



Community shared ES-PV system for managing electric vehicle loads via multi-agent reinforcement learning

Baligen Talihati^{a,b}, Shiyi Fu^{a,b}, Bowen Zhang^{a,b}, Yuqing Zhao^{a,b}, Yu Wang^{a,b,*}, Yaojie Sun^{a,b,*}

^a School of Information Science and Technology, Fudan University, 200433 Shanghai, China

^b Shanghai Engineering Research Center for Artificial Intelligence and Integrated Energy System, Fudan University, 200433 Shanghai, China

HIGHLIGHTS

- Community users and the ES-PV system cooperate and compete using MARL algorithms.
- MARL demonstrates significant strengths in coordinating multiple variables and systems in the community.
- ES-PV systems enhance the effectiveness of supporting community electric vehicle loads.
- ES-PV systems can be profitable and reduce community electricity costs.
- ES-PV systems can increase the self-consumption rate of photovoltaic systems.

ARTICLE INFO

Keywords:
 Multi-agent reinforcement learning
 Energy storage
 Photovoltaic
 Intelligent EV charging

ABSTRACT

The rapid growth of electric vehicles (EVs) is an unavoidable trend within the global energy transition. However, the substantial integration of EVs poses significant challenges to the stability and reliability of power systems. This study proposes mitigating EV load through community-shared energy storage and photovoltaic (ES-PV) systems. Within the framework of multi-agent reinforcement learning (MARL), multiple decision-making agents collaborate to manage various variables and systems in community, including energy storage charging and discharging strategies, intelligent EV charging strategies, and ES-PV system electricity pricing strategies. The coordination and optimization achieved through MARL enable these strategies to address the interdependences and dynamic changes of the variables, thereby enhancing overall performance. Case studies in real-world scenarios demonstrate that ES-PV systems can support up to 38.68 % of EV load, increase photovoltaic self-consumption rates by 66.41 %, and significantly reduce community reliance on the distribution grid. In terms of economic performance, implementing the ES-PV system reduced community electricity expenses by up to 7.73 %, resulting in a net profit of €51,924.65 for the ES-PV system in summer. This indicates a win-win solution for both community residents and ES-PV system operators. Therefore, this framework can support a more efficient and resilient community energy utilization paradigm, accommodating the increasing prevalence of EVs and the rapid development of smart communities.

1. Introduction

Electric vehicles (EVs), known for their high energy efficiency and negligible pollutant emissions, are rapidly expanding globally, providing a viable solution for developing cleaner transportation systems [1–5]. It is projected that by 2024, the number of EVs could reach approximately 17 million, accounting for over one-fifth of global car sales [6]. China is also significantly promoting the development of EVs,

with a market share exceeding 76 %, and the demand for new EVs among Chinese consumers continues to rise [7]. However, as the EV industry flourishes, the simultaneous and uncoordinated charging of numerous EVs could lead to the overloading of cables and transformers, placing increasing pressure on distribution networks [8].

To alleviate the pressure on distribution networks, vehicle-to-grid (V2G) technology is considered a feasible solution. This technology allows EVs not only to draw power from the grid but also to feed power

* Corresponding authors at: Department of Light Sources and Illuminating Engineering, Fudan University, Shanghai 200433, China.

E-mail addresses: y_wangfdu@fudan.edu.cn (Y. Wang), yjsun@fudan.edu.cn (Y. Sun).

back during peak demand periods, thereby balancing the grid load [9–13]. Additionally, vehicle-to-building (V2B) technology serves as an important approach, utilizing EVs as flexible energy storage resources to support building energy needs [14,15]. For instance, Kuang et al. analyzed the impact of driver behavior and building categories on the economic performance of V2G/V2B integration, providing valuable insights for balancing energy demand between the grid and EVs [16]. Similarly, Marcos et al. developed an optimal fleet dispatch management framework for an energy community, utilizing V2G capabilities as large-scale energy storage facilities [17]. Moreover, Babaei et al. incorporated V2G functionality into a demand response model to reduce energy costs and enhance overall performance, although their model did not account for the variability in EV charging demands at different times [18]. However, traditional distribution network designs are primarily based on single-source radial structures, with protection devices typically designed for unidirectional current flow [19]. Achieving widespread V2G adoption would require upgrading protection equipment to support bidirectional current, which presents significant technical and economic challenges. Therefore, despite the potential benefits of V2G technology, its implementation remains challenging. On the other hand, intelligent EV charging technology also presents a feasible solution to alleviate distribution network pressure by integrating flexible charging with other community load demands to promote grid stability [20–22]. This charging approach enables intelligent and proactive charging decisions, ensuring that each EV is fully charged before departure while avoiding grid overload [23–25]. For example, Van et al. applied model-free reinforcement learning to design EV charging strategies, enabling efficient operation in complex, large-scale environments without explicit knowledge of environmental dynamics or system parameters [26]. Lyu et al. demonstrated that flexible management of EVs through intelligent charging strategies is a key factor in maximizing social welfare [27]. Additionally, Huang et al. used genetic algorithms to determine the optimal charging/discharging rates for EVs, improving the match between demand and supply and reducing building power costs. Despite the advantages of intelligent charging, using EVs as energy storage resources may lead to battery degradation, posing a challenge for widespread adoption [28]. In [29], three different intelligent EV charging control methods were proposed to increase the photovoltaic power utilization in microgrids. The results showed that coordinated EV control could increase the self-consumption rate of photovoltaic power from 49 % to 62–87 %. However, relying solely on intelligent EV charging to shift fleet peak loads is insufficient to significantly reduce community dependence on the grid, thereby providing limited relief to the distribution network pressure.

Utilizing energy storage and photovoltaic (ES-PV) systems to meet the demand of EVs, combined with the flexible demand-shifting capabilities of intelligent EV charging, provides a new approach for community energy utilization paradigm. ES-PV systems use PV generation to meet a portion of the community's electricity demand, while storage batteries facilitate the absorption of solar power and can be flexibly dispatched according to demand and electricity prices, generating economic benefits [30]. Chen et al. optimized the operation strategy of grid-connected PV battery systems using reinforcement learning, studying the economic impacts of battery capacity, rooftop PV penetration, and electricity price fluctuations [31]. However, their study employed a simplistic linear model for battery modeling, resulting in a significant deviation in the battery charging strategy from real-world scenarios. Community energy utilization involves multiple variables and systems, such as energy storage charging strategies, intelligent EV charging strategies, and the uncertainties of solar power generation, all of which are interdependent and dynamically changing. Therefore, more coordinated control strategies are required. Due to the heterogeneity of scenarios, these control strategies must also be adaptable across different contexts. In this regard, Multi-Agent Reinforcement Learning (MARL) emerges as an effective control strategy. MARL employs a collaborative control approach, where multiple decision-making agents

interact cooperatively, excelling at managing various systems in complex environments, thus demonstrating significant potential in energy optimization [32]. For instance, Ye et al. utilized MARL to implement local energy trading market functions and provide flexibility services to broader system operators [33]. Wu et al. addressed the dynamic participation problem in energy management by learning control strategies through MARL's interaction with the environment [34]. Li et al. utilized Exploration-enhanced MARL to solve the scheduling problem of distributed PV-ES in smart distribution grids [35]. Si et al. employed MARL for distribution system restoration, contributing to enhancing the resilience of the distribution network [36]. Yu proposed an heating, ventilation, and air conditioning (HVAC) system control algorithm based on MARL and attention mechanisms, significantly enhancing energy efficiency [37]. MARL algorithms can effectively coordinate different units in energy dispatch problems, meet load demands, reduce operational costs, and maximize the self-utilization of renewable energy, all without requiring real-time communication [38]. In home energy management, MARL also achieves exceptional performance in economic and efficient scheduling [39]. Thus, the integration of ES-PV systems and MARL algorithms can provide a more efficient and intelligent solution for optimal energy utilization within a community, further reducing dependence on the grid. The summary of current studies and their limitations are detailed in Table 1.

Therefore, this paper proposes utilizing ES-PV systems to handle the increasing EV load via MARL algorithm. The efficiency and profitability of ES-PV systems under MARL guidance are systematically evaluated. Firstly, agent1 utilizes MARL algorithms, such as MAPPO, MADDPG, and MATD3, to optimize battery charging and discharging strategies in the face of uncertainties in community power demand, PV generation, and fluctuating electricity prices. Secondly, agent2 formulates intelligent EV charging strategies to shift peak EV demand. Thirdly, agent3 devises pricing strategies for selling electricity from ES-PV systems to the community. Finally, MARL coordinates the collaboration and competition among these three agents. Case studies are then conducted in real-world settings to validate this framework's impact on ES-PV system benefits and community economics, exploring efficiency changes based on PV self-consumption rates. Finally, the analysis extends to distribution grid nodes to explore whether this framework effectively alleviates the pressure on the distribution grid.

The structure of this paper is as follows: Section 2 describes the implementation of the EV-PS system in the community and the deployment of the MARL algorithm in detail. Section 3 outlines the

Table 1
Analysis of algorithm, battery model and limitations in existing literature.

Literature	Algorithm	Battery Model	Limitations
[1]	Multi-objective optimization	linear model	Without smart charging, the grid may become overloaded during peak periods.
[8]	MILP	linear model	Guides EV charging and discharging, but does not alleviate load pressure.
[16]	MILP	linear model	V2G is not suitable for traditional grids as they cannot handle the reverse current from EVs.
[18]	Data-mining	linear model	Does not account for the variability in user charging demands at different times.
[27]	DC-ADMM	linear model	Smart charging can shift EV demand, but its effectiveness in reducing fleet dependence on the grid is limited.
[28]	GA	linear model	Using EVs as energy storage units can impact their battery lifespan.
[31]	DRL	linear model	DRL struggles to achieve effective agent coordination and collaboration.

experimental steps for obtaining parameters and presents detailed case studies to illustrate the practical application of our approach. Section 4 discusses the results of the case studies. Finally, Section 5 summarizes our findings and offers suggestions for future research directions.

2. Methodology

Fig. 1 summarizes the research framework. This paper models a community energy utilization paradigm at IEEE community nodes, integrating a community-shared ES-PV system with energy storage and photovoltaic technologies to enhance the stability and efficiency of local electricity supply, particularly addressing the rising demand from EVs. The study employs a MARL approach for control strategies, encompassing battery charge-discharge management, electric vehicle charging scheduling, and pricing strategies for ES-PV systems. The first step in the research involves data integration, which includes both real-world data and data generated by the MARL algorithm during the decision-making process. In the second step, a multi-constraint battery model is incorporated as part of the ES-PV system to ensure that battery modeling closely reflects real-world conditions. The third step involves the interaction of three agents within the MARL framework with the environment to generate optimal decisions. Finally, the fourth step is to validate the effectiveness of the proposed model.

The Residential IEEE-14 distribution network depicted in **Fig. 2** serves as the foundational framework for this study. Given our focus on utilizing intelligent management of ES-PV systems to address the challenges posed by EVs at the community level, its implementation within

the IEEE-14 distribution network nodes is essential. The structure of the IEEE-14 distribution network is well-suited to meet the requirements of community energy utilization paradigm. As a widely used standard model, the IEEE-14 distribution network has been employed in power research for many years [40–42]. It provides a simple yet effective testing environment, enabling us to validate the effectiveness and stability of the MARL approach without introducing unnecessary complexity. This network exemplifies an intelligent community comprising ES-PV systems, residential loads, and electric vehicles, laying a solid foundation for further research expansion.

2.1. Multi-Agent Reinforcement Learning (MARL) application

Our main objective is to establish an ES-PV system to manage the EV load within the community based on its electricity demand, solar PV generation, and time-of-use electricity prices. This process requires us to achieve economically efficient battery charging and discharging, and to determine pricing strategies that enable the ES-PV system and the community to be profitable. This leads to a complex and dynamic interaction process, which involves both cooperation and competition. Multi-agent learning emphasizes the complex dynamics and interdependence in shared environments, highlighting the unique challenges and opportunities for collaboration in these complex systems [32]. Therefore, multi-agent learning is an ideal choice for this work. Additionally, to achieve this goal, we have implemented ES agent, EV agent, and Price agent.

Additionally, the management of battery charging and discharging

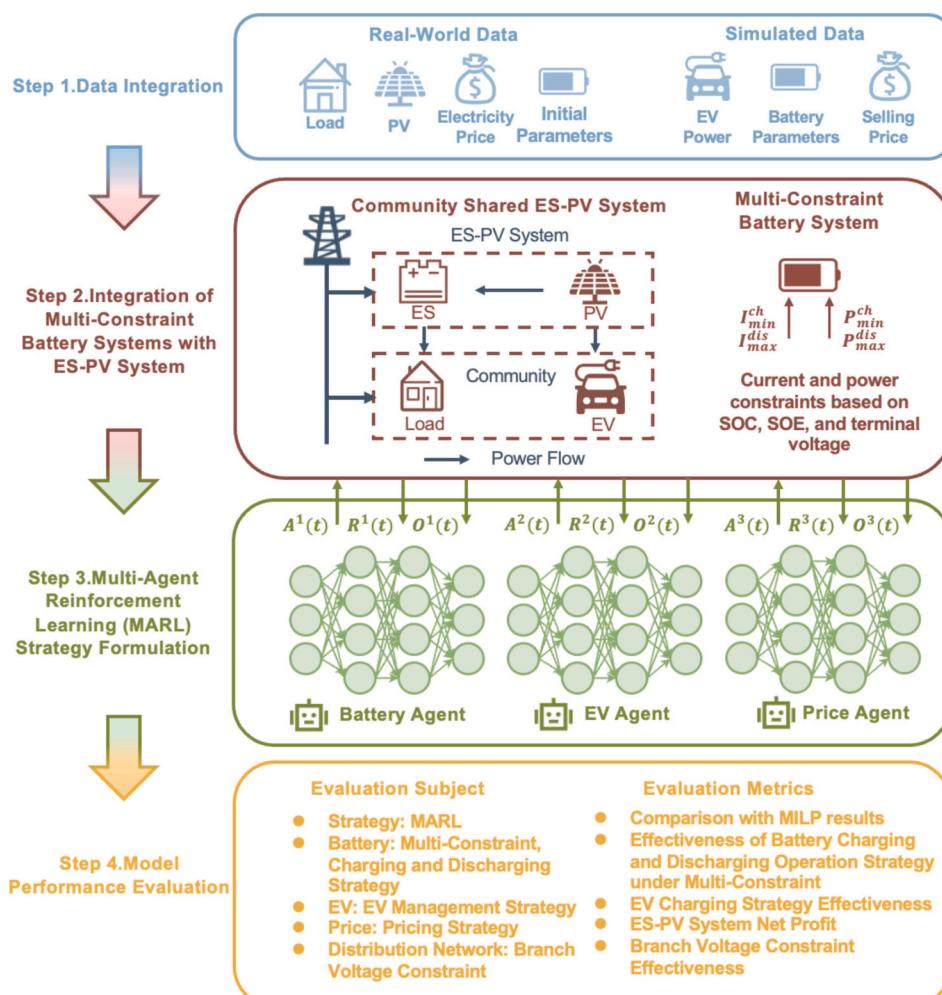


Fig. 1. Framework of the research.

