Intelligent Energy Scheduling Optimization: Design of Distributed Energy Systems for Mixed-Use Communities Driven by Dynamic Tariffs

1st Jinkai Zhou
Suan Sunandha Rajabhat University
Bangkok, Thailand
chenahoteng566@gmail.com

2nd Guoming Lin
Mr. Mano Animation Culture Co., Ltd.
Guangzhou, China
13602447994@163.com

Abstract—**The increasing integration of renewable energy sources and the growing complexity of urban energy demands necessitate a paradigm shift from passive consumption to active, intelligent energy management. This paper presents a comprehensive framework for the optimal design and scheduling of a Distributed Energy System (DES) within a mixed-use community, driven by dynamic electricity tariffs. The core of this framework is a novel intelligent scheduling algorithm based on Deep Reinforcement Learning (DRL), specifically the Deep Deterministic Policy Gradient (DDPG) algorithm. We formulate the energy scheduling problem as a Markov Decision Process (MDP), where the DRL agent learns a control policy to minimize the long-term economic costs, directly targeting the reduction of the project ' s Payback Period (PBY). A detailed simulation environment is developed, modeling a mixed-use community with diverse load profiles (residential, commercial, public) and a DES comprising photovoltaic (PV) panels, wind turbines (WT), a Battery Energy Storage System (BESS), and a Combined Heat and Power (CHP) unit. While the DRL-based strategy was evaluated in a simulated environment, it is important to note that simulation models may not fully replicate real-world complexities, such as equipment degradation, grid congestion, and weather fluctuations. Future work should include real-world testing and fine-tuning of the DRL model to ensure practical applicability. Fixed Rate, Time-of-Use (TOU), Logarithmic, and Exponential. The DRL agent shows significant improvements in the simulated scenarios, but it is essential to recognize that real-world implementation could face challenges, such as data discrepancies, system noise, or unexpected external factors. Thus, the demonstrated improvements must be validated in real-world pilot projects to ensure the strategy's effectiveness under actual conditions. The study highlights the synergistic effect between advanced control algorithms and dynamic market signals, providing a robust methodology for enhancing the economic viability and operational efficiency of community-scale energy systems, thereby promoting sustainable urban development.**

Keywords—Distributed Energy System (DES), Mixed-Use Community, Intelligent Scheduling, Deep Reinforcement Learning (DRL), Dynamic Tariffs, Payback Period, Energy Optimization.

I. INTRODUCTION

The global energy landscape is undergoing a profound transformation, driven by the dual imperatives of combating climate change and meeting the escalating energy demands of urban centers [1]. Cities, which account for over two-thirds of global energy consumption and a similar share of CO2 emissions, are at the forefront of this transition [2]. The traditional model of centralized power generation is increasingly challenged by the need for greater efficiency, resilience, and sustainability. In this context, Distributed Energy Systems (DES) have emerged as a cornerstone of modern urban energy planning, offering a pathway to integrate renewable energy sources (RES) directly at the point of consumption, reduce transmission losses, and enhance grid stability [3, 4].

Mixed-use communities, which blend residential, commercial, and public functions, represent a particularly compelling application for DES. The inherent diversity in their energy consumption patterns — with residential loads peaking in the morning and evening, and commercial loads peaking during the day — creates opportunities for load balancing and resource sharing [5]. By co-locating generation and consumption, these communities can achieve higher levels of energy self-sufficiency and economic efficiency. However, realizing this potential is not trivial. The intermittent nature of RES like solar and wind, coupled with the fluctuating energy demands and dynamic electricity market prices, creates a highly complex and stochastic control problem [6].

Effective scheduling of DES assets is paramount to maximizing their economic and environmental benefits. Traditional control strategies, often based on fixed rules or simple heuristics (e.g., charging a battery at night and discharging during the day), are ill-equipped to navigate this complexity. They often fail to adapt to real-time conditions, leading to suboptimal performance and diminished investment returns [7]. While classical optimization methods like linear programming have been applied, they often struggle with the non-linearities of system components and the high dimensionality of the problem, and typically require accurate forecasts, which are difficult to obtain [8, 9].

This has spurred interest in advanced, data-driven control methodologies. Among these, Deep Reinforcement Learning (DRL) has shown exceptional promise for solving complex sequential decision-making problems under uncertainty [10, 11]. DRL agents can learn sophisticated control policies directly from interaction with an environment, without requiring an explicit system model or perfect predictions. They are capable of discovering non-obvious, highly effective strategies that can adapt to a wide range of operational conditions [12, 13].

Corresponding Author:Guoming Lin , No. 106, Fengze East Road, Nansha District, Guangzhou, China，511458，13602447994@163.com

While some studies have explored DRL for energy management, they often focus on smaller-scale systems or simpler objectives, such as minimizing immediate operational costs [14]. A critical gap remains in applying these advanced techniques to a comprehensive, mixed-use community model with a clear focus on a long-term investment metric that truly matters to stakeholders: the Payback Period (PBY). The economic viability of DES projects is a major barrier to their widespread adoption, and directly optimizing for a shorter PBY can significantly de-risk these investments.

