Reinforcement learning-based profit maximization for battery energy storage systems with electric vehicles and photovoltaic systems

Dashen Chen[1], Haisheng Li[1], Chun Sing Lai[1,2,], Loi Lei Lai[1]*

[1]Department of Electrical Engineering, School of Automation, Guangdong University of Technology,
Guangzhou 510006, China
[2]Brunel Interdisciplinary Power Systems Research Centre, Department of Electronic and Electrical
Engineering, Brunel University London, London UB8 3PH, UK
**chunsing.lai@brunel.ac.uk*

Keywords: Battery energy storage system, Reinforcement learning, Photovoltaic, Electric vehicles

Abstract

With the growing penetration of renewable energy and the increasing adoption of electric vehicles, the reliable and secure operation of the power grid is facing significant challenges. The inherent randomness and uncertainty associated with renewable energy generation and electric vehicle charging are major factors contributing to grid instability. To address this issue, this paper proposes the utilization of energy storage systems for actively regulating active and reactive power to mitigate grid supply-demand imbalances. Reinforcement learning algorithms are employed to schedule the active and reactive power of the energy storage system, and sensitivity and economic analyses are conducted. The results demonstrate that the integration of energy storage systems into the grid can effectively mitigate the uncertainties and randomness associated with electric vehicle charging and renewable energy generation. The real-time scheduling strategy outputted by the reinforcement learning algorithm reduces computation time, while the economic and sensitivity analyses confirm the profitability and robustness of the energy storage system.

1 Introduction

In response to the global efforts to reduce carbon emissions, countries worldwide are increasing their investments in renewable energy sources. As of 2022, the global installed capacity for wind power has reached 840 GW, while the total installed capacity for photovoltaic (PV) power has reached 1.1 TW. Notably, by the end of 2019, China alone had achieved an installed capacity of 200 GW for both photovoltaic and wind power [1]. Driven by supportive national policies, the integration of uncontrollable renewable energy sources into the power grid is steadily increasing. This growing trend is accompanied by an increase in electric vehicle sales, which introduces greater demand uncertainty on the load side of the grid [2]. To address the challenges posed by the high penetration of renewable energy and electric vehicle loads in the grid, the installation of large-scale energy storage batteries is an effective solution. These batteries can help mitigate the uncertainty and randomness associated with these variable energy sources and provide stability and flexibility to the grid. By incorporating a large-scale energy storage system into the power grid, the uncertainty of load and the intermittency of renewable energy can be effectively addressed through optimal scheduling. To tackle the intermittent nature of renewable power generation and the uncertain load of the electricity network, a virtual energy hub is proposed. This hub integrates electric buses, electric vehicle parking lots, photovoltaic power generation systems, and energy storage systems, creating a dynamic and interconnected system that optimizes energy flow and utilization. Ref. [3] introduces a virtual energy hub as a solution to address the intermittent nature of renewable energy generation and the uncertain load in the electricity network. This hub integrates various components such as electric buses, electric vehicle parking lots, photovoltaic power generation systems, and energy storage systems. By optimizing the coordination and utilization of these elements, the virtual energy hub aims to enhance the stability and efficiency of the overall energy system. The virtual energy hub has been proven effective in energy management, reducing operating costs through extensive testing in various scenarios. Its optimization algorithms and control mechanisms ensure efficient utilization of energy resources, maximizing cost-effectiveness and energy efficiency. The community energy management system proposed in [4], which utilizes multi-intelligence strengthening learning, has demonstrated excellent performance in handling the intermittency of renewable energy and optimizing system economy. Through detailed case analyses, it has been determined that the proposed method effectively addresses the challenges posed by renewable energy integration. However, it is worth noting that as the penetration rate of renewable energy increases, there is a possibility of transmission equipment overload during peak generation periods. Ref. [5] presents two schemes for addressing the challenges posed by the increasing penetration of renewable energy and the random charging load of electric vehicles. The investment costs of these schemes are compared, and relevant suggestions are provided. Additionally, the paper explores the impact of electricity pricing strategies on reducing the randomness of electric vehicle charging behavior.