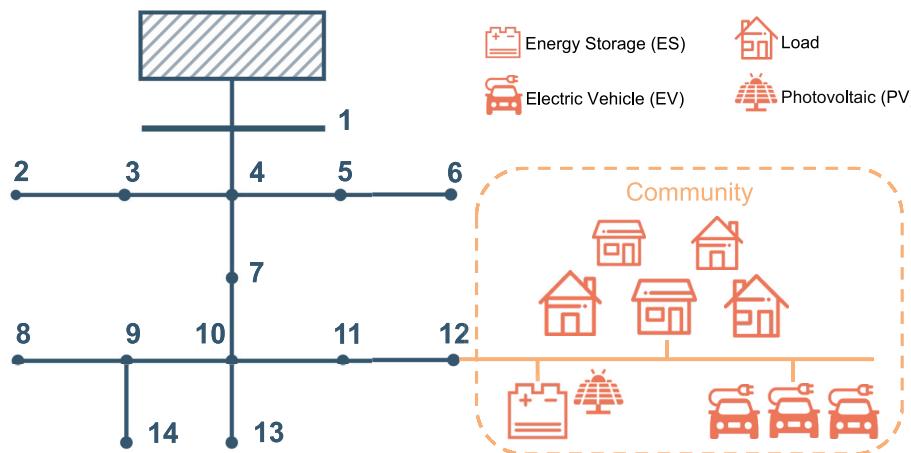


Fig. 2. The residential IEEE-14 distribution network.

can be formulated as a Markov decision process (MDP), which can be modeled using reinforcement learning. MDP comprises five essential elements (S, A, P, R, γ), representing state, action, transition probability, reward, and discount factor, respectively.

2.1.1. Definition of action space

The action space A represents the decisions made by MARL in interaction with the environment, all of which are continuous values. Specifically, A^{bat} represents the decisions of the ES agent, with an action space of $[-1, 1]$, where negative values indicate charging and positive values indicate discharging. A^{ev} represents the decisions of the EV agent, with an action space of $[0, 1]$. A^{pri} represents the decisions of the Price agent, with an action space of $[0.8, 1.2]$.

2.1.2. Definition of state space and observation space

The state space S represents the useful information available for making decisions in the environment. In this work, the state space includes the community's base load power, time-of-use electricity prices, solar power generation, battery SOC, EV power, selling price, EV penalty, and power penalty. The community's base load power, time-of-use electricity prices, and solar power generation are derived from the real world, providing real environmental information for the agents' decision-making.

The observation space O consists of the useful information each agent needs for decision-making. Each agent's observation space O is derived from the state space S .

2.1.3. Definition of transition probability

State transition P represents the subsequent state after executing an action A . In this work, the updates involve battery SOC, EV power, selling price, EV penalty, and power penalty. These updates are based on the interactions between MARL and the environment.

2.1.4. Definition of reward function

The reward function in this work consists of both individual and shared reward function. The ES agent and EV agent share the same reward function. This design is based on the fact that EVs are part of the community load, and both agents aim to reduce the community's electricity costs. On the other hand, the Price agent has its individual reward function, focusing solely on the interests of the ES-PV system by setting the electricity selling price. Therefore, the Price agent aims to maximize the ES-PV system's profitability. This leads to a complex coordination and cooperation process among the three agents, as the community desires to purchase electricity from the ES-PV system at a lower price to minimize its electricity costs, while the ES-PV system aims to sell as much electricity as possible to the community. This requires the pricing

to be competitive while also ensuring that the price is not too low to maximize profitability.

2.2. Introduction of ES-PV System

The ES-PV system within the community optimizes the charging and discharging power of the battery based on the community's electricity demand, solar PV generation, and time-of-use electricity prices using a MARL ES agent. This optimization aims to minimize electricity costs.

Given the significant EV power and uncertain EV charging times, they play a crucial role in the community's electricity demand. The EV power in the smart charging mode is optimized by the EV agent.

The ES-PV system sells electricity to the community at a selling price, which is determined by the Price agent. This pricing strategy aims to benefit the ES-PV system through peak shaving and valley filling, while also benefiting community residents by allowing them to purchase electricity from the ES-PV system at a reasonable price.

2.3. Battery model

The enhanced self-correcting (ESC) battery model (Fig. 3c) is chosen for its precision, reliability, and excellent performance in handling large datasets and complex computations [43]. This makes it highly suitable for integration into the large-scale ES system depicted in Fig. 3b. Consequently, our overall battery modeling strategy involves detailed modeling of the battery pack, followed by integration into the ES system, and ultimately deployment of the ES system at IEEE-14 distribution network nodes. Next, we will establish various constraints for the battery pack based on the ESC battery model to ensure that our work simulates performance close to real-world conditions and maintains safety standards.

The electrical behaviour of the ESC model can be expressed by:

$$\begin{cases} \dot{U}_p = -\frac{1}{C_p R_p} U_p + \frac{1}{C_p} I_t \\ U_t = U_{ocv} - U_p - I_t R_0 + U_{hyst} \end{cases} \quad (1)$$

U_{ocv} represents the open-circuit voltage (OCV) of the battery. Fig. 4a shows the OCV curves during charging and discharging. As illustrated in the figure, the OCV under charging conditions is significantly higher than that under discharging conditions. Ignoring this hysteresis loop may lead to substantial errors. Therefore, it is imperative to include the hysteresis voltage in the battery modeling. U_{hyst} represents the hysteresis voltage of the battery. R_p and C_p form an RC circuit to describe the voltage polarization in the battery, a phenomenon caused by the slow diffusion processes of lithium in a lithium-ion cell. U_p is the voltage polarization of the battery through the RC network. R_0 models the ohmic

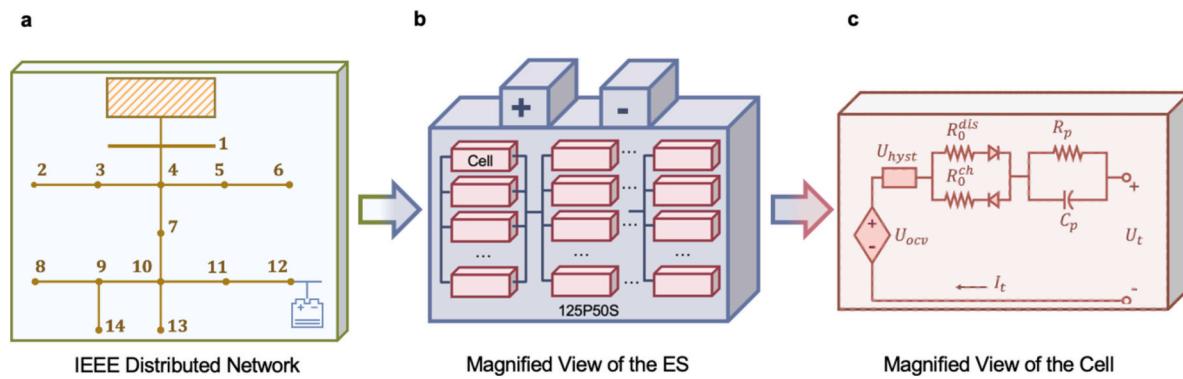


Fig. 3. a. Residential IEEE-14 distribution network. b. The energy storage (ES) system. c. The enhanced self-correcting (ESC) cell model for batteries.

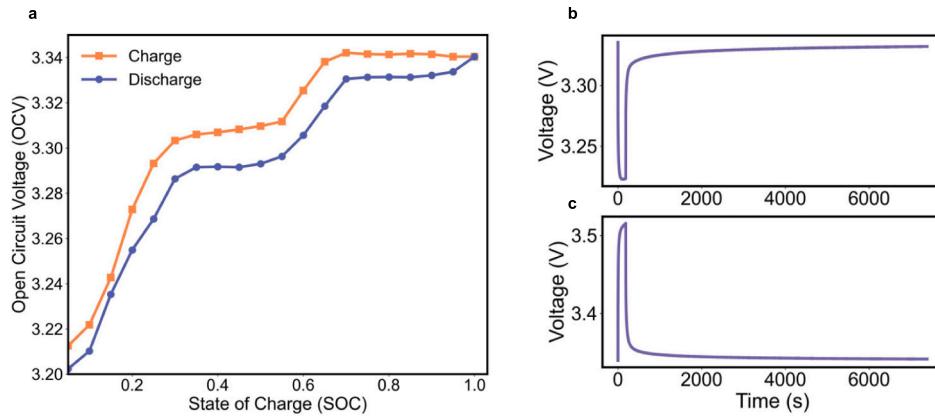


Fig. 4. a. The open-circuit voltage of a lithium iron phosphate battery. b. Discharge phase hybrid pulse power characterization (HPPC) voltage profile. c. Charge phase HPPC voltage profile.

resistance of the battery, denoted respectively by R_0^{dis} for discharge resistance and R_0^{ch} for charge resistance. I_t denotes the battery current, positive for discharge and negative for charge. U_t represents the terminal voltage of the entire model. The internal battery parameters such as R_0 , R_p , and C_p vary with SOC. In this study, the MARL method necessitates a continuous action space, requiring internal parameters to be fit to a polynomial representation in Equation (2). These battery parameters were obtained from Hybrid Pulse Power Characterization (HPPC) tests, yielding HPPC small cycle voltage plots shown in Fig. 4b and c. Curve fitting was performed using the ‘curve fit’ function from the ‘scipy.optimize’ library in Python to derive polynomial parameters as listed in Table 2.

$$R_0, R_p, C_p = a_0 + a_1 z + a_2 z^2 + a_3 z^3 + a_4 z^4 + a_5 z^5 + a_6 z^6 + a_7 z^7 \quad (2)$$

2.3.1. Hysteresis voltages

The change in hysteresis voltage can be modeled as a function of the change in SOC [43], which is transformed into a discrete-time difference equation for application to hysteresis voltage as follows:

$$U_{hyst,k} = A_H U_{hyst,k-1} - (1 - A_H) \text{sgn}(I_{t,k-1}) \quad (3)$$

$$A_H = \exp\left(-\left|\frac{\eta I_{t,k-1} \gamma \Delta t}{Q}\right|\right) \quad (4)$$

where $U_{hyst,k}$ and $U_{hyst,k-1}$ are the hysteresis voltages at the k th and $(k-1)$ th sampling times respectively, $I_{t,k-1}$ is the current flowing through the battery at the $(k-1)$ th sampling time. The term has a positive constant γ , which tunes the rate of decay.

2.3.2. Diffusion voltages

We derive an expression for the resistor current $I_{p,t}$ as follows. First, we acknowledge that the sum of the current through R_p and the current through C_p must equal I_t . Additionally, $I_{p,t} = C_p \dot{U}_{p,t}$, and since $U_{p,t} = R_p I_{p,t}$, we can derive a discrete-time equation for the resistor current as follows:

$$I_{p,k} = F_p I_{p,k-1} + (1 - F_p) I_{t,k-1} \quad (5)$$

Table 2
Parameters of battery internal characteristics and open-circuit voltage.

Condition	Parameter	a_0	a_1	a_2	a_3	a_4	a_5	a_6	a_7
Discharging	$R_0(\Omega)$	0.003	0.025	- 0.29	1.34	- 3.03	3.62	- 2.18	0.52
	$R_p(\Omega)$	0.007	- 0.08	0.58	- 1.98	3.53	- 3.24	1.36	- 0.17
	$C_p(F)$	1.17^5	- 3.64 ⁶	4.18 ⁷	- 2.06 ⁸	5.16 ⁸	- 6.81 ⁸	4.53 ⁸	- 1.19 ⁸
	$R_0(\Omega)$	0.0035	- 0.029	0.33	- 1.75	4.69	- 6.67	4.77	- 1.35
Charging	$R_p(\Omega)$	0.003	- 0.024	0.29	- 1.71	5.13	- 7.97	6.12	- 1.83
	$C_p(F)$	7.64^4	- 8.28 ⁵	1.29 ⁷	- 7.93 ⁷	2.34 ⁸	- 3.55 ⁸	2.67 ⁸	- 7.86 ⁷
OCV	U_{ocv}	3.36	- 0.02	7.73	0.06	0.0002	/	/	/

$$F_p = \exp\left(-\frac{\Delta t}{R_p C_p}\right) \quad (6)$$

2.3.3. Open-Circuit Voltages (OCV)

The OCV U_{ocv} represents the average voltage of the OCV during charging and discharging:

$$U_{ocv} = (U_{ocv}^{dis} + U_{ocv}^{ch})/2 \quad (7)$$

According to reference [44], the polynomial fitting function for the OCV can be expressed as follows:

$$U_{ocv} = a_0 + a_1 z + \frac{a_2}{z} + a_3 \ln(z) + a_4 \ln(1-z) \quad (8)$$

z represents the SOC, and a_0 to a_4 are the polynomial coefficients obtained by fitting the OCV curve. These coefficients are listed in Table 2.

2.3.4. SOC constraint current

As mentioned earlier, we have $\frac{dz}{dt} = -\eta I_t Q$, thus the discrete-time formula for SOC is as follows:

$$z_k = z_{k-1} - \frac{\Delta t}{Q} \eta I_{t,k-1} \quad (9)$$

where z_k and z_{k-1} are the SOC at the k th and $(k-1)$ th sampling times. According to Equation(11), assuming the SOC is constant, the continuous minimum charging current and maximum discharging current, constrained by SOC, can be expressed as follows:

$$\begin{cases} I_{min}^{ch,SOC} = \frac{(z_k - z_{max})Q}{\eta \Delta t} \\ I_{max}^{dis,SOC} = \frac{(z_k - z_{min})Q}{\eta \Delta t} \end{cases} \quad (10)$$

$I_{min}^{ch,SOC}$ represents the minimum charging current, $I_{max}^{dis,SOC}$ represents the maximum discharging current.

2.3.5. SOE constraint current

The State of Energy (SOE) is a crucial quantity in battery models and therefore, its constraint is particularly important. According to reference [45], the formula for SOE is as follows:

$$s_k = s_{k-1} - \frac{\Delta t}{E_N} \eta_e U_{t,k-1} I_{t,k-1} \quad (11)$$

where s_k and s_{k-1} are the SOE at the k th and $(k-1)$ th sampling times respectively. η_e represents the energy efficiency of the battery, and E_N represents available energy of the battery. Then, the continuous minimum charging current and maximum discharging current are expressed as follows, constrained by SOE:

$$\begin{cases} I_{min}^{ch,SOE} = \frac{(s_k - s_{max})E_N}{\eta_e U_{t,k+1} \Delta t} \\ I_{max}^{dis,SOE} = \frac{(s_k - s_{min})E_N}{\eta_e U_{t,k+1} \Delta t} \end{cases} \quad (12)$$

$I_{min}^{ch,SOE}$ represents the minimum charging current, $I_{max}^{dis,SOE}$ represents the maximum discharging current, and $U_{t,k+1}$ is the terminal voltage at the $(k+1)$ th sampling time.