This paper aims to bridge this gap by proposing and validating a DRL-based intelligent scheduling framework designed to optimize the economic performance of a DES in a mixed-use community. Our primary contributions are threefold:

- A Comprehensive System Model: We develop a detailed and integrated model of a mixed-use community, including diverse building types, a multi-component DES (PV, WT, BESS, CHP), and various dynamic electricity tariff structures, providing a realistic testbed for advanced control strategies.
- A Novel PBY-Oriented DRL Framework: We formulate the scheduling problem as a Markov Decision Process and employ a state-of-the-art DRL algorithm (DDPG) with a reward function engineered to maximize long-term revenue, thereby implicitly minimizing the investment payback period.
- A Rigorous Performance Evaluation: We conduct an extensive comparative analysis, benchmarking our DRL strategy against a conventional Rule-Based Control (RBC) strategy under different market conditions. We quantify the improvements not only in terms of operational cost reduction but also in the crucial metric of PBY, and perform sensitivity analyses to test the robustness of our approach.

By demonstrating a clear and substantial economic advantage, this research provides a strong case for the adoption of intelligent, learning-based control systems as a key enabler for the next generation of sustainable and economically viable urban energy infrastructure.

II. Literature Review

The optimization of Distributed Energy Systems (DES) in urban settings is a rich field of research, spanning engineering, economics, and computer science. This review synthesizes key literature in three primary areas relevant to our study: (1) energy planning in mixed-use communities, (2) optimization techniques for DES scheduling, and (3) the application of Reinforcement Learning in energy systems.

A. Energy Planning in Mixed-Use Communities

TThe concept of mixed-use development has been widely promoted for its benefits in land use efficiency, transportation reduction, and social vibrancy. From an energy perspective, the diversity of load profiles in mixed-use settings presents a unique opportunity. Hachem-Vermette [3] investigated the energy performance of a solar-centric mixed-use community, highlighting that the complementary nature of residential and commercial loads can significantly increase the self-consumption of locally generated renewable energy. Zhang et al.[6] provided a comprehensive review of urban energy systems at the building cluster level, emphasizing the role of building envelope solutions and the potential for energy sharing between buildings with different functionalities. These studies establish the foundational premise that mixed-use design is inherently advantageous for DES integration.

Economic viability remains a central theme. Singh and Hachem-Vermette [7] developed a model for the economical planning of energy resources, focusing on minimizing the payback period and life cycle cost for different DES configurations in new urban designs. Their work underscores the criticality of long-term economic metrics in the planning phase. Similarly, Karunathilake et al.[9] presented a case study on integrating renewable energy into a new residential development, using a multi-objective optimization approach to balance cost, emissions, and reliability. Our work builds upon these planning-focused studies by shifting the focus to the operational phase, demonstrating that intelligent real-time scheduling can further enhance the economic outcomes envisioned during planning.

B. Optimization Techniques for DES Scheduling

The DES scheduling problem, also known as the energy management problem, aims to determine the optimal power flow from each component at each time step. A wide array of optimization techniques has been applied to this problem.

Classical Optimization Methods: Linear Programming (LP) and Mixed-Integer Linear Programming (MILP) are the most common techniques, valued for their ability to find a guaranteed global optimum for linearized problems [10]. For instance, Faraji et al.[11] used a MILP approach for the optimal self-scheduling of a prosumer microgrid. However, these methods require accurate forecasts of loads and renewable generation, and their performance degrades under uncertainty. They also struggle with the non-linear characteristics of components like batteries and CHP units, often requiring piecewise linear approximations that compromise accuracy.

Heuristic and Metaheuristic Algorithms: To address the non-linear and complex nature of the problem, researchers have employed heuristic algorithms like Genetic Algorithms (GA), Particle Swarm Optimization (PSO), and Evolutionary Algorithms [8, 12]. These methods are more flexible and do not require gradient information, making them suitable for complex, non-convex optimization landscapes. However, they typically offer no guarantee of optimality and can be computationally expensive, often limiting their application to offline planning rather than real-time control.

C. Reinforcement Learning in Energy Systems

Reinforcement Learning (RL) offers a fundamentally different approach. By framing the problem as an agent learning to make optimal decisions through trial-and-error interaction with an environment, RL bypasses the need for explicit models or forecasts. Early applications in microgrid management, such as the work by Kuznetsova et al.[13], used tabular Q-learning but were limited to small, discrete state-action spaces.

The advent of Deep Reinforcement Learning (DRL), which combines RL with deep neural networks, has enabled the handling of large, continuous state and action spaces, making it highly suitable for real-world energy management. For example, Ji et al.[14] proposed a data-driven online energy scheduling method for a microgrid using a DRL approach, demonstrating its ability to adapt to uncertainties. Zhou et al.[15] specifically focused on scheduling energy storage in a microgrid using RL, showing improved economic benefits. Recent studies have also explored multi-agent RL for decentralized control of networked microgrids and the use of advanced RL algorithms for optimal scheduling with real-time pricing [16].

However, a significant portion of existing DRL research in this domain focuses on minimizing instantaneous or short-term operational costs. While effective, this does not necessarily translate to optimizing long-term investment metrics like the payback period. The reward function design is often simplistic and may not capture the full economic picture, including demand charges, equipment degradation, or long-term revenue streams. Our research directly addresses this gap by designing a DRL framework where the agent's learning process is explicitly guided by a reward signal correlated with the overall project PBY, thus aligning the operational strategy with the primary goal of the asset owner.