Ref. [6] focuses on the combined scheduling of electric vehicle pricing and power management in charging stations. The study

demonstrates operational enhancements and efficiency improvements through turnover comparisons and detailed power management analysis.
In conclusion, installing large-scale energy storage batteries in the grid is a favorable solution given the rising penetration of renewable energy and electric vehicle charging loads. The remaining sections of this paper are organized as follows: Section 2 introduces the research problem model, the energy storage battery degradation model, and the economic and sensitivity analysis indicators. Section 3 provides an example analysis of the proposed method. Finally, Section 4 summarizes the findings and conclusions of this study.

2. Methodology

2.1 Model framework based on Markov decision process

A schematic diagram of the Markov decision process is shown in Fig. 1, which contains two parts: the environment (ENV) and the agent. Agent output the corresponding policy based on the status, reward and Isdone output of ENV. This is repeated until all simulation moments are completed [7].

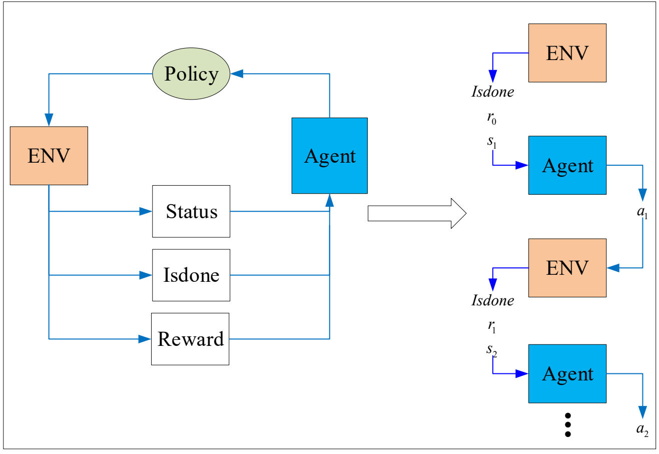


Fig.1 Schematic diagram of Markov decision process

Fig. 2 illustrates the connection of the battery energy storage system (BESS), photovoltaic (PV) power plant, and electric vehicle charging station (EVCS) to node 9. A reinforcement learning agent is employed to generate scheduling strategies for controlling the active and reactive power output of the BESS. The agent makes decisions based on environmental state information, which encompasses grid voltage, photovoltaic power generation, and electric vehicle charging load.

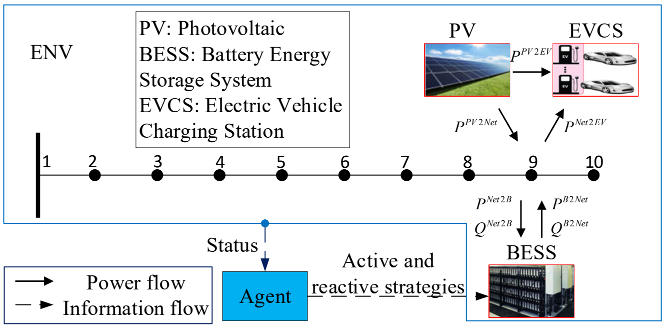


Fig. 2 Battery energy storage system to mitigate the uncertainty of renewable energy and electric vehicle charging station structure framework.

2.2 Methods and models

2.2.1Profit maximization modelling of BESS

$$Z_{BESS} = \max_{P_t^B} \sum_{t=1}^{T} R_t \tag{1}$$

$$R_t = \begin{cases} \alpha_t (P_t^B - P_t^{EV})\Delta t + \beta_t P_t^{EV} \Delta t - C_t^B, \\ \qquad P_t^B > 0, P_t^B > P_t^{EV} \\ \beta_t P_t^B \Delta t - C_t^B, P_t^B > 0, P_t^B \le P_t^{EV} \\ \qquad \alpha_t P_t^B \Delta t - C_t^B, P_t^B < 0 \end{cases} \tag{2}$$