2.3.6. Voltage constraint current

From the expression in Equation (1), we can derive the terminal voltage at the $(k+1)$ th sampling time. The OCV at the $(k+1)$ th sampling time can also be calculated using the first-order Taylor expansion formula. It is important to note that $R(\cdot)$ represents the remainder term in the Taylor series expansion:

$$\begin{aligned} U_{ocv,k+1} &= f\left(z_k - I_{t,k} \frac{\Delta t \eta}{Q}\right) \\ &= U_{ocv,k} - I_{t,k} \frac{\Delta t \eta}{Q} \frac{\partial U_{ocv}}{\partial z} \Big|_{z=z_k} - R\left(z_k, I_{t,k} \frac{\Delta t \eta}{Q}\right) \approx U_{ocv,k} - I_{t,k} \frac{\Delta t \eta}{Q} \frac{\partial U_{ocv}}{\partial z} \Big|_{z=z_k} \end{aligned} \quad (13)$$

Combining the above expressions, we can obtain the formula for the current under the constraint of terminal voltage:

$$\begin{cases} I_{min}^{ch,volt} = \frac{U_{ocv,k} - U_{p,k} F_p + U_{hyst,k} - U_{t,max}}{\frac{\Delta t \eta}{Q} \frac{\partial U_{ocv}}{\partial z} \Big|_{z=z_k}} + R_p(1 - F_p) + R_0 \\ I_{max}^{dis,volt} = \frac{U_{ocv,k} - U_{p,k} F_p + U_{hyst,k} - U_{t,min}}{\frac{\Delta t \eta}{Q} \frac{\partial U_{ocv}}{\partial z} \Big|_{z=z_k}} + R_p(1 - F_p) + R_0 \end{cases} \quad (14)$$

2.3.7. Current constraint

Combining equations (10), (12)and (14), we can derive the formula for the charging and discharging currents under the constraint of terminal voltage. Here, I_{min} and I_{max} refer to the current limits specified by the battery manufacturer:

$$\begin{cases} I_{min}^{ch} = \max(I_{min}^{ch,SOC}, I_{min}^{ch,SOE}, I_{min}^{ch,volt}, I_{t,k-1}^{ch}, I_{min}) \\ I_{max}^{ch} = \min(I_{max}^{dis,SOC}, I_{max}^{dis,SOE}, I_{max}^{dis,volt}, I_{t,k-1}^{dis}, I_{max}) \end{cases} \quad (15)$$

2.3.8. Power constraint

In summary, the charging and discharging power formula for a battery under power constraints can be expressed as follows, where P_{min} and P_{max} represent the power limits specified by the battery manufacturer:

$$\begin{cases} P_{min}^{ch} = \max(P_{min}, U_{t,k+1} \max(I_{min}^{ch,SOC}, I_{min}^{ch,SOE}, I_{min}^{ch,volt}, I_{min})) \\ P_{max}^{ch} = \min(P_{max}, U_{t,k+1} \min(I_{max}^{dis,SOC}, I_{max}^{dis,SOE}, I_{max}^{dis,volt}, I_{max})) \end{cases} \quad (16)$$

2.3.9. Battery Penalty

In order to achieve effective management of the charging and discharging processes of energy storage batteries using MARL, it is imperative to establish specific constraints on each battery pack. This necessitates the implementation of penalty mechanisms, thereby enhancing the management of the batteries and prolonging their operational lifespan.

$$P_k^{bat} = (z_{k-1} - z_k) E_N \quad (17)$$

$$\begin{cases} \text{Penalty}_t^{es,z} = 0 \text{ } z_{min} \leq z_k \leq z_{max} \\ \text{Penalty}_t^{es,z} = 10 \text{ } \text{otherwise} \end{cases} \quad (18)$$

$$\begin{cases} \text{Penalty}_t^{es,p} = 0 \text{ } P_{min}^{ch} \leq P_k^{bat} \leq P_{max}^{dis} \\ \text{Penalty}_t^{es,p} = 10 \text{ } \text{otherwise} \end{cases} \quad (19)$$

$$\begin{cases} \text{Penalty}_t^{es,i} = 0 \text{ } I_{min}^{ch} \leq I_k \leq I_{max}^{dis} \\ \text{Penalty}_t^{es,i} = 10 \text{ } \text{otherwise} \end{cases} \quad (20)$$

$$\text{Penalty}_t^{es} = \text{Penalty}_t^{es,z} + \text{Penalty}_t^{es,p} + \text{Penalty}_t^{es,i} \quad (21)$$

Equation (17) provides the power of the battery. Equation (18) defines the SOC constraints, stipulating that the SOC range must be between 0.05 and 0.95, beyond which a penalty is applied. This prevents overcharging and deep discharging, thereby protecting the battery's lifespan. Equation (19) imposes certain power limits on the battery, while Equation (20) sets the current constraints for the battery. Equation (21) represents the total penalty for the battery pack, which will be

incorporated into the MARL algorithm. Consequently, during the iterative process, the algorithm will progressively avoid actions that trigger these penalties, thereby achieving safe battery management.

2.4. Electric Vehicles (EVs) Model

Due to its rapid growth, EVs have become an indispensable load in communities. EVs in the community follow a home charging mode, where owners start charging when they return home from work and stop charging when they leave for work. The start time of EVs charging follows a normal distribution [46], from which the EVs charging schedules in the community can be obtained.

$$f_t(x) = \begin{cases} \frac{1}{\sigma_s \sqrt{2\pi}} \exp\left(-\frac{(x - \mu_s)^2}{2\sigma_s^2}\right) & \mu_s - 12 < x < 24 \\ \frac{1}{\sigma_s \sqrt{2\pi}} \exp\left(-\frac{(x + 24 - \mu_s)^2}{2\sigma_s^2}\right) & 0 < x < \mu_s - 12 \end{cases} \quad (22)$$

The capacity of EVs in the community follows the distribution described in Equation (23), where 20-60 kWh represents the common capacity range of EVs in the market. The probability of EVs in the community having capacities within this range is uniform:

$$f_c(x) = \begin{cases} \frac{1}{40} & 20 < x < 60 \\ 0 & \text{others} \end{cases} \quad (23)$$

The charging process of EVs is as follows:

$$SOC_{i,t+1} = SOC_{i,t} + \frac{P_{i,t}^{ev}}{Q_{i,t}^{ev}} \Delta t \quad (24)$$

$$SOC_{min} \leq SOC_{i,t} \leq SOC_{max} \quad (25)$$

Among them, $P_{i,t}^{ev}$ is the EV charging power, and $Q_{i,t}^{ev}$ is the EV battery capacity, determined by Equation (23). SOC_{min} ensures that the EV is charged to a minimum SOC required for operation, while SOC_{max} is set to avoid overcharging the EV.

Consider two charging modes for EVs, namely fast charging mode and smart charging mode. The fast charging mode provides a shorter charging time, suitable for residents in urgent need of charging. However, the high current and voltage of fast charging can affect the battery life. Therefore, smart charging mode is provided for residents who are not in a hurry to charge.

$$\begin{cases} P_{i,t}^{ev} = P^{ev,max} & \text{quickcharging} \\ 0 \leq P_{i,t}^{ev} \leq P^{ev,max} & \text{smartcharging} \end{cases} \quad (26)$$

In the fast charging mode, the charging process is completed using the fast charging power $P^{ev,max}$. In the smart charging mode, the charging power $P_{i,t}^{ev}$ of the EV is generated by the EV agent, i.e., $P_{i,t}^{ev} = A^{ev} P^{ev,max}$, where A^{ev} is a continuous action in the range of 0 to 1.

Whether the vehicles in the community use maximum power charging or smart charging depends on the general charging habits of community users. Equation (22) produces the charging schedule. Generally, vehicles starting to charge in the evening have a high probability of choosing smart charging, as there is usually no need to use the vehicle in the middle of the night. Therefore, the probability of selecting smart charging is set to 0.8 for nighttime charging. In contrast, vehicles charging during the day are likely to use maximum power charging, with the probability set to 0.5. These charging probabilities can also be set by users themselves, considering that some users may reject smart charging, and thus 5% of the vehicles are always set to select fast charging only.

$$P_t^{fleet} = \sum_{i=1}^m P_{i,t}^{ev} \quad (27)$$

The fleet power is the total power of all EVs. This power is considered as part of the community load and will be incorporated into the MARL along with other community loads to calculate the MARL reward. Thus, under the guidance of this reward function, this study aims to align EV charging as closely as possible with community users' habits, while also using the smart charging strategies determined by the EV agent to adjust the EV charging load curve to some extent, thereby mitigating the pressure on the distribution grid during peak community load times.

$$\begin{cases} \text{Penalty}_t^{ev} = 10 \text{ chargingduration} > h^{max} \\ \text{Penalty}_t^{ev} = 0 \text{ } 0 \leq \text{chargingduration} \leq h^{max} \end{cases} \quad (28)$$

This is the penalty calculation for the fleet, which will also be passed to the MARL for subsequent reward calculation. During the fleet charging process, SOC_{min} ensures that the fleet is charged to a sufficient level for use. While the fast charging mode completes in a very short time, in the smart charging mode, to avoid excessively low power, which would lead to long charging times, an EV penalty is set.

2.5. Community power balance

This section primarily integrates the previously developed battery and EV models into the community to achieve power balance. This integration aims to optimize the utilization of the community-shared ES-PV system, effectively addressing the community's power demand and enhancing the overall benefits for both the community residents and the ES-PV system.

Fig. 5a illustrates the power balance within the community. As shown, the community-shared ES-PV system consists of energy storage and solar power components. The electricity generated by the system is shared among community residents. The energy storage can store solar power and supply it to community residents at a certain price during periods of high electricity prices or power outages. In addition, the ES-PV system plays an important role by utilizing its energy storage characteristics to further delay part of the EV load curve, thereby alleviating the pressure that EV loads impose on the distribution grid. This function operates synergistically with the load adjustments previously made by the EV smart charging agent.

It is important to clarify that the ES-PV system sells electricity to the community at a certain price, while internal electricity usage within the system is free of charge. If the energy storage system's charging demand exceeds the supply from solar generation, additional electricity must be purchased from the grid.

Clearly, the community's total load cannot be fully covered by the ES-PV system, and therefore, the community still needs to purchase electricity from the grid. However, any surplus electricity in the ES-PV system cannot be fed back into the grid, leading to potential wastage. The community's total load consists of base loads (such as household appliances) and EV loads. As shown in **Fig. 5b**, the demand information, energy information, and other relevant data from the community and the ES-PV system are transmitted to the MARL agent for processing, which then issues decisions to guide the community's energy utilization.

$$Load_t = L_t + P_t^{fleet} \quad (29)$$

$$\text{Penalty}_t^{waste} = P_t^{waste} \quad (30)$$

$$G_t = P_t^{p,c} + P_t^{p,e} + P_t^{pv,waste} \quad (31)$$

$$Load_t = P_t^{p,c} + P_t^{dis} + P_t^{g,c} - P_t^{e,waste} \quad (32)$$

$$P_t^{ch} = P_t^{g,e} + P_t^{p,e} \quad (33)$$

$$P_t^{waste} = P_t^{pv,waste} + P_t^{e,waste} \quad (34)$$

$$P_t^{e,c} = P_t^{p,c} + P_t^{dis} \quad (35)$$

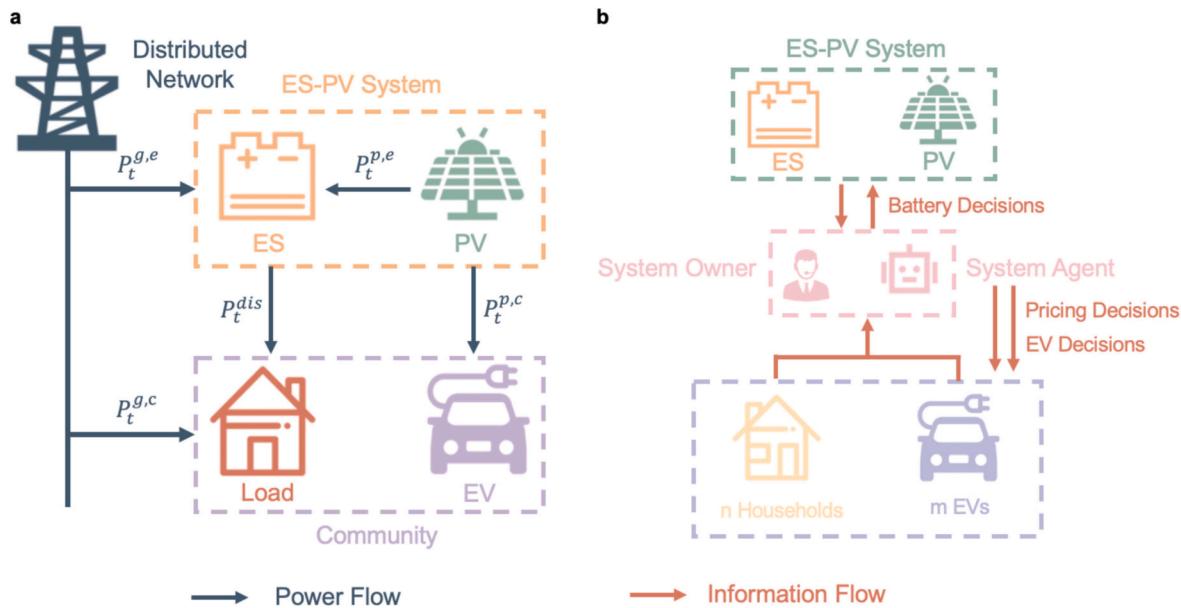


Fig. 5. a. Community power balance diagram; b. Community information exchange diagram.

In Equation (29), $Load_t$ represents the total load of the community, composed of the community's general load L_t (derived from real-world load data) and the electric vehicle load P_t^{fleet} . Equation (30) addresses the wastage penalty. As previously mentioned, any surplus energy in the ES-PV system that cannot be fed back into the grid is wasted. To minimize this wastage, a penalty is imposed; the more energy wasted, the greater the penalty. This penalty is also incorporated into the reward calculation.

Equation (31) represents the power balance equation for solar generation G_t . The power generated by the photovoltaic system is allocated as $P_t^{p,c}$ for the community, for which a certain fee is charged. Another portion is allocated as $P_t^{p,e}$ for the energy storage system, which is free of charge since it constitutes an internal transaction within the ES-PV system. If any surplus power remains, it is denoted as $P_t^{pv,waste}$.

Equation (32) represents the energy balance on the community side. The total load of the community $Load_t$ is supplied by the power generated from the photovoltaic system ($P_t^{p,c}$), the discharge power from the battery (P_t^{dis}), and the power purchased from the grid ($P_t^{g,c}$). If the discharge from the battery storage system exceeds what the load can absorb, the surplus is denoted as P_t^{waste} . This value is positive, indicating that some energy from the battery storage is wasted.

Equation (33) represents the charging power of the battery (P_t^{ch}), which is met through the power supplied by the photovoltaic system ($P_t^{p,e}$) and the power purchased from the grid ($P_t^{g,e}$). Equation (34) represents the total wasted energy in the ES-PV system. Equation (35) represents the total power ($P_t^{e,c}$) sold by the ES-PV system to the community.