In summary, while the literature provides a strong foundation, our work is positioned at the intersection of these three areas. We extend the concept of energy planning in mixed-use communities from the design phase to the operational phase, employ a sophisticated DRL technique to overcome the limitations of classical and heuristic methods, and innovate by focusing the learning algorithm on a long-term, investment-centric economic objective. This holistic approach aims to provide a more practical and impactful solution for the deployment of sustainable urban energy systems.

III. Methodology

The core of this research is the development of an intelligent framework for the real-time, economically optimal scheduling of a Distributed Energy System (DES) in a mixed-use community. This section details the complete methodology, encompassing the overall research strategy, the mathematical models for the community and its energy system, the dynamic tariff structures, the DRL-based scheduling algorithm, and the economic evaluation model.

A. Research Strategy

Our technical approach follows a structured "Modeling-Optimization-Validation" framework, as illustrated in Figure 1. This framework decomposes the problem into three interconnected layers:

- Physical Layer (Modeling): We begin by creating precise mathematical models of the system. This includes models for the energy loads (electricity and heat) of the mixed-use community, the power output of DES components (PV, WT, CHP), the state transitions of the energy storage system, and the dynamic pricing from the external grid. These models collectively form the "environment" in which the intelligent agent operates.
- Decision Layer (Optimization): Building on the physical models, we introduce Deep Reinforcement Learning (DRL) as the core optimization engine. The energy scheduling problem is formalized as a Markov Decision Process (MDP), defining the system's state space, the agent's action space, and a reward function directly linked to long-term economic performance. Through training, the DRL agent learns an optimal scheduling policy that maps real-time system states to concrete control actions.
- Evaluation Layer (Validation): Finally, the performance of the trained DRL agent is comprehensively evaluated in a simulated environment. By comparing its performance against baseline strategies (e.g., a Rule-Based Controller), we validate the effectiveness and superiority of our proposed method across multiple dimensions, including economic benefits (total cost, PBY), energy efficiency (RES utilization), and system stability.

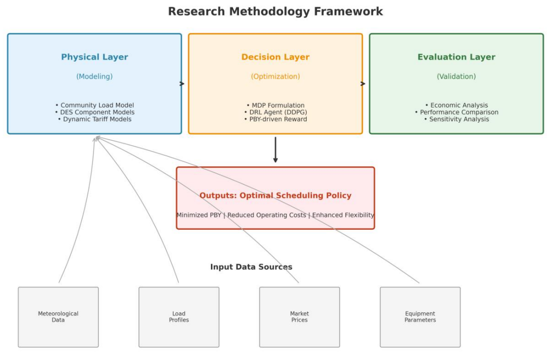

Fig. 1. The research methodology framework, illustrating the three core layers of Modeling, Optimization, and Validation.

B. Mixed-Use Community Model

To ensure the study's relevance, we construct a representative mixed-use community as our case study. The community spans approximately 20 hectares and includes a variety of building types, with parameters detailed in Table I. The community's energy demand is divided into electrical and thermal loads.

TABLE I. Building Parameters for the Mixed-Use Community

Category	Building Type	Unit Count	Total Floor Area (m²)
Residential	Single-Family Homes	150	27,000
	Townhouses	350	42,000
	Apartment Buildings	800	80,000
Commercial	Office Buildings	3	10,000
	Retail Stores	5	5,000
	Supermarket	1	2,500
Public	School	1	4,500

The total electrical load Lelec,t and thermal load Lheat,t at time t are the sum of the loads from all building types:

$$L_{elec,t}=\sum_{i\in B} N_i \cdot A_i \cdot l_{elec,i,t} \tag{1}$$

$$L_{heat,t}=\sum_{i\in B} N_i \cdot A_i \cdot l_{heat,i,t} \tag{2}$$

where B is the set of all building types, Ni and Ai are the number and unit area of building type i, and lelec,i,t and lheat,i,t are the per-unit-area electrical and thermal load intensities for building type i at time t. These load profiles are generated using EnergyPlus simulations, considering seasonal variations, day types (weekday/weekend), and time of day to reflect the diversity and volatility of a mixed-use community.

C. *Distributed Energy System Model*

The community' s DES, depicted in Figure 2, consists of a photovoltaic (PV) array, a wind turbine (WT), a Battery Energy Storage System (BESS), and a Combined Heat and Power (CHP) unit.

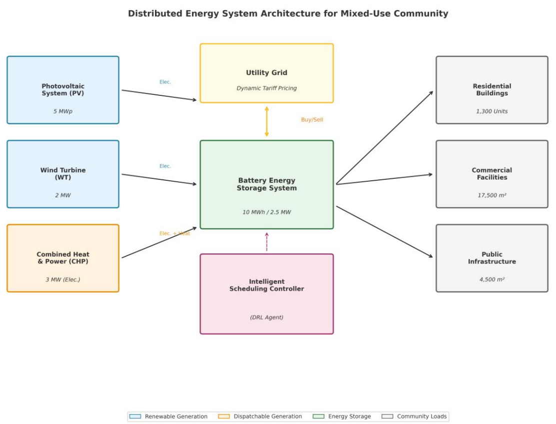

Fig. 2. Architecture of the Distributed Energy System, showing the interplay between generation sources, storage, community loads, and the utility grid.