Equation (1) is the objective function, which represents the profit of the battery energy storage system in the whole simulation cycle. Here R_t represents the profit of the battery energy storage system at time t, which can be determined by Equation (2). In Equation (2), P_t^B and P_t^{EV} represent the BESS output power and EVCS load, respectively. α_t and β_t represent the electricity market price and the electric vehicle charging price respectively. C_t^B and Δt represent the degradation cost of the BESS and simulation step size, respectively.

$$L_b \le V_{k,t} \le U_b, \forall k, \forall t \tag{3}$$

$$SOC_{\min} \le SOC_t \le SOC_{\max}, \forall t \tag{4}$$

$$SOC_t = \begin{cases} SOC_{t-1} - \dfrac{\eta_{ch} P_t^B \Delta t}{E_{BESS}}, P_t^B < 0 \\ SOC_{t-1} - \dfrac{P_t^B \Delta t}{\eta_{dis} E_{BESS}}, P_t^B \ge 0 \end{cases}, \forall t \tag{5}$$

Equations (3-5) are the constraints of the problem. Equation (3) is the network voltage constraint, where $V_{k,t}$ represents the voltage at time t of the k node. U_b and L_b are the upper and lower voltage boundaries, respectively. Equation (4) is the constraint condition of the state of charge of the battery energy storage system. SOC_t indicates the state of charge of the battery energy storage system at time t. $SOC_{\max}$ and $SOC_{\min}$ represent the maximum and minimum SOC of the BESS, respectively. Equation (5) represents the SOC updating process of BESS. η_{ch} and η_{dis} represent charging efficiency and discharge efficiency respectively. E_{BESS} represents the capacity of the battery energy storage system.

2.2.2 BESS degradation model

The rain-flow counting method is employed in this study to compute the battery degradation cost. This method simplifies the complex load process into several simple load cycles, which are then used to estimate the fatigue life. Specifically, the rain-flow counting method is utilized to determine the cycle number, cycle depth, and cycle average state of charge

(SOC) of the battery energy storage system (BESS) within one simulation cycle.

$$RC_c = e^{\frac{\ln\left(\frac{100^{0.453}\cdot(100-NDC)}{a_c}\right)}{0.453}}, \forall c \tag{6}$$

$$\begin{aligned} a_c = 3.25\cdot SOC_{mean,c}\cdot(1+3.25\cdot\Delta SOC_c - \\ 2.25\cdot(\Delta SOC_c)^2), \forall c \end{aligned} \tag{7}$$

$$C_{BESS} = \sum_{c=1}^{C}\left(\frac{E_{BESS}\cdot C_{price}}{RC_c}\cdot\Delta SOC_c\right), \forall c \tag{8}$$

$$\lambda_{i+1} = \frac{C_{BESS,i}}{\sum_{t=1}^{T}\left|P_{t,i}^{B}\right|}, \forall i, \forall t \tag{9}$$

$$C_{t,i} = \gamma_i \cdot \left|P_{t,i}^{B}\right|, \forall i, \forall t \tag{10}$$

Equation (6) calculates the equivalent number of full cycles under the c cycle condition. NDC is normalized discharge capacity. When the capacity of the battery energy storage system degrades to the NDC value, it is considered that the battery needs to be replaced. In Equation (7), ΔSOC_c and $SOC_{mean,c}$ are respectively the cycle depth and cycle average calculated by the rain-flow counting method. Equation (8) is the degradation cost of energy storage battery in a complete simulation cycle. Where C_{price} is the unit price of the battery energy storage system. Equation (9) is the degradation coefficient of the battery energy storage system, which is obtained by dividing the degradation cost by the absolute value of the charging and discharging power. Equation (10) is the degradation cost of one simulation step of the battery energy storage system. The degradation cost of a simulation step of the battery energy storage system can be obtained by calculating Equations (6-10).