2.6. Pricing model

In this section, we provide auxiliary strategies for the price agent using specific formulas. The price agent is integrated as a part of the MARL system because the ES-PV system needs to sell electricity to the community, and the current MARL manages battery charging and discharging as well as smart electric vehicle charging, which affects price setting. Therefore, it is essential to include this aspect in the model. We employ a Stackelberg game in this part, where the ES-PV system acts as the leader by setting the price to maximize its profit, while other users act as followers by responding to the price set by the leader.

$$Gain_t^p = \Pi_{t+1}^{sell} P_t^{e,c} \quad (36)$$

$$Cost_t^p = \Pi_t^* P_t^{g,e} \quad (37)$$

$$Reward_t^p = Gain_t^p - Cost_t^p \quad (38)$$

The leader, which is the ES-PV system, aims to maximize the profit $Reward_t^p$ by optimally setting the price Π_{t+1}^{sell} . The followers (other community residents) will adjust their demand $P_t^{g,c}$ based on the price Π_{t+1}^{sell} set by the leader. If the energy storage in the ES-PV system requires electricity from the grid, represented as $P_t^{g,e}$, this portion of electricity should be purchased at the time-of-use price Π_t^* .

2.7. Power flow constraints

This work aims to use the ES-PV system to shoulder the community load, thereby minimizing the stress on the distribution grid. Ultimately, our goal is to integrate the entire community and the ES-PV system into the IEEE 14-bus distribution network, implemented using the PYPOWER library in Python, and apply certain penalties to ensure that the distribution grid's loading stress remains within acceptable limits [47].

$$Penalty_t^{power} = \sum_i \begin{cases} |V_i - V_{min}|, & \text{if } V_i < V_{min} \\ |V_i - V_{max}|, & \text{if } V_i > V_{max} \end{cases} \quad (39)$$

$Penalty_t^{power}$ in equation (39) is the voltage penalty. For the voltage magnitude in the distribution network, we set a lower limit V_{min} and an upper limit V_{max} . If the voltage at the node where the community is located exceeds these limits, a penalty will be recorded.

2.8. Community reward

Previously, we introduced the reward function for the price agent, which focuses solely on the benefits of the ES-PV system. However, for the community, the overall community reward (benefit) is more important. Therefore, this section addresses the reward for both the ES agent and the EV agent. Specifically, the ES agent and EV agent share the same reward function.

$$Gain_t^s = \Pi_t^{sell} P_t^{e,c} \quad (40)$$

$$Cost_t^s = \Pi_t^* (P_t^{g,c} + P_t^{g,e}) \quad (41)$$

$$Cost_t^{wear} = \frac{C_E}{\eta D L_C} |P_t^{bat}| \quad (42)$$

$$\begin{aligned} Reward_t^s = & Gain_t^s - Cost_t^s - Cost_t^{wear} - Penalty_t^{power} - Penalty_t^{es} - Penalty_t^{ev} \\ & - Penalty_t^{waste} \end{aligned} \quad (43)$$

Equation (40) represents the profit of the ES-PV system, which is also the profit of the entire community, similar to the previous Equation (36). Equation (41) details the total electricity cost for the community, including the costs incurred by both the ES-PV system and the community purchasing electricity.

In Equation (42), the term $Cost_t^{wear}$ refers to the wear cost of the battery at time t . C_E is the unit energy cost of the battery, which is either the initial cost or the replacement cost of the battery. D is the depth of discharge of the battery, representing the extent of discharge. L_C is the cycle life of the battery, indicating the number of charge-discharge cycles the battery can undergo under specific depth of discharge and conditions. This formula indicates that the wear cost of the battery is related to the unit energy cost, efficiency, depth of discharge, cycle life, and charging/discharging power of the battery. By using this formula, the wear cost of the battery under different charging and discharging conditions can be calculated, providing a basis for battery management and optimization.

Equation (43) represents the reward function for the entire community, which is a crucial basis for guiding the decision-making of the ES agent and EV agent. The actions of the ES agent determine the extent of battery charging and discharging, while the EV agent decides the smart charging power of the electric vehicles. These actions will be adjusted based on the reward function, influencing the strategies and ultimately making decisions that are beneficial to the interests of the community.

2.9. Objective optimization algorithm

In this subsection, this study incorporates an objective optimization algorithm for comparison with the results obtained from the MARL approach. A Mixed-Integer Linear Programming (MILP) algorithm is selected due to its high accuracy and effectiveness in solving large-scale problem [48]. However, in practical computation, the equations of this algorithm differ slightly from those of the MARL approach. Therefore, this section will outline and explain these differences, while the common aspects will not be elaborated extensively.

2.9.1. Battery Model

The battery model employed is a classic linear model. This simplified approach abstracts the concept of individual battery packs within the actual energy storage system, instead modeling the entire storage system as a single, large battery pack. The advantage of this assumption is its ability to significantly reduce computational load and provide a rough estimation in many planning scenarios. However, its drawback lies in potential discrepancies with the real state of the battery system in actual control or simulation scenarios, which may result in significant errors in the optimization results.

The choice of this modeling approach in MILP is due to the inability of MILP to process the battery model presented in Section 2.3, which involves 6,250 battery packs and requires planning across 168 time points (equivalent to 7 days of data). Such complexity requires an algorithm capable of handling intricate environments, and MILP cannot effectively plan for this level of battery modeling. Nevertheless, despite the differences between this battery model and that used in MARL, a rough comparison of results between the two algorithms remains feasible, provided these differences are understood.

$$S_{d,t+1} = S_{d,t} + \left(\eta P_{d,t}^{ch} - \frac{P_{d,t}^{dis}}{\eta} \right) \Delta T \quad \forall t \in T \quad (44)$$

$$S_d^{min} \leq S_{d,t} \leq S_d^{max} \quad \forall t \in T \quad (45)$$

$$P_{d,t} = P_{d,t}^{ch} - P_{d,t}^{dis} \quad \forall t \in T \quad (46)$$

$$s_t^{ch} P_d^{ch,min} \leq P_{d,t}^{ch} \leq s_t^{ch} P_d^{ch,max} \quad \forall t \in T \quad (47)$$

$$s_t^{dis} P_d^{dis,min} \leq P_{d,t}^{dis} \leq s_t^{dis} P_d^{dis,max} \quad \forall t \in T \quad (48)$$

$$s_t^{ch} + s_t^{dis} \leq 1 \quad \forall t \in T \quad (49)$$

$$s_t^{ch}, s_t^{dis} \in \{0, 1\} \quad \forall t \in T \quad (50)$$

Equation (44) describes the battery state of charge at time t , denoted as $S_{d,t}$. The charging and discharging efficiencies are given by η , respectively, while $P_{d,t}^{ch}$ and $P_{d,t}^{dis}$ indicate the respective power levels. Equations (45), (47), and (48) set the operational ranges for these variables to ensure that the charging and discharging processes can proceed safely without damaging the battery. Equation (46) calculates the net power contributed by energy storage to residences. To prevent simultaneous charging and discharging, constraints are imposed in Equations (49) and (50). Here, s_t^{ch} and s_t^{dis} are binary variables: when $s_t^{ch} = 1$ and $s_t^{dis} = 0$, charging is active, preventing discharging; conversely, when $s_t^{ch} = 0$ and $s_t^{dis} = 1$, discharging is active.

2.9.2. Pricing model

The pricing model remains essentially unchanged; however, in Equation (36), $Gain_t^p$ is a product of two variables, which is not permitted in MILP. Therefore, McCormick relaxation is applied to transform this equation into a linear form.

$$\begin{cases} Gain_t^p \geq 0.8 \cdot \Pi_{t+1}^* \cdot P_t^{e,c} + \Pi_{t+1}^{sell} \cdot P_{min}^{e,c} - 0.8 \cdot \Pi_{t+1}^* \cdot P_{min}^{e,c} \\ Gain_t^p \geq 1.2 \cdot \Pi_{t+1}^* \cdot P_t^{e,c} + \Pi_{t+1}^{sell} \cdot P_{max}^{e,c} - 1.2 \cdot \Pi_{t+1}^* \cdot P_{max}^{e,c} \\ Gain_t^p \leq 0.8 \cdot \Pi_{t+1}^* \cdot P_t^{e,c} + \Pi_{t+1}^{sell} \cdot P_{max}^{e,c} - 0.8 \cdot \Pi_{t+1}^* \cdot P_{max}^{e,c} \\ Gain_t^p \leq 1.2 \cdot \Pi_{t+1}^* \cdot P_t^{e,c} + \Pi_{t+1}^{sell} \cdot P_{min}^{e,c} - 1.2 \cdot \Pi_{t+1}^* \cdot P_{min}^{e,c} \end{cases} \quad (51)$$

Here, since Π_{t+1}^{sell} is obtained by multiplying the electricity price Π_{t+1}^* with the price action A^{pri} , the minimum value of Π_{t+1}^{sell} is 0.8 times the electricity price, and the maximum value is 1.2 times the electricity price. The maximum value of the electricity sold by the ES-PV system, $P_t^{e,c}$, denoted as $P_{max}^{e,c}$, is determined by the combined maximum output of solar generation and the battery system's maximum discharge capacity at each time interval. The minimum value, $P_{min}^{e,c}$, is 0, indicating that the ES-PV system does not sell any power. Through this approach, the product of the two variables in Equation (36) is converted into a set of linear constraints, allowing for MILP-based solution.

2.9.3. Objective function

In MARL, there are two reward functions: a community reward that guides the battery and EV agents, and a sales reward for the ES-PV system that directs the behavior of the pricing agent. Correspondingly, in MILP, there should ideally be two objective functions. However, MILP can only solve for a single objective function. Therefore, the community reward, represented by Equation (43), is selected as the MILP objective function, while the profit reward (Equation 38) is incorporated as a constraint using the ϵ -constraint method.

$$\max \sum_{t \in T} Reward_t^s \quad (52)$$

$$Reward_t^s \geq \epsilon \quad (53)$$

In Equation (52), the aim is to maximize the community reward, with constraints including the modifications made in Section 2.9 and the parts from the MARL algorithm that do not require changes (specifically, Equations (17)–(35), (37)–(39), and (41)–(51)). Equation (53) transforms the original Equation (38) into a constraint, where ϵ is set to the minimum value of $Reward_t^d$ as calculated in the MARL algorithm for convenience.

3. Case study

3.1. Experimental setup

To achieve a more precise energy storage model, experiments were conducted on a pouch cell lithium iron phosphate battery with a nominal capacity of 25Ah to obtain battery parameters [49]. The detailed information of the tested battery is listed in Table 3. The experimental setup includes NEWARE Battery Testing System (CT-8016-5V50A-NTFA, Shenzhen, China), a temperature chamber (NEWARE WGDW Series) [50], and a host computer, as shown in Fig. 6. The battery testing equipment facilitates the charge and discharge processes of the battery. The temperature chamber controls the temperature during the battery experiments. The host computer inputs the experimental procedures, controls both the battery testing equipment and the temperature chamber, and finally records the experimental data.

3.2. Experimental procedure

To obtain battery parameters and complete the HPPC test, the following procedure is performed at each predetermined SOC value. In each HPPC mini-cycle, a 25A discharge is applied for 30 seconds, followed by a 1800-second rest period. Then, a 25A charge is applied for 30 seconds, ensuring no change in SOC within a single mini-cycle. Subsequently, a 2-hour long rest period is initiated, awaiting the next HPPC mini-cycle. Pulse data corresponding to various SOC points are obtained from the HPPC test. The ‘curve fit’ function from the ‘scipy.optimize’ library in Python is then used to identify the battery parameters.

3.3. Training process

The primary objective of this study is to establish an ES-PV system to accommodate the increasing electric vehicle load in the community. To achieve this, we employ a MARL algorithm, focusing on the MAPPO framework, which has shown superior performance, as illustrated in Algorithm 1. The training data for this framework is derived from open-source data provided by Ausgrid, specifically selecting load, solar, and electricity price data from July 1 to July 7, 2011 (winter), and January 1 to January 7, 2012 (summer) as the training dataset. To partially validate the model’s generalizability and transferability, data from July 1 to July 7, 2012 (winter) and January 1 to January 7, 2013 (summer) are used as the validation dataset [51,52]. The validation data (i.e., winter 2012 and summer 2013) are depicted in Fig. 7.

When modeling EVs, representing the EV battery capacity using a distribution model is a highly rational approach. This consideration stems from the variability in battery capacities across different EV models, which directly influences the charging curve. For instance, BMW’s fully electric series can feature battery capacities as high as 78.9

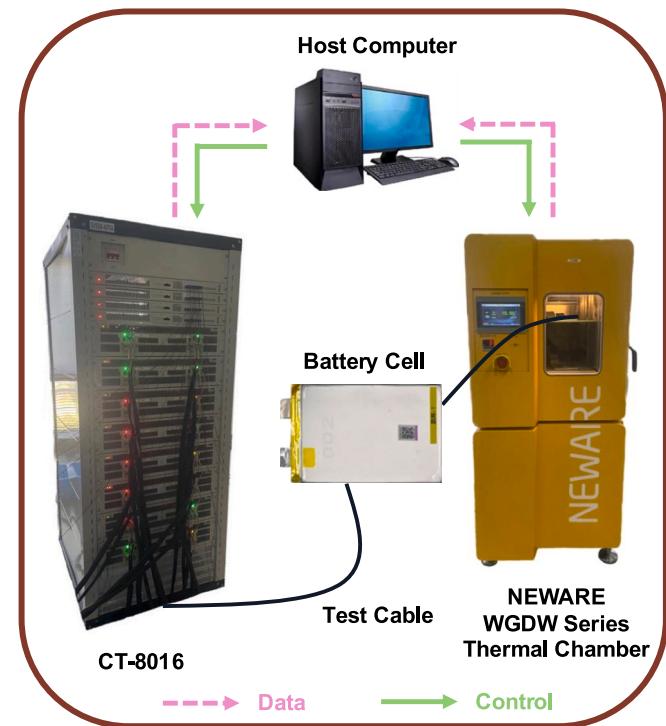


Fig. 6. Experimental setup schematic.

kWh, while their hybrid models only offer capacities of 17.72 kWh [53]. Tesla’s EVs vary between 60 kWh and 80 kWh [54], whereas the capacities of BYD’s Dynasty series range from 12.9 kWh to 26.6 kWh [55]. To more accurately simulate the fleet capacity within a community, the model adopts a uniform distribution for battery capacities within the range of 20–60 kWh. Additionally, the EV penetration rate within the community is set at 20%, a relatively high figure that effectively represents a community with significant EV ownership. Given that the community load data encompasses 300 users, the number of EVs in the fleet is consequently set at 60.

Algorithm 1. Multi-agent proximal policy optimization (MAPPO) algorithm.