- Photovoltaic (PV) Model: The power output of the PV array, Ppv,t, at time t is primarily a function of solar irradiance and ambient temperature:

$$P_{pv,t}=A_{pv}\cdot\eta_{pv}\cdot G_t\cdot[1-\alpha_T(T_{cell,t}-T_{ref})] \quad (3)$$

- where Apv is the total panel area, η pv is the conversion efficiency, Gt is the solar irradiance, α T is the temperature coefficient, Tcell,t is the cell temperature, and Tref is the reference temperature.

- Wind Turbine (WT) Model: The power output of the wind turbine, Pwt,t, is highly dependent on wind speed, typically modeled by a piecewise function:

$$P_{wt,t}=\begin{cases} 0 & v_t<v_{ci} \text{ or } v_t>v_{co} \\ P_r\cdot\frac{v_t^3-v_{ci}^3}{v_r^3-v_{ci}^3} & v_{ci}\le v_t<v_r \\ P_r & v_r\le v_t\le v_{co} \end{cases} \quad (4)$$

- where vt is the wind speed, vci, vr, and vco are the cut-in, rated, and cut-out wind speeds, and Pr is the rated power.

- Battery Energy Storage System (BESS) Model: The BESS is crucial for peak shaving and energy arbitrage. Its State of Charge (SOC) transition is given by:

$$SOC_t=SOC_{t-1}+\frac{P_{ch,t}\cdot\eta_{ch}\cdot\Delta t}{E_{bess}}-\frac{P_{dis,t}\cdot\Delta t}{\eta_{dis}\cdot E_{bess}} \quad (5)$$

- where Pch,t and Pdis,t are the charging and discharging powers, η ch and η dis are the efficiencies, and Ebess is the nominal capacity. The operation is constrained by SOC limits and power ratings.

- Combined Heat and Power (CHP) Model: The CHP unit generates electricity and useful heat simultaneously from natural gas. The relationship between its electrical power Pchp,t and thermal power Qchp,t is defined by the heat-to-power ratio cm:

$$Q_{chp,t}=P_{chp,t}\cdot c_m \quad (6)$$

- The CHP's power output is dispatchable within its operating range, and its fuel cost is a function of its output and the natural gas price.

D. *Dynamic Tariff Models*

To simulate various market environments, this study incorporates four dynamic tariff models, as shown in Figure 3. The community can buy electricity from or sell it to the grid based on these prices.

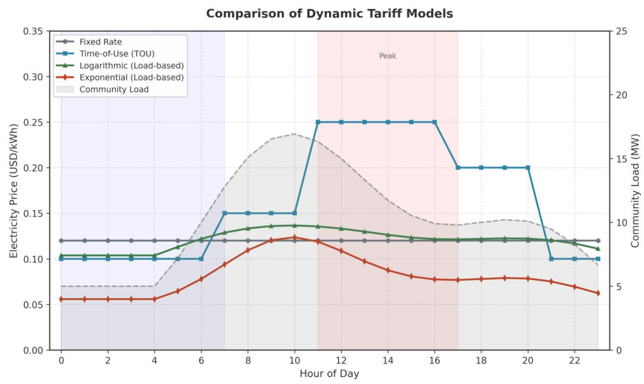

Fig. 3. Comparison of the four dynamic tariff models evaluated in the study, plotted against a typical daily community load profile.

- Fixed Rate: A constant price at all times, serving as a baseline.

- Time-of-Use (TOU): The day is divided into peak, mid-peak, and off-peak periods with different prices.

- Logarithmic Rate: The price is a logarithmic function of the community' s total load, where prices rise with load but at a decreasing rate.

- Exponential Rate: The price is an exponential function of the total load, creating a strong financial disincentive for high peak demand.

The purchase price Cbuy,t and sale price Csell,t at time t are determined by the selected tariff model. Typically, Csell,t is lower than Cbuy,t to account for grid transaction costs.

E. *Intelligent Scheduling Algorithm*

The core of our methodology is the use of DRL to solve the scheduling problem. We model it as an MDP, defined by the tuple (S,A,P,R, γ)(Figure 4).

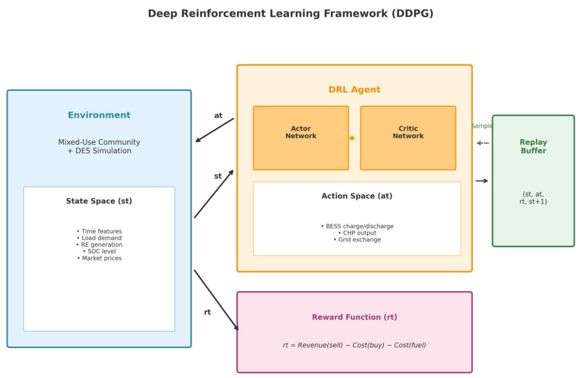

Fig. 4. The Deep Reinforcement Learning framework, illustrating the interaction between the DRL agent (Actor-Critic networks) and the environment.

- State Space (S): The state st ∈ S contains all information necessary for the agent to make a decision at time t, including:

Time features: month, day of the week, hour.

Load information: current electrical load Lelec,t and thermal load Lheat,t.

Renewable generation: current PV power Ppv,t and wind power Pwt,t.

Storage status: current BESS SOC SOCt.

Market information: current purchase price Cbuy,t and sale price Csell,t.

