, Sydney, Australia, Aug. 2017, pp. 2721–2730. Available from: https://proceedings.mlr.press/v70/ostrovski17a.html

[47] Paszke et al., "PyTorch: An imperative style, high-performance deep learning library," in *Advances in Neural Information Processing Systems*, vol. 32, 2019, pp. 8026–8037. Available from: https://proceedings.neurips.cc/paper/2019/hash/bdbca288fee7f92f2bfa9f7012727740-Abstract.html

[48] M. Abadi et al., "TensorFlow: A system for large-scale machine learning," in *Proceedings of the 12th USENIX Symposium on Operating Systems Design and Implementation*, Savannah, GA, USA, Nov. 2016, pp. 265–283. Available from: https://www.usenix.org/conference/osdi16/technical-sessions/presentation/abadi

[49] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015. Available from: https://doi.org/10.1038/nature14539

[50] Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA: MIT Press, 2016, pp. 326–380. Available from: https://www.deeplearningbook.org

[51] X. Glorot and Y. Bengio, "Understanding the difficulty of training deep feedforward neural networks," in *Proceedings of the 13th International Conference on Artificial Intelligence and Statistics*, Sardinia, Italy, May 2010, pp. 249–256. Available from: https://proceedings.mlr.press/v9/glorot10a.html

[52] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in *Proceedings of the 32nd International Conference on Machine Learning*, Lille, France, Jul. 2015, pp. 448–456. Available from: https://proceedings.mlr.press/v37/ioffe15.html

[53] N. Srivastava et al., "Dropout: A simple way to prevent neural networks from overfitting," *Journal of Machine Learning Research*, vol. 15, pp. 1929–1958, 2014. Available from: https://jmlr.org/papers/v15/srivastava14a.html

[54] Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Advances in Neural Information Processing Systems*, vol. 25, 2012, pp. 1097–1105. Available from: https://proceedings.neurips.cc/paper/2012

[55] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, Sep. 2014. Available from: https://arxiv.org/abs/1409.1556

[56] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, NV, USA, Jun. 2016, pp. 770–778. Available from: https://doi.org/10.1109/CVPR.2016.90

[57] Vaswani et al., "Attention is all you need," in *Advances in Neural Information Processing Systems*, vol. 30, 2017, pp. 5998–6008. Available from: https://proceedings.neurips.cc/paper/2017

[58] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," *arXiv preprint arXiv:1810.04805*, Oct. 2018. Available from: https://arxiv.org/abs/1810.04805

[59] Radford, K. Narasimhan, T. Salimans, and I. Sutskever, "Language models are unsupervised multitask learners," *OpenAI Blog*, Feb. 2019. Available from: https://openai.com/blog/better-language-models/

[60] T. Brown et al., "Language models are few-shot learners," *arXiv preprint arXiv:2005.14165*, May 2020. Available from: https://arxiv.org/abs/2005.14165

[61] X. Liu, A. Shahidehpour, and Z. Li, "Microcontroller-based microgrid energy management systems: Design principles and deployment challenges," *IEEE Transactions on Smart Grid*, vol. 12, no. 3, pp. 2156–2168, May 2021. Available from: https://doi.org/10.1109/TSG.2020.3039876

[62] P. Siano, "Demand response and smart grids—A survey," *Renewable and Sustainable Energy Reviews*, vol. 30, pp. 461–478, Jun. 2014. Available from: https://doi.org/10.1016/j.rser.2013.10.022

[63] S. Maharjan, Q. Zhu, Y. Zhang, S. Gjessing, and T. Başar, "Dependence of demand response and renewable energy penetration on the electricity market," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 6, pp. 1031–1041, Dec. 2014. Available from: https://doi.org/10.1109/JSTSP.2014.2332296

[64] D. E. Olivares et al., "Trends in microgrid control," *IEEE Transactions on Smart Grid*, vol. 5, no. 4, pp. 1905–1919, Jul. 2014. Available from: https://doi.org/10.1109/TSG.2013.2295514

[65] Oudalov, D. Chartouni, and C. Ohler, "Deploying active distribution networks for the smooth integration of distributed generation," *Energy Policy*, vol. 35, no. 10, pp. 4999–5012, Oct. 2007. Available from: https://doi.org/10.1016/j.enpol.2007.04.003

[66] R. H. Lasseter, "Smart distribution: Coupled microgrids," *Proceedings of the IEEE*, vol. 99, no. 6, pp. 1074–1082, Jun. 2011. Available from: https://doi.org/10.1109/JPROC.2011.2106670

[67] F. A. Cleveland, "Cyber security issues for control systems," in *IEEE Conference on Technologies for Homeland Security and Information Networking*, Oct. 2011, pp. 1–6. Available from: https://doi.org/10.1109/CNSIV.2011.6080957

ABOUT THE AUTHOR

Professor Adel Elgammal is currently a Professor of Electrical Engineering and Assistant Vice President - Research, Impact and Postgraduate Studies AVP-RIPS at the University of Trinidad and Tobago UTT. He received his B.Sc. Degree in Electrical Power Engineering from Helwan University-EGYPT in 1996. He completed his M.Eng. Degree in Electric Drives and Machines Engineering in 2002 and Ph.D. Degree in Jan-2007 from the Faculty of Engineering (Helwan University-EGYPT). In May 2008, Professor Elgammal has joined the University of Trinidad and Tobago (UTT) as an Assistant Professor and then he was promoted to Associate Professor in May 2010 at Utilities and Sustainable Engineering, UTT. In 2010, Professor Elgammal has been elevated to senior member of Institute of Electrical and Electronics Engineers (IEEE). Professor Elgammal has been professionally registered as Chartered Engineer (CEng) IET. Professor Elgammal was promoted to Professor in February 2023 at Utilities and Sustainable Engineering, UTT. Professor Elgammal authored and co-authored over 47 Scholarly Technical Journals, and over 77 Refereed Conference Publications and three Engineering Book Chapters. His main research emphasis the application of Intelligent Systems (Particle Swarm Optimization PSO, fuzzy logic, neutral networks, and Genetic algorithms) to Power Systems, renewable / Green Energy systems, demand-side sustainable management, and Electric Drives, Efficient control and utilization of renewable energy and green technology, Hybrid renewable systems (PV, wind, diesel, small hydro, hydrogen, etc.), E-Vehicle (Battery, PV, Hydrogen FC powered).