Algorithm MAPPO for community energy management

Input: Initial parameters θ_i for critic V and ϕ_i for actor π , Replay buffer D , learning rate α .

for $episode = 1$ to M **do**

 Retrieve initial state space S and observation space O : community load, electricity price, SOC, EV power, sell price.

for $t = 1$ to T **do**

for $agent i = 1$ to N **do**

 Select action A_t^i according to policy $\pi_{\theta_i}(O_t^i; \theta_i)$;

 Calculate value according to $V(S_t^i; \phi_i)$;

end

 Execute action A_t^i in the environment;

 Observe individual reward R_t^i and new state S_{t+1}^i ;

 Store transition $[S_t, O_t, R_t, S_{t+1}, O_{t+1}]$ in buffer D ;

end

 Update θ_i using Adam optimizer on minibatch;

 Update ϕ_i using Adam optimizer on minibatch;

$\theta_{i+1} \leftarrow \theta_i + \alpha \nabla J_{\pi}(\theta)|_{\theta=\theta_i}$

end

Table 3
Detailed information of the tested battery.

Items	Parameters
Type	pouch cell
Material system	$LiFePO_4$
Nominal capacity	25 Ah
Nominal voltage	3.2 V
Cut-off voltage	3.65 V/2.5 V

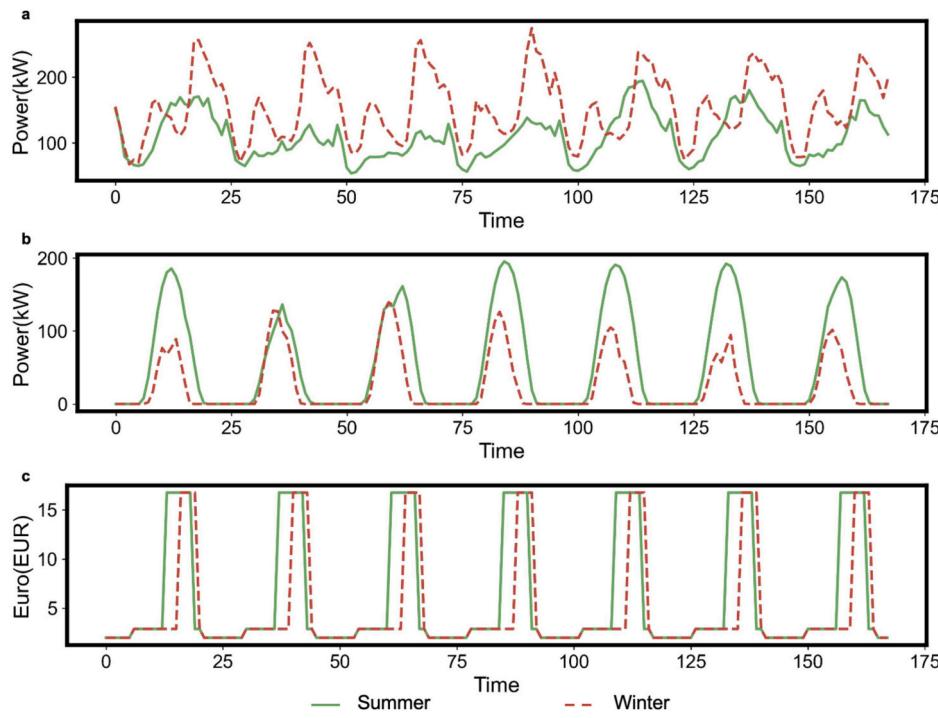


Fig. 7. Real-world data. (a. Power consumption of base loads. b. Power generation of PV system. c. Electricity price.)

The parameters σ_s and μ_s represent the standard deviation and mean, respectively, of the normal distribution governing the start time of EV charging, where $\sigma_s = 3.3$ and $\mu_s = 18$ [46]. In the market, residential charging stations typically offer three power options: 7 kW, 11 kW, and 21 kW [56]. Consequently, the maximum fast charging power $P_{ev,max}$ in the community is set to 20 kW, which aligns with a reasonable power capacity for residential charging stations. The smart charging power, however, is determined by the EV agent's decision, ranging between 0 and 20 kW. Some of these parameters are presented in Tables 2 and 3, while the rest are listed in Table 4.

The training iterates until the algorithm converges, enabling the evaluation of the learned policies' effectiveness in practice. With the exception of the baseline control, all computational models are accelerated using an NVIDIA RTX 4090 (24GB) GPU. All experimental environment code is developed using Pytorch, scikit-learn, and Keras optimizers. The primary hyperparameters for the algorithm design are listed in Table 5. The optimization algorithm is based on the MILP method, which is solved using the CPLEX optimizer. The various data for the community are also derived from the data presented in Fig. 7.

The ultimate goal of the algorithm is to develop optimal charging and discharging strategies for energy storage devices, smart charging strategies for electric vehicles, and pricing strategies for selling

Table 4
All parameters considered in this paper.

Category	Parameters	Category	Parameters
Q	25 Ah	$[s_{min}, s_{max}]$	[0.05, 0.95]
$[I_{min}, I_{max}]$	[0.25A]	$[z_{min}, z_{max}]$	[0.05, 0.95]
η, η_e	0.975	σ_s	3.3
μ_s	18	$[SOC_{min}, SOC_{max}]$	[0.2, 0.8]
h^{max}	8 h	$p_{ev,max}$	20 kW
A^{bat}	$[-1, 1]$	A^{ev}	[0, 1]
A^{pri}	[0.8, 1.2]	$[V_{min}, V_{max}]$	[0.92, 1.08]
C_E	0.05	D	0.95
L_C	4996	$[P_{min}^{ch,c} P_{max}^{ch,c}]$	[0, 450 kW]
$[S_d^{min} S_d^{max}]$	[0, 500 kWh]	$[P_d^{ch,min} P_d^{ch,max}]$	[0, 250 kW]
$Cost_{cell}, Cost_{inv}$	€211.9/kWh, €1271.2	$rate_c$	0.5

Table 5
Algorithm parameters.

Items	Parameters
Actor learning rate	0.001
Critic learning rate	0.005
Batch size	512
Hidden dimension	256
Replay buffer capacity	100,000
Target network update interval	400
Soft update parameter	0.01
Discount factor	0.99

electricity within the ES-PV system. After completing the battery pack modeling, it is necessary to series and parallel the batteries to form the energy storage system model. The charging and discharging power and penalties of the batteries must be input into the MARL ES agent's rewards. The smart charging power of the electric vehicles is determined by the actions of the EV agent. Subsequently, the power of the entire fleet must be transmitted to the MARL and incorporated into the community load, meaning the EV agent and ES agent share the same rewards. Conversely, the price agent focuses more on the interests of the ES-PV system and has its own individual rewards. These aspects are also described in detail in Section 2. Achieving the algorithm's objectives requires complex cooperation and competition among these agents, a process that is depicted in detail in Fig. 8.

4. Result and discussion

4.1. Result of the training process

In this study, three commonly used MARL algorithms were selected for training under the same environment and parameters: MADDPG, MAPPO, and MATD3. As previously mentioned, MAPPO performed the best in this learning task, as evidenced by the training results shown in Fig. 9. All three algorithms tended to converge at around 2000 episodes; however, MADDPG exhibited significant fluctuations post-convergence. This instability is attributed to the algorithm's susceptibility to strategy

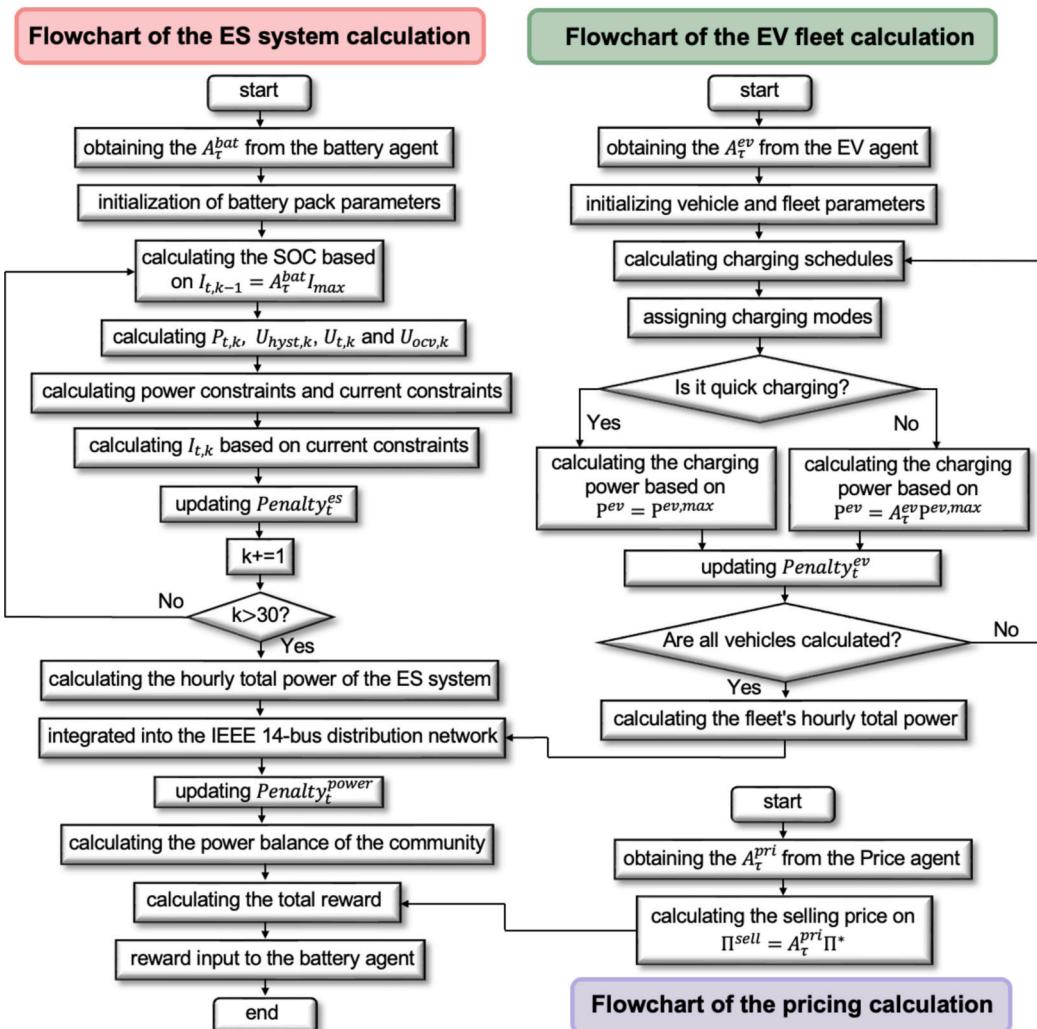


Fig. 8. Flowchart of energy storage (ES) system, electric vehicle (EV) fleet calculation and pricing calculation.

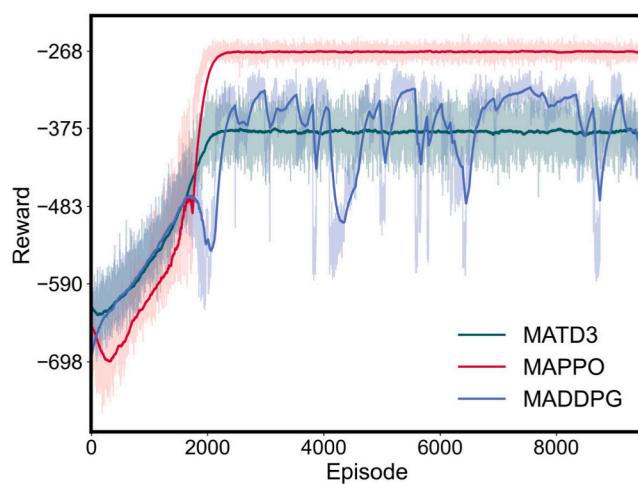


Fig. 9. Training results of multi-agent reinforcement learning algorithm.

and Q-value function instability, resulting in inconsistent performance. MATD3 showed slightly better performance with minor fluctuations post-convergence, but overall, it remained satisfactory. In contrast, MAPPO achieved higher reward values and exhibited minimal

fluctuations post-convergence, making it evidently more suitable for this learning task. Therefore, in the subsequent section showcasing the testing results, we selected the results from the best-performing MAPPO for demonstration.

4.2. Testing results

In this section, we focus on analyzing the contribution of the ES-PV system within the community. We selected the test results of MAPPO as a representative example. First, we demonstrate the effectiveness of the MARL algorithm in managing battery charging and discharging, and evaluate the contribution of the ES-PV system through the photovoltaic self-consumption rate. Next, we present the performance of the EV agent in controlling EV smart charging, and the extent to which the ES-PV system can accommodate the EV load within the community. By shouldering the community load, the ES-PV system can alleviate some of the stress on the distribution network; therefore, we will display the voltage magnitude and power at the distribution network nodes. Finally, we will show the pricing results set by the price agent for the ES-PV system and explain the impact of adopting the ES-PV system on the overall community electricity costs and the profitability of the ES-PV system. The corresponding results for the MILP algorithm will also be presented to enable a comparison between the outcomes of both algorithms.

4.3. Optimal charging and discharging strategies

Fig. 10 illustrates the dynamic performance of battery charging and discharging optimization using the ES agent. The yellow bars represent the SOC of the energy storage battery, with the left vertical axis corresponding to the SOC values. The brown line indicates the community load (including EV load), and the blue line represents the PV generation, with the right vertical axis showing the power values. By observing the changes in SOC, we can see that the energy storage battery tends to charge when there is significant solar power generation. In the evening, the strategic use of the energy storage becomes particularly evident, coinciding with the time when solar generation decreases and the community load increases.

Fig. 11 illustrates battery charging and discharging optimization using the MILP algorithm, where the purple bars represent the SOC, and the red line indicates the community load. The solar generation curve is the same as in **Fig. 10**. It can be observed that the energy storage tends to charge during periods of high solar power generation. Unlike in **Fig. 10**, the MILP algorithm seldom schedules charging outside of solar generation periods, suggesting that this operation may be relatively insensitive to changes in electricity prices.

Table 6 provides a quantitative analysis of the impact of integrating the ES-PV system on PV self-consumption rates. In the MARL, during summer, without the energy storage system, the PV system produces a significant amount of surplus power, resulting in a self-consumption rate of 24.49%. After implementing the ES-PV system, the PV self-consumption rate increases to 90.90%, an improvement of 66.41%. In the winter, due to the significantly reduced PV generation, the unabsorbed energy by the community is also lower, resulting in a relatively higher self-consumption rate of 54.73% even without storage. With the ES-PV system, the self-consumption rate further increases to 99.78%, representing an improvement of 45.05%.

In the MILP algorithm, the initial conditions are the same as those in the MARL algorithm, meaning the PV generation and the PV self-consumption rate without the ES-PV system are identical. However, the self-consumption rate after implementing the system shows differences. In the MILP algorithm, after implementing the ES-PV system, the surplus PV energy is entirely eliminated, resulting in a self-consumption rate of 100%. This indicates that the MILP algorithm is more sensitive to penalizing wasted solar energy, demonstrating better performance in improving community energy efficiency.

Overall, these analyses highlight the effectiveness of the ES-PV system in managing instantaneous energy consumption and strategically utilizing storage under both algorithm frameworks, significantly enhancing the energy efficiency of the entire community.