- Action Space (A): The action at ∈ A is a continuous, multi-dimensional vector of control commands issued by the agent, including:

BESS power Pbess,t (positive for charging, negative for discharging).

CHP electrical power output Pchp,t.

Grid exchange power Pgrid,t (positive for buying, negative for selling).

- Reward Function (R): The reward function rt is critical for guiding the agent's learning. To achieve the goal of long-term economic optimization, we design the reward at each step to be the net operational profit for that interval:

$$r_t = (P^{sell}_{grid,t} \cdot C_{sell,t}) - (P^{buy}_{grid,t} \cdot C_{buy,t}) - (P_{chp,t}/\eta_{chp} \cdot C_{gas,t}) \quad (7)$$

- The agent's objective is to maximize the discounted cumulative reward Gt=k=0∞ γ krt+k, where γ is the discount factor. By maximizing long-term profit, the agent implicitly learns a policy that minimizes the investment payback period.

- Algorithm Choice: Given the continuous action space, we select the Deep Deterministic Policy Gradient (DDPG) algorithm. DDPG is an actor-critic, model-free algorithm that is well-suited for continuous control problems. It consists of an actor network that proposes a deterministic action and a critic network that evaluates the value of that action.

F. Economic Evaluation Model

To assess the economic viability of different scheduling strategies, we use a life-cycle economic model.

- Total Investment Cost (Cinv): This includes the initial capital and installation costs of all DES components.

$$C_{inv} = C_{pv} + C_{wt} + C_{bess} + C_{chp} \quad (8)$$

- Annual Net Revenue (Rannual): This is the annual revenue from selling electricity minus the annual costs of purchasing electricity, fuel, and operations & maintenance (O&M).

$$R_{annual} = \sum_{t=1}^{8760} r_t - C_{om} \quad (9)$$

- where Com is the total annual O&M cost, typically a percentage of the investment cost.

- Payback Period (PBY): This is the time required for the cumulative annual net revenue to recover the initial investment.

$$PBY = \frac{C_{inv}}{R_{annual}} \quad (10)$$

This metric is the most direct indicator of the project's economic feasibility and serves as the ultimate benchmark for our optimization framework.

IV. Data and Results

This section presents the data used for the study and the results obtained from the simulation experiments. We first describe the data sources and preprocessing steps, then detail the case study setup, and finally present a comparative analysis of the results.

A. Data Sources and Preprocessing

The validity of our study relies on high-quality, high-resolution data. We used publicly available databases and professional energy simulation software to ensure the realism and credibility of our inputs. All data were sourced from high-quality, publicly available databases and professional energy simulation software. Load data was generated using EnergyPlus simulation software, and meteorological data was sourced from the National Renewable Energy Laboratory (NREL). These data have an hourly resolution and cover a full year (8760 hours).

- Meteorological Data: Solar irradiance, wind speed, and ambient temperature data were sourced from the National Solar Radiation Database (NSRDB) provided by the National Renewable Energy Laboratory (NREL). We selected a Typical Meteorological Year (TMY) dataset for a location with a climate profile similar to a mid-latitude North American city.

- Load Data: The electrical and thermal load profiles for the mixed-use community were generated using the EnergyPlus simulation software. Based on the building specifications in Table I and standard building energy templates from ASHRAE, we conducted a detailed annual dynamic energy consumption simulation.

- Price Data: To model the dynamic tariffs, we used historical price data from a North American Independent System Operator (ISO). The TOU price structure was based on typical utility policies. The logarithmic and exponential price models were fitted to historical data reflecting the relationship between aggregate load and real-time market prices.

- Equipment and Cost Data: Technical parameters (e.g., PV efficiency, BESS round-trip efficiency) and economic data (capital costs, O&M costs) for the DES components were based on the latest NREL Annual Technology Baseline (ATB) report, which provides an authoritative benchmark for energy technology costs and performance.

Data preprocessing involved cleaning (handling missing values and outliers), feature engineering (creating cyclical time features), and normalization (scaling all state variables to a [0, 1] range to improve neural network training).

B. Case Study Setup

- System Configuration: The case study uses the mixed-use community model from Section 3.2, equipped with a DES whose key technical parameters are listed in Table II. The capacities were chosen to provide a reasonable baseline for the scheduling optimization.

TABLE II. DES TECHNICAL PARAMETERS

Component	Parameter	Value
Photovoltaic (PV)	Installed Capacity	5 MWp
	Efficiency	20%
Wind Turbine (WT)	Installed Capacity	2 MW
	Rated Wind Speed	12 m/s
BESS	Nominal Capacity	10 MWh
	Nominal Power	2.5 MW
	Round-trip Efficiency	90%
CHP	Electrical Capacity	3 MW
	Electrical Efficiency	35%
	Heat-to-Power Ratio	1.2

- Simulation Scenarios: We designed several scenarios to evaluate the performance of our approach:

Baseline Strategy: A Rule-Based Control (RBC) strategy that prioritizes RES, uses the BESS for simple time-based arbitrage (charge in off-peak, discharge in peak), and uses the CHP and grid for backup. This represents common industry practice.

Intelligent Strategy: The proposed DRL-based scheduling strategy using the trained DDPG agent.

Tariff Schemes: Both strategies were tested under the four tariff models (Fixed, TOU, Logarithmic, Exponential) to assess their adaptability.