2.2.3 Transform the research problem into a Markov decision process

Since the reinforcement learning algorithm based on Markov decision process can not directly solve the mathematical model of the problem studied, it is necessary to transform the problem into Markov decision process.

$$s_t = \left(SOC_t, \alpha_t, \beta_t, P_t^{PV}, P_t^{EV}, V_t^{\max}, V_t^{\min}\right) \tag{11}$$

Equation (11) presents the system state output by the environment. Where $V_t^{\max}$ and $V_t^{\min}$ represent the highest and lowest voltages of all nodes in the power network.

$$Isdone = \begin{cases} yes, V_t^{\max} > U_b, V_t^{\min} < L_b, SOC_t > SOC_{\max}, \\ SOC_t < SOC_{\min}, \left(P_t^B\right)^2 + \left(Q_t^B\right)^2 > \left(S^B\right)^2 \\ no, V_t^{\max} \le U_b, V_t^{\min} \ge L_b, SOC_t \le SOC_{\max}, \\ SOC_t \ge SOC_{\min}, \left(P_t^B\right)^2 + \left(Q_t^B\right)^2 \le \left(S^B\right)^2 \end{cases} \tag{12}$$

Equation (12) is the indication of whether the battery energy storage system violates the constraints. When the executed scheduling strategies violates the constraint, Isdone is triggered to end the training. Where P_t^B and Q_t^B represent the active and reactive power output of the battery energy storage system respectively. S^B indicates the maximum apparent power of the battery energy storage system.

Equation (13) is a rewrite of Equation (2) in order to make the objective function more suitable for reinforcement learning algorithm training. By comparing Equations (2) and (13), it can be found that when Isdone is not triggered, a reward will be added, and when Isdone is triggered, it will be punished once and the current round of training will end. In Equation (13), Δr and r_p are the extra reward obtained by completing one simulation step and the penalty for violating the constraint, respectively. After transformation, Equations (11-13) can be solved by reinforcement learning algorithm.

$$R_t = \begin{cases} \alpha_t\left(P_t^B - P_t^{EV}\right)\Delta t + \beta_t P_t^{EV}\Delta t - C_t^B + \Delta r, \\ P_t^B > 0, P_t^B > P_t^{EV}, Isdone = no \\ \beta_t P_t^B \Delta t - C_t^B + \Delta r, P_t^B > 0, P_t^B \le P_t^{EV}, \\ Isdone = no \\ \alpha_t P_t^B \Delta t - C_t^B + \Delta r, P_t^B \le 0, Isdone = no \\ \alpha_t\left(P_t^B - P_t^{EV}\right)\Delta t + \beta_t P_t^{EV}\Delta t - C_t^B + \Delta r - \\ r_p, P_t^B > 0, P_t^B > P_t^{EV}, Isdone = yes \\ \beta_t P_t^B \Delta t - C_t^B + \Delta r - r_p, P_t^B > 0, P_t^B \le P_t^{EV}, \\ Isdone = yes \\ \alpha_t P_t^B \Delta t - C_t^B + \Delta r - r_p, P_t^B \le 0, Isdone = yes \end{cases} \tag{13}$$

2.2.4 Reinforcement learning algorithm and its sensitivity analysis and economic analysis index

In this paper, the TD3 algorithm of reinforcement learning algorithm will be used to solve the model. TD3 algorithm is a reinforcement learning algorithm based on Actor-Critic model. The action space of the algorithm is continuous, so the scheduling strategy formulated is more delicate. Reinforcement learning algorithm has the function of offline training and online execution, so it can realize real-time output scheduling strategy. In the online execution stage, there may be a violation of constraints, and corresponding countermeasures will be made when there is a violation of constraints.

Fig. 3 shows the response plan for constraint violation in the real-time scheduling stage. The main steps are as follows:

1) The scheduling strategies P_t^B and Q_t^B output by the TD3 algorithm are loaded, and then the power flow is calculated.
2) Determine whether the condition of Equation (12) is satisfied. If it is satisfied, output the result directly; otherwise, proceed to the following steps.

3) When the SOC of BESS violates the constraint, the SOC is adjusted to the critical value (upper and lower boundary) by adjusting P_t^B.

4) When the voltage violates the constraint, the voltage is restored to normal by adjusting Q_t^B.

5) After steps 3) and 4), we can ensure that the SOC of BESS and grid voltage are within the appropriate range. In order to ensure that constraint $\left(P_t^B\right)^2+\left(Q_t^B\right)^2 \le \left(S^B\right)^2$ is satisfied, backup power is introduced if necessary. The backup power supply can be used to measure the robustness of the algorithm.