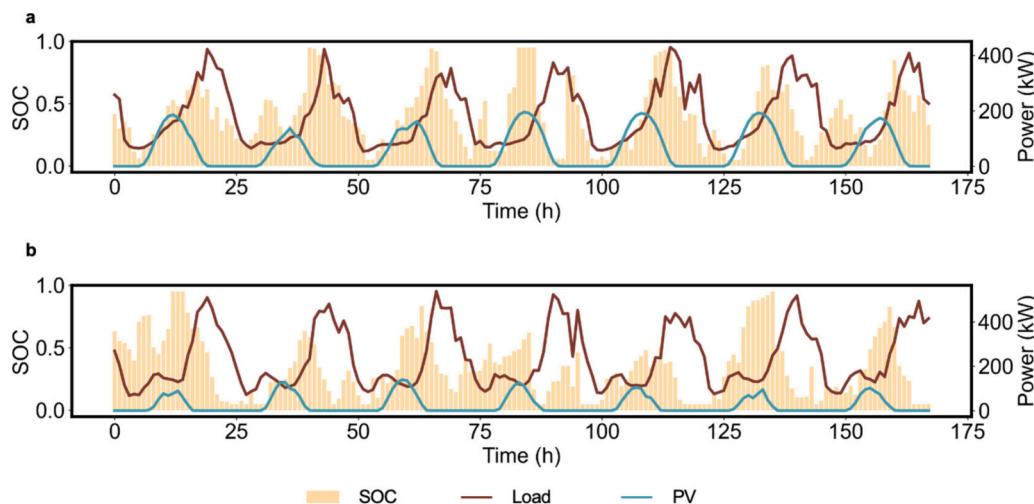


Fig. 10. The optimized charging and discharging profile of the energy storage battery system within MARL algorithm. (a. Summer case; b. Winter case.)

4.4. Smart charging strategies for electric vehicles

Fig. 12 provides a detailed depiction of the EV fleet charging profiles before and after the implementation of the EV agent control within the MARL algorithm. For clarity, data for only one day is shown. In **Fig. 12a**, the blue curve represents the smart charging profile of the fleet under the control of the EV agent, while the purple curve indicates the charging profile without smart charging. The yellow areas highlight the portions where the non-smart charging curve exceeds the smart charging curve, whereas the green areas indicate the opposite. **Fig. 12b** presents the base load (excluding EV demand) and the solar power generation curve, facilitating the evaluation of the effectiveness of smart EV charging.

By comparing the two figures, it is evident that without smart charging, the peak charging fleet load (17:00-20:00) nearly coincides with the peak community baseline load. This implies that the overall community load is significantly high during this period, while solar power generation is zero, resulting in substantial stress on the distribution grid. After implementing smart charging, the fleet's charging load shifts significantly later, preventing large load spikes at certain times and alleviating the pressure on the community distribution grid. Additionally, a considerable portion of the load (yellow area) is shifted out of peak electricity pricing periods, ensuring economic efficiency for the community. This illustrates an advanced EV charging power management.

Within the MARL framework, the fleet charging can flexibly avoid peaks in other community loads, demonstrating the system's capability to adaptively manage loads, thereby enhancing the stability of the community distribution grid and ensuring cost-effectiveness for the community.

Within the MARL framework, not only does the EV agent enable intelligent charging and load shifting for the fleet, but the ES-PV system also plays a crucial role in sharing the fleet load, thereby alleviating the burden on the social distribution grid. **Fig. 12** illustrates the fleet power's dependence on the grid before and after the integration of the ES-PV system. The purple curve represents the load conditions prior to integration, while the blue curve depicts the load conditions post-integration. It is evident from the figure that, regardless of whether it is summer or winter, the integration of the ES-PV system significantly reduces the fleet's reliance on the community distribution grid, thereby substantially mitigating grid stress. This integrated strategy not only optimizes energy utilization but also enhances the system's flexibility and reliability in responding to variable load demands.

In addition to the aforementioned smart charging, the ES-PV system also plays a crucial role in sharing the fleet load, thereby alleviating the

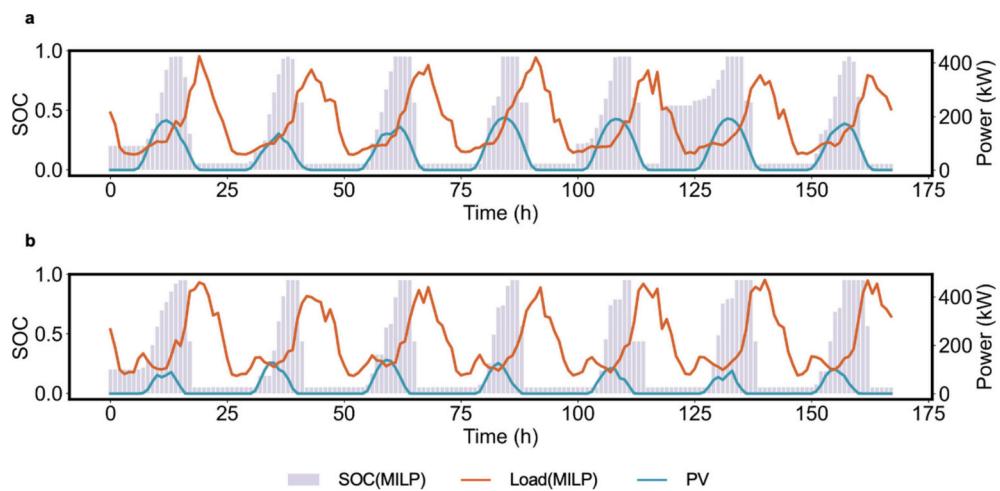


Fig. 11. The optimized charging and discharging profile of the energy storage battery system within MILP algorithm. (a. Summer case; b. Winter case.)

Table 6
Self-consumption ratio comparison before and after ES-PV system installation.

Season	Algorithm	Parameter	Before ES-PV system implementation	After ES-PV system implementation
Summer	MARL	PV generation (kWh)	9272.42	9272.42
		Surplus energy (kWh)	7001.55	843.42
		Self-consumption ratio	24.49%	90.90%
Winter	MILP	Surplus energy (kWh)	7001.55	0
		Self-consumption ratio	24.49%	100%
		PV generation (kWh)	4135.52	4135.52
Summer	MARL	Surplus energy (kWh)	1980.01	9.29
		Self-consumption ratio	54.73%	99.78%
		Surplus energy (kWh)	1980.01	0
Winter	MILP	Self-consumption ratio	54.73%	100%

burden on the distribution grid. Fig. 13 illustrates the dependence of the fleet on the grid before and after the integration of the ES-PV system within the MARL algorithm. The solid line represents the EV load conditions before integration, while the dashed line represents the load conditions post-integration. As shown in the figure, regardless of whether it is summer or winter, the dashed line is significantly lower than the solid line after the integration of the ES-PV system, indicating that the ES-PV system significantly reduces the fleet's reliance on the community distribution grid, thereby substantially mitigating grid stress. This integrated strategy not only optimizes energy utilization but also enhances the system's flexibility and reliability in responding to variable load demands. Fig. 14 shows the results of the MILP algorithm for addressing the same issue, also demonstrating that the ES-PV system can undertake a substantial portion of the EV load.

Table 7 presents a quantitative analysis corresponding to Figs. 13 and 14, providing a clearer perspective on the ES-PV system's contribution to fleet load management. In the MARL framework, during

summer, the fleet's total weekly demand amounts to 12,452.87 kWh. Prior to the implementation of the ES-PV system, this demand was entirely met by the grid. Following the integration of the ES-PV system, 4816.15 kWh of this demand was supplied by the system, thereby accounting for 38.68% of the fleet's load. In winter, due to the reduction in PV generation, the load-sharing capacity of the ES-PV system decreased, yet it still managed to support 23.76% of the load. Therefore, irrespective of the season, the deployment of the ES-PV system significantly alleviates the burden on the community distribution grid imposed by the increasing electric vehicle load.

In the MILP algorithm results, it can be observed that the ES-PV system also undertakes a substantial portion of the EV load, covering 41.15% in summer and 31.73% in winter, which are both higher than the results of the MARL algorithm. This indicates that, irrespective of the solution method, the ES-PV system demonstrates an excellent capacity to support the EV load.

4.5. Analysis of voltage magnitude and power at distribution network nodes

Fig. 15a illustrates the voltage magnitude within the MARL framework. For the IEEE-14 node system, specifically at node 12 in our installed community, voltage limit penalties were set, with penalties applied for values exceeding 1.08 or falling below 0.92. As depicted in Fig. 15a, node 12 remains within the voltage limits. It is important to note that the voltage magnitude at node 12 shows minimal variation over different times, thus only one hour is depicted for clarity. Fig. 15b contrasts the power at node 12 before and after the installation of the ES-PV system. The blue curve represents the scenario prior to the ES-PV installation, while the pink curve represents the scenario post-installation. Generally, before the ES-PV system installation, community loads (including EV loads) are fully dependent on the distribution grid. However, post-installation, the dependency on the distribution grid is significantly reduced. This aligns with the previously discussed reduction in EV load dependency on the grid following the ES-PV system installation, and thus requires no further elaboration. It is noteworthy that the reduced dependency on the distribution grid not only alleviates grid stress but also reduces the overall electricity costs for the community, an economic advantage that will be elaborated upon subsequently.

4.6. Pricing strategies for ES-PV system

Fig. 16 illustrates the pricing outcomes determined by the Price Agent in the MARL algorithm. The pink curve represents the market electricity price, while the green curve denotes the selling price set for

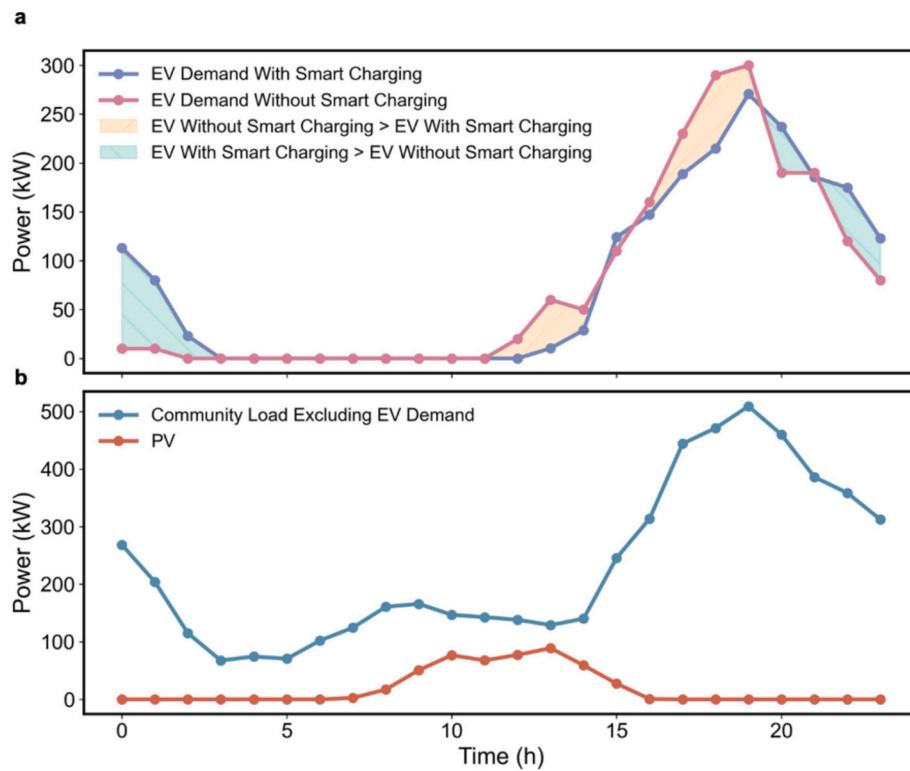


Fig. 12. a. Comparison between smart EV charging mode and conventional charging mode within MARL algorithm; b. Community Load excluding EV demand and PV generation.

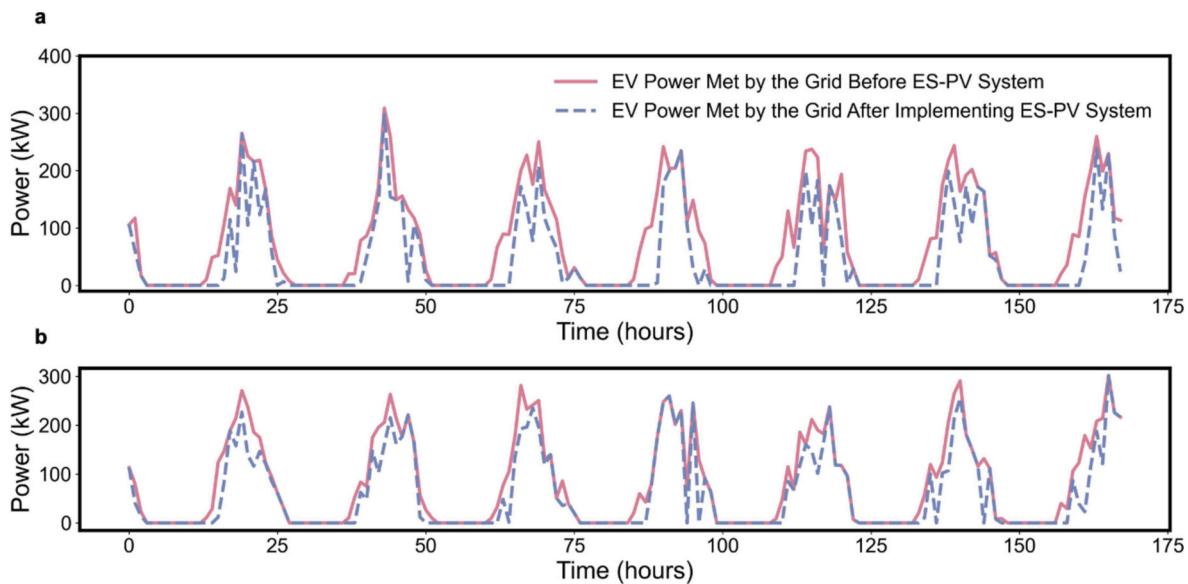


Fig. 13. EV power dependence on the grid before and after implementing the ES-PV system within MARL algorithm. (a. Summer case; b. Winter case.)

the ES-PV system. It is evident that the selling price of the ES-PV system is generally lower than the market electricity price. This selling price is established within the Price Agent's reward mechanism, based on the economic interests of the ES-PV system. In essence, the objective is to increase the ES-PV system's revenue through electricity sales while minimizing the cost of purchasing electricity from the grid. The achievement of this objective will be analyzed in detail later. Nonetheless, it can be observed from the figure that the selling price of the ES-PV system is significantly more competitive compared to the market electricity price, thereby highlighting the economic attractiveness of this

system.

Fig. 17 presents the pricing outcomes in the MILP algorithm. The pink curve still represents the market electricity price, while the blue dashed line indicates the selling price variable for the ES-PV system as guided by the objective in the MILP framework. Unlike the MARL algorithm, the price in the MILP algorithm consistently takes the minimum value within the range of electricity prices.