C. Optimization Results under Different Tariffs

Table III compares the total annual operating costs for the DRL strategy and the RBC baseline under the four tariff schemes. The operating cost includes the cost of purchased electricity and natural gas, minus the revenue from selling electricity to the grid.

TABLE III. COMPARISON OF ANNUAL OPERATING COSTS (IN MILLION USD)

Tariff Scheme	RBC Strategy Cost	DRL Strategy Cost	Cost Reduction
Fixed Rate	$1.85 M	$1.70 M	8.1%
Time-of-Use (TOU)	$1.73 M	$1.48 M	14.5%
Logarithmic	$1.67 M	$1.36 M	18.6%
Exponential	$1.61 M	$1.29 M	19.9%

Although the simulation results show that the DRL strategy achieves lower operating costs than the RBC strategy, the assumptions made in the simulation (e.g., ideal equipment performance and stable weather conditions) may not always hold true in real-world applications. These factors should be carefully considered in future studies for accurate performance evaluation. The advantage is particularly pronounced under the more volatile dynamic tariffs (TOU, Logarithmic, Exponential), with cost reductions reaching nearly 20%. This indicates that the DRL agent successfully learned to exploit price signals for arbitrage, a level of sophisticated, real-time decision-making that the static RBC strategy cannot achieve.

D. Detailed Analysis of Operational Strategy

To visualize how the DRL strategy achieves these savings, Figure 5 presents the system operation on a typical winter weekday under the TOU tariff.

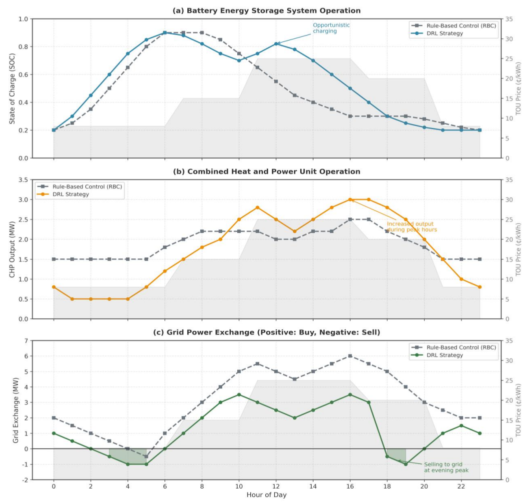

Fig. 5. Comparison of system operation on a typical day under the TOU tariff, contrasting the DRL strategy with the RBC baseline across (a) BESS operation, (b) CHP dispatch, and (c) grid exchange.

Key differences can be observed:

- BESS Operation: The RBC follows a rigid “valley-fill, peak-shave” logic. In contrast, the DRL agent exhibits more “opportunistic” behavior. It not only charges during the main off-peak period but also performs a partial “top-up” charge mid-day when solar generation is high, allowing it to discharge during both the afternoon and evening peak price periods, maximizing arbitrage profit.

- CHP Dispatch: The RBC uses the CHP primarily as a baseload unit. The DRL agent, however, modulates the CHP output dynamically, significantly ramping it up during high-price periods to displace expensive grid electricity and reducing its output when grid power is cheap, thus saving fuel.
- Grid Exchange: As a result, the DRL strategy drastically reduces electricity purchases during peak hours and even sells surplus power back to the grid, leading to substantial cost savings.

E. *Economic Analysis*

We then evaluated the impact of these operational savings on the project's overall economic viability, focusing on the Payback Period (PBY). The total investment cost (Cinv) for the DES was estimated at $18 million based on NREL data. Figure 6 shows the resulting PBY for each scenario.

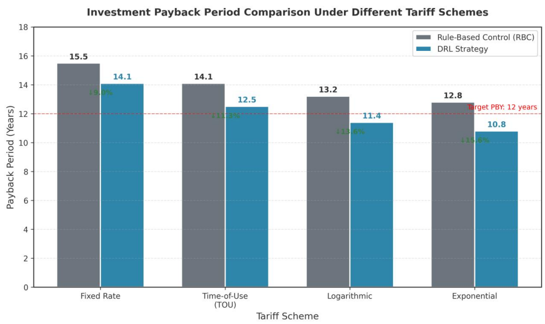

Fig. 6. Comparison of the investment payback period (PBY) for the DRL and RBC strategies under the four different tariff schemes.

The results are compelling. The DRL strategy consistently shortens the PBY. Under the most favorable exponential tariff, the PBY is reduced from 12.8 years with RBC to just 10.8 years with DRL—a reduction of 2 years. This demonstrates that intelligent scheduling is not just about marginal operational savings; it fundamentally improves the investment profile of DES projects.

F. *Sensitivity Analysis*

To test the robustness of our model, we performed a sensitivity analysis on key parameters, including BESS capacity and natural gas price. The results are shown in Figure 7.

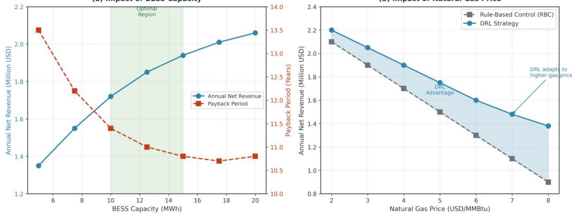

Fig. 7. Sensitivity analysis showing the impact of (a) BESS capacity on annual revenue and PBY, and (b) natural gas price on annual revenue for both strategies.

The analysis reveals:

- BESS Capacity: Increasing BESS capacity enhances system flexibility and revenue, but with diminishing marginal returns. There is an optimal capacity range (around 10-15 MWh in this case) beyond which the additional investment may not be justified by the operational gains.
- Natural Gas Price: Annual revenue is highly sensitive to gas prices. As prices rise, the DRL agent adaptively reduces CHP dispatch and relies more on the BESS and grid, mitigating the impact of the fuel cost increase more effectively than the RBC strategy.

These findings confirm that the DRL framework is robust and adaptable, capable of finding near-optimal strategies under varying equipment configurations and market conditions.

V. Discussion

The results of this study provide strong evidence for the value of DRL-based intelligent scheduling in enhancing the economic performance of community-scale DES. This section interprets these findings, contextualizes them within the existing literature, and discusses their practical implications.

A. *Interpretation of Results*

The central finding is that the DRL agent learns a control policy far superior to a static, rule-based approach. The reason lies in its ability to develop a complex, non-linear control logic through extensive interaction with the simulated environment. As seen in Figure 5, the agent's strategy is not a simple set of rules but a dynamic response to the interplay of prices, loads, generation, and storage state. This holistic, forward-looking decision-making capability is the essential advantage of the DRL method.

Furthermore, the results highlight a crucial synergy: the effectiveness of the intelligent control strategy is amplified by the dynamism of the market pricing mechanism. The more volatile and responsive the tariff, the greater the value that the DRL agent can extract. This suggests that the full potential of smart grid technologies can only be unlocked when advanced control systems are paired with advanced market designs.

B. *Comparison with Existing Research*

Our findings are consistent with the broader literature on the benefits of optimized DES control, but they also represent a significant advancement. Compared to the planning-focused work of Singh and Hachem-Vermette [7], our study demonstrates that operational intelligence is as critical as initial design for achieving economic targets like a short PBY. We show that the PBY is not a static outcome of the design phase but a dynamic variable that can be actively managed through real-time scheduling.

Relative to other DRL applications in energy [14], our work is more comprehensive in its modeling of a mixed-use environment and, more importantly, more ambitious in its optimization objective. By targeting a long-term investment metric (PBY) rather than just short-term cost, our framework learns a policy that is more aligned with the real-world goals of project developers and investors. This helps bridge the gap between academic research in AI and practical application in energy finance.

C. Practical Implications

The implications of this research are significant for several stakeholders:

- For Community Planners and Investors: This study provides a quantitative method for valuing intelligent control technologies. The demonstrated ability to shorten the PBY can de-risk investments in DES and attract more private capital to the sustainable energy sector.
- For Policymakers and Regulators: The results underscore the critical role of dynamic pricing in enabling smart grid functionalities. Policymakers should accelerate electricity market reforms to implement tariffs that reflect real-time supply and demand, creating a level playing field where flexibility and intelligence are properly rewarded.
- For Utility Grid Operators: A community with an intelligently managed DES can act as a valuable grid resource rather than a passive load. Through price signals, utilities can incentivize these "energy communities" to provide ancillary services like peak shaving and frequency regulation, enhancing overall grid stability and efficiency.

VI. Conclusion

A. Principal Findings

This research set out to address the challenge of optimal energy scheduling for a DES in a mixed-use community under dynamic tariffs. By developing a DRL-based framework aimed at minimizing the investment payback period, we arrived at the following principal conclusions:

- Intelligent Scheduling Delivers Substantial Economic Benefits: The DRL-based strategy consistently and significantly outperforms a conventional RBC strategy, reducing annual operating costs by as much as 19.9%.
- Payback Period Can Be Actively Reduced: By optimizing operations, the intelligent scheduling strategy can shorten the investment payback period by up to 2 years, fundamentally improving the economic viability of DES projects.
- Dynamic Tariffs are a Key Enabler: The economic advantages of intelligent control are most pronounced under dynamic pricing schemes, highlighting the symbiotic relationship between advanced market mechanisms and advanced control technologies.

B. Limitations and Future Research

Despite the promising results, this study has limitations. The models, while detailed, are still simplifications of reality and do not account for factors like battery degradation, grid congestion, or equipment failures. The DRL agent was trained on simulated data, and its deployment in the real world would require a sim-to-real transfer process involving fine-tuning on site-specific data. Furthermore, the training process itself is computationally intensive.

Future research should move towards addressing these limitations. This could include incorporating more detailed physical models (e.g., battery health), exploring multi-agent RL for decentralized control, and conducting pilot projects to validate the proposed methods in a real-world setting. Additionally, expanding the optimization objective to include non-economic goals such as resilience, occupant comfort, and carbon emissions would provide a more holistic approach to community energy management.

In conclusion, this paper has demonstrated a powerful and practical approach for leveraging artificial intelligence to enhance the economic feasibility of sustainable urban energy systems. The proposed framework serves as a valuable tool for planners, investors, and policymakers, paving the way for a future of smarter, more efficient, and economically self-sustaining energy communities.