Table 8 provides a further analysis of the community's overall economic status under both the MARL and MILP frameworks. In both algorithms, the energy storage system is identical; thus, the capital cost of

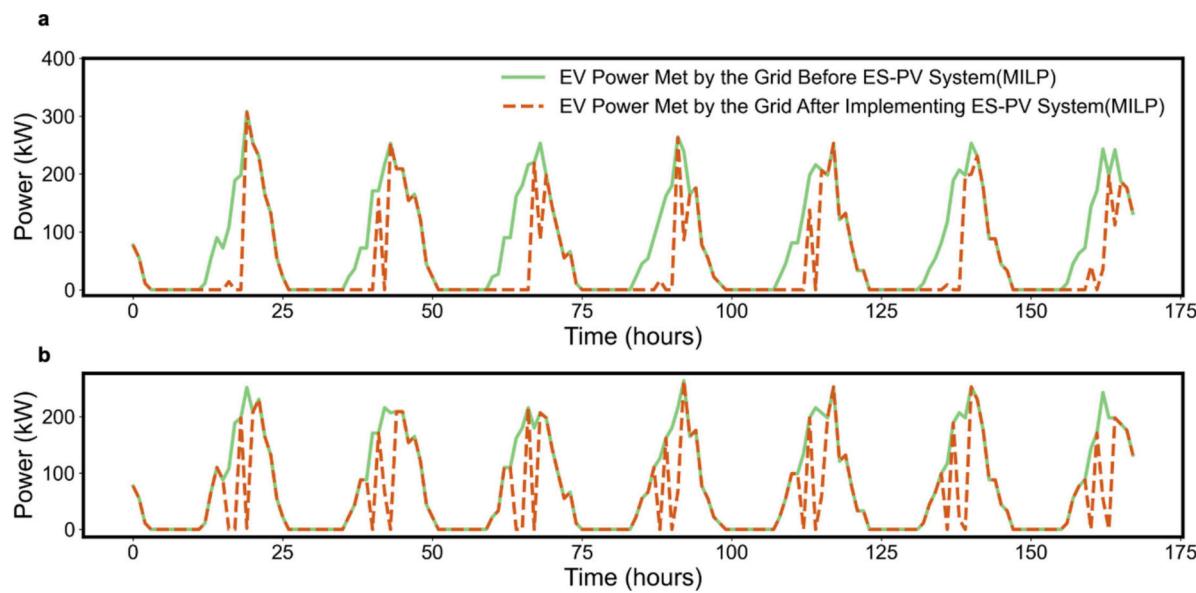


Fig. 14. EV power dependence on the grid before and after implementing the ES-PV system within MILP algorithm. (a. Summer case; b. Winter case.)

Table 7
Extent of electric vehicle load supported by ES-PV system.

Season	Algorithm	Parameter	Before ES-PV system implementation	After ES-PV system implementation	
Summer	MARL	E^{g_ev} (kWh)	12,452.87	7636.72	
		E^e_{-ev} (kWh)	0	4816.15	
		$Ratio_{ev}$	0%	38.68 %	
Winter	MILP	E^{g_ev} (kWh)	12,810.10	7537.85	
		E^e_{-ev} (kWh)	0	5272.15	
		$Ratio_{ev}$	0%	41.15%	
Summer	MARL	E^{g_ev} (kWh)	13,251.73	10,103.30	
		E^e_{-ev} (kWh)	0	3148.43	
		$Ratio_{ev}$	0%	23.76%	
Winter	MILP	E^{g_ev} (kWh)	13,512.26	9224.95	
		E^e_{-ev} (kWh)	0	4287.31	
		$Ratio_{ev}$	0%	31.73%	
E^{g_ev} Electric vehicle load met by the grid					
E^e_{-ev} Electric vehicle load met by the ES-PV system					
$Ratio_{ev}$ Extent of EV load supported by ES-PV system					

the storage system can be estimated as €134,055.22 using Equation (54) [57]. According to data from the National Renewable Energy Laboratory (NREL), in 2011, the installation cost of solar energy included a photovoltaic module cost of €2.78/W and an inverter cost of €0.89/W. With a community PV capacity of approximately 200 kWh, the total capital cost is estimated to be around €733,954.88 [58].

In the MARL framework, before integrating the ES-PV system, the community's electricity was entirely dependent on the grid, purchased at the market electricity price (depicted by the pink curve in Fig. 16). The total expenditure during the summer was €217,236.69. After the integration of the ES-PV system, the community could purchase electricity from both the grid and the ES-PV system at its selling price (depicted by the green curve in Fig. 16). This resulted in a reduced total expenditure of €200,437.70, representing a 7.73% reduction. The ES-PV system incurred costs of €28,477.46 for purchasing electricity from the

grid and generated €80,402.11 by selling electricity to the community, yielding a net profit of €51,924.65. Using this net profit, the return on investment (ROI) can be calculated as 5.98% using Equation (55). In the MILP framework, the ES-PV system achieved a high net profit of €88,851.14, resulting in a high ROI. Additionally, community residents also benefited significantly, with community electricity costs reduced by 13.48%.

$$Cost_{capital}^{ev} = Cost_{cell} \times S^{max} + Cost_{inv} \times \left(\frac{rate_c \times S^{max}}{3} \right)^{0.7} \quad (54)$$

$$ROI_{weekly} = \frac{Profit_{Net}}{Cost_{capital}} \quad (55)$$

During winter, the MARL framework achieved a reduction in community electricity costs by 2.50%, although this reduction was smaller compared to summer, indicating that the community still benefited overall. The ES-PV system achieved a net profit of €23,601.87 during winter. Thus, the ES-PV system not only resulted in significant cost savings for the community but also generated considerable net profit, making it a win-win solution. In the MILP framework, the community also benefited during winter, with community electricity costs reduced by 5.91%. The ES-PV system continued to generate high profits, amounting to €34,907.71.

Overall, the MILP algorithm simplifies many aspects during implementation, which is why the results from MILP tend to outperform those from MARL. While MILP excels in high-level planning, it may face challenges in adapting to complex environmental changes as a real-time control algorithm in realistic settings. Additionally, it lacks the capability to effectively coordinate collaboration and competition among different entities. Nevertheless, using MILP as a benchmark shows that the MARL results are reasonable and accurate. These analyses also highlight the advantages of MARL in complex environments, where it can achieve a sophisticated balance in managing community loads and strategically utilizing storage and solar energy in more realistic battery system models. The system not only ensures operational efficiency but also optimizes economic benefits. This balance underscores the system's exceptional capability to optimize energy resources, reduce dependency on the distribution grid, and maintain high energy efficiency and economic viability.

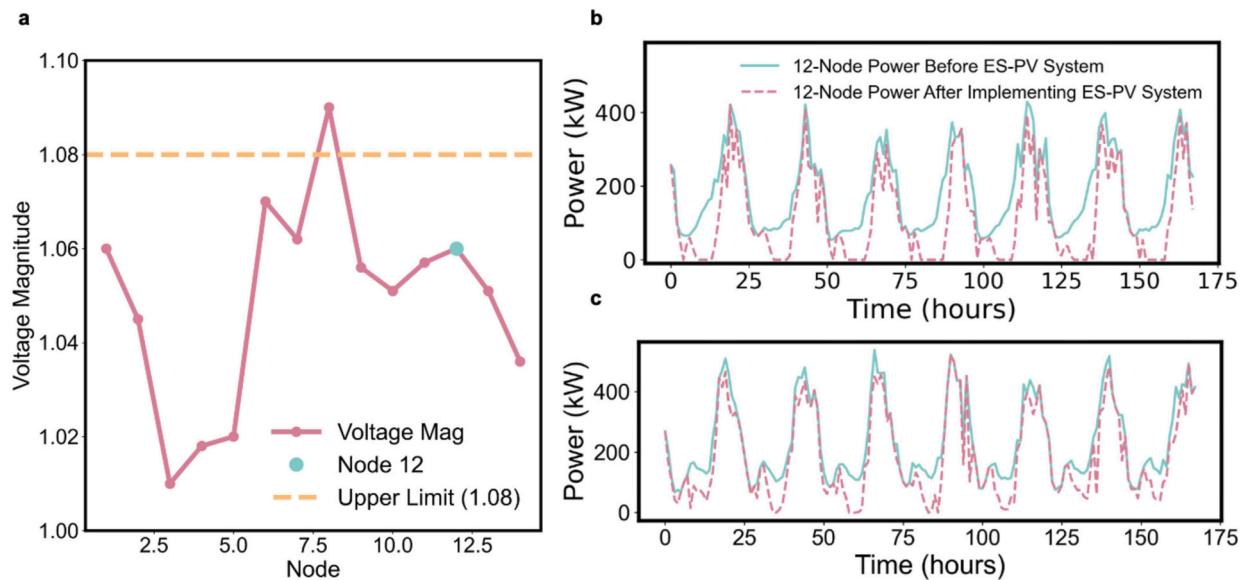


Fig. 15. Voltage magnitude and 12-Node power comparison before and after implementing the ES-PV system in IEEE-14 Node Network within MARL algorithm. (a. Voltage magnitude of each node; b. 12-Node power comparison in summer; c. 12-Node power comparison in winter.)

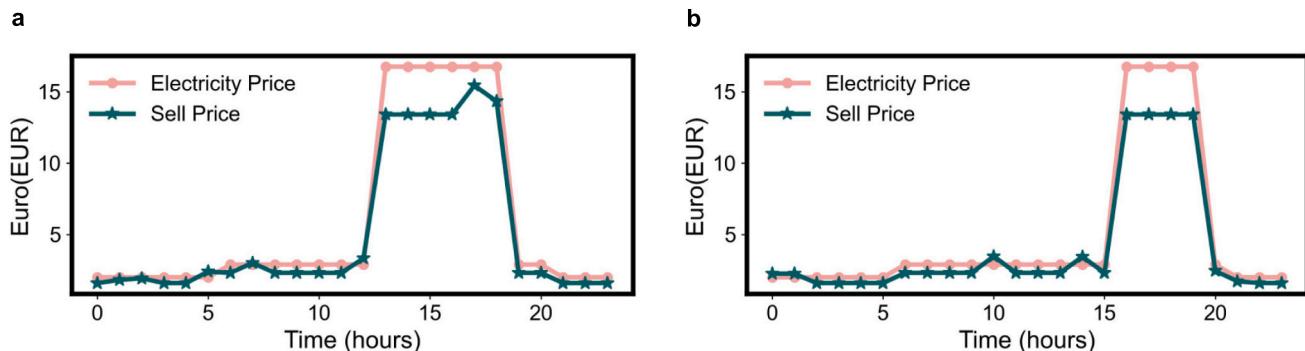


Fig. 16. ES-PV system electricity selling price within MARL algorithm. (a. Summer case; b. Winter case.)

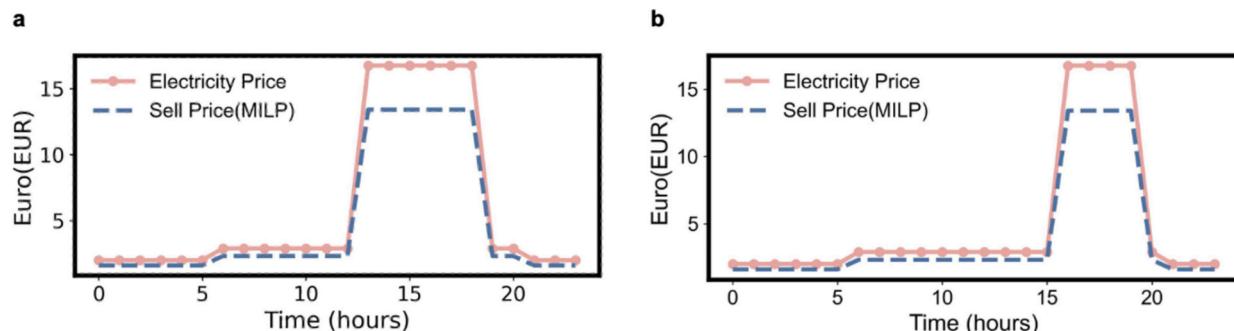


Fig. 17. ES-PV system electricity selling price within MILP algorithm. (a. Summer case; b. Winter case.)

5. Conclusion

This paper proposes a community-shared ES-PV system to meet load demands, employing a MARL algorithm for battery charging/discharging operations, intelligent EV charging, and pricing for the ES-PV system. The framework was trained and validated using real-world data from Australia, demonstrating its effectiveness. The key conclusions drawn from this study can be summarized as follows:

- (1) The MARL algorithm accurately identifies trends in community load and solar generation curves, providing more effective guidance for the charging and discharging operations of storage batteries, thereby improving energy efficiency.
- (2) After implementing the community-shared ES-PV system, the self-consumption rate of the PV system increased by 66.41% in summer and 45.05% in winter, significantly enhancing the system's energy utilization efficiency for the community.

Table 8

Comparative analysis of weekly community costs before and after ES-PV system installation.

Season	Algorithm	Parameters	Before ES-PV system implementation	After ES-PV system implementation
Summer	MARL	Capital costs (€)	0	868,010.10
		Community costs (€)	217,236.69	200,437.7
		ES-PV costs (€)	0	28,477.46
	MILP	ES-PV benefits (€)	0	80,402.11
		ROI _{weekly}	0	5.98%
		Community costs (€)	231,270.54	200,100.33
Winter	MARL	ES-PV costs (€)	0	32,029.57
		ES-PV benefits (€)	0	120,880.71
		ROI _{weekly}	0	10.24 %
	MILP	Community costs (€)	256,209.61	249,788.04
		ES-PV costs (€)	0	11,687.77
		ES-PV benefits (€)	0	35,289.64
	MILP	ROI _{weekly}	0	2.72 %
		Community costs (€)	258,038.88	242,799.20
		ES-PV costs (€)	0	26,015.01
	MILP	ES-PV benefits (€)	0	60,958.72
		ROI _{weekly}	0	4.02 %

- (3) The MARL algorithm effectively guides intelligent EV charging, successfully shifting peak demand, reducing the pressure on the distribution grid, and enhancing grid stability.
- (4) After installing the ES-PV system, with an EV penetration rate of 20% in the community, it was able to support 38.68% of the community's EV load in summer and 23.76% in winter, demonstrating the system's potential in balancing community power demands.
- (5) The MARL algorithm can set pricing for the ES-PV system, resulting in net profits of €51,924.65 in summer, with an ROI of 5.98%, and €23,601.87 in winter, with an ROI of 2.72%. This finding underscores the potential for optimizing the economic performance of ES-PV systems through intelligent pricing strategies.
- (6) After installing the ES-PV system, community residents benefited from a 7.73% reduction in electricity costs in summer and a 2.50% reduction in winter. This result indicates that the ES-PV system not only benefits the grid and operators but also significantly enhances the economic welfare of community residents.
- (7) An MILP algorithm was introduced as a benchmark, enhancing the validity and credibility of the MARL results.

Overall, the MARL algorithm demonstrated strong performance in effectively guiding the coordinated decision-making of various agents in this study, achieving a win-win outcome for both community residents and ES-PV system operators while ultimately reducing the pressure on the distribution grid. Future research should further explore the potential application of this algorithm in larger-scale and more complex environments, as well as its performance under different climatic conditions and load types. Additionally, optimizing battery lifespan management and grid integration technologies are crucial research directions to further enhance the system's sustainability and economic efficiency.