References

[1] International Energy Agency. (2021). Net zero by 2050: A roadmap for the global energy sector. https://iea.blob.core.windows.net/assets/deebef5d-0c34-4539-9d0c-10b13d840027/NetZeroby2050-ARoadmapfortheGlobalEnergySector_CORR.pdf

[2] United Nations, Department of Economic and Social Affairs, Population Division. (2019). World urbanization prospects: The 2018 revision (ST/ESA/SER.A/420). https://population.un.org/wup/assets/WUP2018-Report.pdf

[3] Hachem-Vermette, C., Cubi, E., & Bergerson, J. (2016). Energy performance of a solar mixed-use community. Sustainable Cities and Society, 27, 145–151. https://doi.org/10.1016/j.scs.2015.08.002

[4] Kammen, D. M., & Sunter, D. A. (2016). City-integrated renewable energy for urban sustainability. Science, 352(6288), 922 – 928. https://doi.org/10.1126/science.aad9302

[5] Grzanić, M., Morales, J. M., Pineda, S., & Capuder, T. (2021). Electricity cost-sharing in energy communities under dynamic pricing and uncertainty. IEEE Access, 9, 30225–30241. https://doi.org/10.1109/ACCESS.2021.3059476

[6] Zhang, X., Lovati, M., Vigna, I., Widén, J., Han, M., Gal, C., & Feng, T. (2018). A review of urban energy systems at building cluster level incorporating renewable-energy-source (RES) envelope solutions. Applied Energy, 230, 1034–1056. https://doi.org/10.1016/j.apenergy.2018.09.041

[7] Singh, K., & Hachem-Vermette, C. (2021). Economical energy resource planning to promote sustainable urban design. Renewable and Sustainable Energy Reviews, 137, 110619. https://doi.org/10.1016/j.rser.2020.110619

[8] Fadaee, M., & Radzi, M. A. M. (2012). Multi-objective optimization of a stand-alone hybrid renewable energy system by using evolutionary algorithms: A review. Renewable and Sustainable Energy Reviews, 16(5), 3364–3369. https://doi.org/10.1016/j.rser.2012.02.071

[9] Karunathilake, H., Perera, P., Ruparathna, R., Hewage, K., & Sadiq, R. (2018). Renewable energy integration into community energy systems: A case study of new urban residential development. Journal of Cleaner Production, 173, 292 – 307. https://doi.org/10.1016/j.jclepro.2016.10.067

[10] Ghatikar, G., Mashayekh, S., Stadler, M., Yin, R., & Liu, Z. (2016). Distributed energy systems integration and demand optimization for autonomous operations and electric grid transactions. Applied Energy, 167, 432–448. https://doi.org/10.1016/j.apenergy.2015.10.117

[11] Faraji, J., Ketabi, A., Hashemi-Dezaki, H., Shafie-Khah, M., & Catalão, J. P. S. (2020). Optimal day-ahead self-scheduling and operation of prosumer microgrids using hybrid machine learning-based weather and load forecasting. IEEE Access, 8, 157284–157305. https://doi.org/10.1109/ACCESS.2020.3019562

[12] Khan, F. A., Pal, N., & Saeed, S. H. (2018). Review of solar photovoltaic and wind hybrid energy systems for sizing strategies optimization techniques and cost analysis methodologies. Renewable and Sustainable Energy Reviews, 92, 937 – 947. https://doi.org/10.1016/j.rser.2018.04.107

[13] Kuznetsova, E., Li, Y.-F., Ruiz, C., Zio, E., Ault, G., & Bell, K. (2013). Reinforcement learning for microgrid energy management. Energy, 59, 133–146. https://doi.org/10.1016/j.energy.2013.05.060

[14] Ji, Y., Wang, J., Xu, J., & Li, D. (2021). Data-driven online energy scheduling of a microgrid based on deep reinforcement learning. Energies, 14(8), 2120. https://doi.org/10.3390/en14082120

[15] Zhou, K., Zhou, K., & Yang, S. (2022). Reinforcement learning-based scheduling strategy for energy storage in microgrid. Journal of Energy Storage, 51, 104379. https://doi.org/10.1016/j.est.2022.104379

[16] Li, Y., Wang, R., & Yang, Z. (2022). Optimal scheduling of isolated microgrids using automated reinforcement learning-based multi-period forecasting. IEEE Transactions on Sustainable Energy, 13(1), 159–169. https://doi.org/10.1109/TSTE.2021.3105529

ACKNOWLEDGEMENTS

The authors would like to express their gratitude to all those who contributed to the development of this research. Special thanks are extended to the organizations and institutions that provided the necessary data and resources, enabling the comprehensive modeling and simulation of the distributed energy system. We also acknowledge the valuable support from the academic community in advancing the understanding of energy optimization techniques. Lastly, the authors appreciate the insightful feedback and suggestions that helped improve the quality of this study.

FUNDING

None.

AVAILABILITY OF DATA

Not applicable.

AUTHOR CONTRIBUTIONS

Jinkai Zhou: Conceptualization, Methodology, Formal Analysis, Writing – Original Draft, Visualization. Jinkai Zhou contributed to the design of the research framework, the development of the simulation environment, and the formulation of the deep reinforcement learning-based optimization algorithm. He was also responsible for conducting the analysis and preparing the initial manuscript.

Guoming Lin: Supervision, Writing – Review & Editing, Funding Acquisition. Guoming Lin provided overall supervision throughout the research process, reviewed and edited the manuscript, and contributed to securing resources and funding for the project.

COMPETING INTERESTS

The authors declare no competing interests.