CRediT authorship contribution statement

Baligen Talihiati: Writing – original draft, Validation, Methodology, Formal analysis, Data curation, Conceptualization. **Shiyi Fu:** Visualization, Validation, Investigation. **Bowen Zhang:** Visualization, Investigation. **Yuqing Zhao:** Validation. **Yu Wang:** Supervision, Funding acquisition. **Yaojie Sun:** Supervision, Methodology, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

This study was supported by National Science and Technology Major Project of Smart Grid (Grant No. 2024ZD0800400), State Grid Shanghai Electric Power Company Project (Grant No. 520940240014), and Shanghai Engineering Research Center for Artificial Intelligence and Integrated Energy System (Grant No. 19DZ2252000).

References

- [1] Perera P, Hewage K, Sadiq R. Electric vehicle recharging infrastructure planning and management in urban communities. *J Clean Prod* Mar. 2020;250:119559. <https://doi.org/10.1016/j.jclepro.2019.119559>.
- [2] Talihiati B, et al. Energy storage sharing in residential communities with controllable loads for enhanced operational efficiency and profitability. *Appl Energy* Nov. 2024;373:123880. <https://doi.org/10.1016/j.apenergy.2024.123880>.
- [3] Tao S, et al. Rapid and sustainable battery health diagnosis for recycling pretreatment using fast pulse test and random forest machine learning. *J Power Sources* Mar. 2024;597:234156. <https://doi.org/10.1016/j.jpowsour.2024.234156>.
- [4] Tao S, et al. Collaborative and privacy-preserving retired battery sorting for profitable direct recycling via federated machine learning. *Nat Commun* 2023;14 (1):8032. <https://doi.org/10.1038/s41467-023-43883-y>.
- [5] Fu S, et al. Data-driven capacity estimation for lithium-ion batteries with feature matching based transfer learning method. *Appl Energy* 2024;353:121991. <https://doi.org/10.1016/j.apenergy.2023.121991>.
- [6] International Energy Agency [Online]. Available: <https://www.iea.org/reports/global-ev-outlook-2024>. [Accessed 5 August 2024].
- [7] National Development and Reform Commission [Online]. Available: https://www.ndrc.gov.cn/fggz/fgyz/shgqhy/202407/t20240730_1392076.html. [Accessed 5 August 2024].
- [8] Li Y, Han M, Yang Z, Li G. Coordinating flexible demand response and renewable uncertainties for scheduling of community integrated energy systems with an electric vehicle Charging Station: a bi-level approach. *IEEE Trans Sustain Energy* Oct. 2021;12(4):2321–31. <https://doi.org/10.1109/TSTE.2021.3090463>.
- [9] Tan KM, Ramachandaramurthy VK, Yong JY. Optimal vehicle to grid planning and scheduling using double layer multi-objective algorithm. *Energy* Oct. 2016;112: 1060–73. <https://doi.org/10.1016/j.energy.2016.07.008>.
- [10] Li S, Gu C, Zeng X, Zhao P, Pei X, Cheng S. Vehicle-to-grid management for multi-time scale grid power balancing. *Energy* Nov. 2021;234:121201. <https://doi.org/10.1016/j.energy.2021.121201>.
- [11] Tan KM, Padmanabhan S, Yong JY, Ramachandaramurthy VK. A multi-control vehicle-to-grid charger with bi-directional active and reactive power capabilities for power grid support. *Energy* Mar. 2019;171:1150–63. <https://doi.org/10.1016/j.energy.2019.01.053>.
- [12] Tarroja B, Zhang L, Wifvat V, Shaffer B, Samuels S. Assessing the stationary energy storage equivalency of vehicle-to-grid charging battery electric vehicles. *Energy* Jul. 2016;106:673–90. <https://doi.org/10.1016/j.energy.2016.03.094>.
- [13] Staudt P, Schmidt M, Gärtner J, Weinhardt C. A decentralized approach towards resolving transmission grid congestion in Germany using vehicle-to-grid technology. *Appl Energy* Nov. 2018;230:1435–46. <https://doi.org/10.1016/j.apenergy.2018.09.045>.
- [14] Buonomano A. Building to vehicle to building concept: a comprehensive parametric and sensitivity analysis for decision making aims. *Appl Energy* Mar. 2020;261:114077. <https://doi.org/10.1016/j.apenergy.2019.114077>.
- [15] Borge-Diez D, Icaza D, Açıkkalp E, Amaral H. Combined vehicle to building (V2B) and vehicle to home (V2H) strategy to increase electric vehicle market share. *Energy* Dec. 2021;237:121608. <https://doi.org/10.1016/j.energy.2021.121608>.

- [16] Kuang Y, Chen Y, Hu M, Yang D. Influence analysis of driver behavior and building category on economic performance of electric vehicle to grid and building integration. *Appl Energy* Dec. 2017;207:427–37. <https://doi.org/10.1016/j.apenergy.2017.07.006>.
- [17] Tostado-Véliz M, Jordehi AR, Mansouri SA, Jurado F. A two-stage IGDT-stochastic model for optimal scheduling of energy communities with intelligent parking lots. *Energy* Jan. 2023;263:126018. <https://doi.org/10.1016/j.energy.2022.126018>.
- [18] Babaei M, Abazari A, Soleymani MM, Ghafouri M, Muyeen SM, Beheshti MTH. A data-mining based optimal demand response program for smart home with energy storages and electric vehicles. *J Energy Storage* Apr. 2021;36:102407. <https://doi.org/10.1016/j.est.2021.102407>.
- [19] Mokryani G. Future distribution networks: planning, operation, and control. AIP Publishing; 2022. <https://doi.org/10.1063/9780735422339>.
- [20] Schwenk K, Meisenbacher S, Briegel B, Harr T, Hagenmeyer V, Mikut R. Integrating battery aging in the optimization for bidirectional charging of electric vehicles. *IEEE Trans Smart Grid* Nov. 2021;12(6):5135–45. <https://doi.org/10.1109/TSG.2021.3099206>.
- [21] Crozier C, Morstyn T, McCulloch M. The opportunity for smart charging to mitigate the impact of electric vehicles on transmission and distribution systems. *Appl Energy* Jun. 2020;268:114973. <https://doi.org/10.1016/j.apenergy.2020.114973>.
- [22] Lauvergne R, Perez Y, Françon M, Tejeda De La Cruz A. Integration of electric vehicles into transmission grids: a case study on generation adequacy in Europe in 2040. *Appl Energy* Nov. 2022;326:120030. <https://doi.org/10.1016/j.apenergy.2022.120030>.
- [23] Zhong Z, Hu W, Zhao X. Rethinking electric vehicle smart charging and greenhouse gas emissions: renewable energy growth, fuel switching, and efficiency improvement. *Appl Energy* May 2024;361:122904. <https://doi.org/10.1016/j.apenergy.2024.122904>.
- [24] Tarroja B, Hittinger E. The value of consumer acceptance of controlled electric vehicle charging in a decarbonizing grid: the case of California. *Energy* Aug. 2021; 229:120691. <https://doi.org/10.1016/j.energy.2021.120691>.
- [25] Geng L, Lu Z, He L, Zhang J, Li X, Guo X. Smart charging management system for electric vehicles in coupled transportation and power distribution systems. *Energy* Dec. 2019;189:116275. <https://doi.org/10.1016/j.energy.2019.116275>.
- [26] Truong VB, Le LB. Electric vehicle charging design: the factored action based reinforcement learning approach. *Appl Energy* Apr. 2024;359:122737. <https://doi.org/10.1016/j.apenergy.2024.122737>.
- [27] Lyu C, Jia Y, Xu Z. Fully decentralized peer-to-peer energy sharing framework for smart buildings with local battery system and aggregated electric vehicles. *Appl Energy* Oct. 2021;299:117243. <https://doi.org/10.1016/j.apenergy.2021.117243>.
- [28] Huang P, Lovati M, Zhang X, Bales C. A coordinated control to improve performance for a building cluster with energy storage, electric vehicles, and energy sharing considered. *Appl Energy* Jun. 2020;268:114983. <https://doi.org/10.1016/j.apenergy.2020.114983>.
- [29] Van Der Kam M, Van Sark W. Smart charging of electric vehicles with photovoltaic power and vehicle-to-grid technology in a microgrid; a case study. *Appl Energy* Aug. 2015;152:20–30. <https://doi.org/10.1016/j.apenergy.2015.04.092>.
- [30] Zhu Z, et al. Energy saving and carbon reduction schemes for families with the household PV-BES-EV system. *Energ Build* Jun. 2023;288:113007. <https://doi.org/10.1016/j.enbuild.2023.113007>.
- [31] Chen Q, Kuang Z, Liu X, Zhang T. Application-oriented assessment of grid-connected PV-battery system with deep reinforcement learning in buildings considering electricity price dynamics. *Appl Energy* Jun. 2024;364:123163. <https://doi.org/10.1016/j.apenergy.2024.123163>.
- [32] Wang Z, Xiao F, Ran Y, Li Y, Xu Y. Scalable energy management approach of residential hybrid energy system using multi-agent deep reinforcement learning. *Appl Energy* Aug. 2024;367:123414. <https://doi.org/10.1016/j.apenergy.2024.123414>.
- [33] Ye Y, Papadaskalopoulos D, Yuan Q, Tang Y, Strbac G. Multi-agent deep reinforcement learning for coordinated energy trading and flexibility services provision in local electricity markets. *IEEE Trans Smart Grid* Mar. 2023;14(2): 1541–54. <https://doi.org/10.1109/TSG.2022.3149266>.
- [34] Wu H, Qiu D, Zhang L, Sun M. Adaptive multi-agent reinforcement learning for flexible resource management in a virtual power plant with dynamic participating multi-energy buildings. *Appl Energy* Nov. 2024;374:123998. <https://doi.org/10.1016/j.apenergy.2024.123998>.
- [35] Li Y, Hou J, Yan G. Exploration-enhanced multi-agent reinforcement learning for distributed PV-ESS scheduling with incomplete data. *Appl Energy* Apr. 2024;359: 122744. <https://doi.org/10.1016/j.apenergy.2024.122744>.
- [36] Si R, Chen S, Zhang J, Xu J, Zhang L. A multi-agent reinforcement learning method for distribution system restoration considering dynamic network reconfiguration. *Appl Energy* Oct. 2024;372:123625. <https://doi.org/10.1016/j.apenergy.2024.123625>.
- [37] Yu L, et al. Multi-agent deep reinforcement learning for HVAC control in commercial buildings. *IEEE Trans Smart Grid* Jan. 2021;12(1):407–19. <https://doi.org/10.1109/TSG.2020.3011739>.
- [38] Jendoubi I, Bouffard F. Multi-agent hierarchical reinforcement learning for energy management. *Appl Energy* Feb. 2023;332:120500. <https://doi.org/10.1016/j.apenergy.2022.120500>.
- [39] Xu X, Jia Y, Xu Y, Xu Z, Chai S, Lai CS. A multi-agent reinforcement learning-based data-driven method for home energy management. *IEEE Trans Smart Grid* Jul. 2020;11(4):3201–11. <https://doi.org/10.1109/TSG.2020.2971427>.
- [40] Veerasamy V, et al. Recurrent network based power flow solution for voltage stability assessment and improvement with distributed energy sources. *Appl Energy* Nov. 2021;302:117524. <https://doi.org/10.1016/j.apenergy.2021.117524>.
- [41] Zhang X, Liu Y, Duan J, Qiu G, Liu T, Liu J. DDPG-based multi-agent framework for SVC tuning in urban Power grid with renewable energy resources. *IEEE Trans Power Syst* Nov. 2021;36(6):5465–75. <https://doi.org/10.1109/TPWRS.2021.3081159>.
- [42] Wang J, Hou Q, Zhuo Z, Jia H, Zhang N. Voltage stability constrained economic dispatch for multi-infeed HVDC Power systems. *IEEE Trans Power Syst* Mar. 2024; 39(2):2598–610. <https://doi.org/10.1109/TPWRS.2023.3277213>.
- [43] Plett GL. *Battery management systems. Volume I, Battery modeling*. Norwood: Artech House; 2015.
- [44] Plett GL. Extended Kalman filtering for battery management systems of LiPB-based HEV battery packs. *J Power Sources* Aug. 2004;134(2):277–92. <https://doi.org/10.1016/j.jpowsour.2004.02.033>.
- [45] He H, Zhang Y, Xiong R, Wang C. A novel Gaussian model based battery state estimation approach: state-of-energy. *Appl Energy* Aug. 2015;151:41–8. <https://doi.org/10.1016/j.apenergy.2015.04.062>.
- [46] Zhou K, Cheng L, Wen L, Lu X, Ding T. A coordinated charging scheduling method for electric vehicles considering different charging demands. *Energy* Dec. 2020; 213:118882. <https://doi.org/10.1016/j.energy.2020.118882>.
- [47] Zimmerman Ray Daniel, Murillo-Sánchez Carlos Edmundo, Thomas Robert John. MATPOWER: Steady-State Operations, Planning, and Analysis Tools for Power Systems Research and Education. *IEEE Transactions on Power Systems* 2011;26(1): 12–9.
- [48] Mulleriyawage UKG, Shen W. Impact of demand side management on peer-to-peer energy trading in a DC microgrid. In: 2021 31st Australasian universities Power engineering conference (AUPEC). Perth, Australia: IEEE; Sep. 2021. p. 1–6. <https://doi.org/10.1109/AUPEC52110.2021.9597792>.
- [49] Great Power [Online]. Available: <https://www.greatpower.net>. [Accessed 3 July 2024].
- [50] Neware [Online]. Available: <https://neware.com.cn/service>. [Accessed 3 July 2024].
- [51] Ausgrid [Online]. Available: <https://www.ausgrid.com.au/Industry/Our-Research/Data-to-share/Solar-home-electricity-data>. [Accessed 16 May 2024].
- [52] Ausgrid [Online]. Available: <https://www.ausgrid.com.au/Your-energyuse/Meters/Time-of-use-pricing>. [Accessed 15 May 2024].
- [53] BMW [Online]. Available: <https://www.bmw.com.cn/zh/topics/experience/electromobility/HVB.html>. [Accessed 10 August 2024].
- [54] Tesla [Online]. Available: <https://www.tesla.cn/model3-choose>. [Accessed 10 August 2024].
- [55] BYD [Online]. Available: <https://www.bydauto.com.cn/pc/carDetail/config?id=124&networkType=dynasty>. [Accessed 10 August 2024].
- [56] Youjia [Online]. Available: <https://www.youjia.com/article/899703474626795993.html>. [Accessed 10 August 2024].
- [57] Chang H-C, Ghaddar B, Nathwani J. Shared community energy storage allocation and optimization. *Appl Energy* Jul. 2022;318:119160. <https://doi.org/10.1016/j.apenergy.2022.119160>.
- [58] NREL [Online]. Available: <https://www.nrel.gov/solar/market-research-analysis/solar-installed-system-cost.html>. [Accessed 30 October 2024].