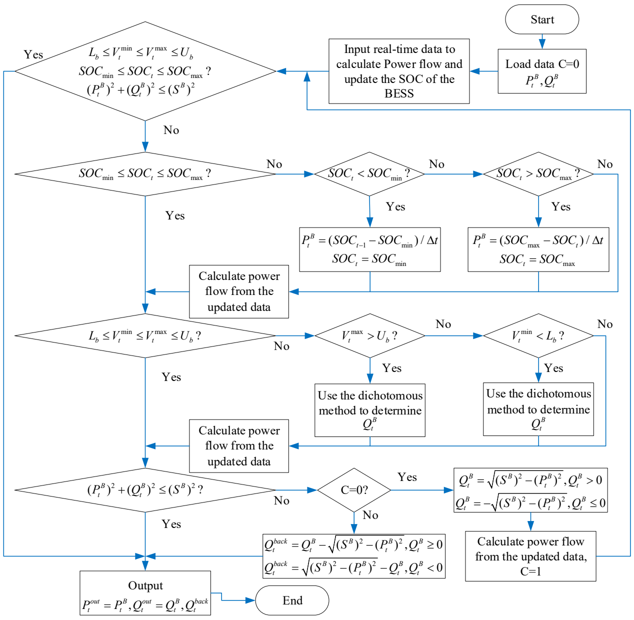

Start
Load data C=0
P_t^B, Q_t^B
Input real-time data to calculate Power flow and update the SOC of the BESS
$L_b \le V_t^{\min} \le V_t^{\max} \le U_b$
$SOC_{\min} \le SOC_t \le SOC_{\max}$?
$(P_t^B)^2+(Q_t^B)^2 \le (S^B)^2$
Yes
No
$SOC_{\min} \le SOC_t \le SOC_{\max}$?
No
$SOC_t < SOC_{\min}$?
No
$SOC_t > SOC_{\max}$?
No
Yes
Yes
Yes
$P_t^B = (SOC_{t-1} - SOC_{\min}) / \Delta t$
$SOC_t = SOC_{\min}$
$P_t^B = (SOC_{\max} - SOC_t) / \Delta t$
$SOC_t = SOC_{\max}$
Calculate power flow from the updated data
$L_b \le V_t^{\min} \le V_t^{\max} \le U_b$?
No
$V_t^{\max} > U_b$?
No
$V_t^{\min} < L_b$?
No
Yes
Yes
Yes
Use the dichotomous method to determine Q_t^B
Use the dichotomous method to determine Q_t^B
Calculate power flow from the updated data
$(P_t^B)^2+(Q_t^B)^2 \le (S^B)^2$?
No
C=0?
Yes
$Q_t^B = \sqrt{(S^B)^2-(P_t^B)^2}, Q_t^B > 0$
$Q_t^B = -\sqrt{(S^B)^2-(P_t^B)^2}, Q_t^B \le 0$
No
Yes
$Q_t^{back} = Q_t^B - \sqrt{(S^B)^2-(P_t^B)^2}, Q_t^B \ge 0$
$Q_t^{back} = \sqrt{(S^B)^2-(P_t^B)^2} - Q_t^B, Q_t^B < 0$
Calculate power flow from the updated data, C=1
Output
$P_t^{out} = P_t^B, Q_t^{out} = Q_t^B, Q_t^{back}$
End

Fig. 3 Violation of constraint response plan

$$Q^{back} = \sum_{t=1}^{T} \left| Q_t^{back} \right| \Delta t \tag{14}$$

$$P_c = \sum_{t=1}^{T} \left| P_t^{out} - P_t^B \right| \Delta t \tag{15}$$

$$Q_c = \sum_{t=1}^{T} \left| Q_t^{out} - Q_t^B \right| \Delta t \tag{16}$$

Equations (14-16) can be used to measure the sensitivity of the algorithm. Equation (14) is the absolute value of the summation of the use of standby reactive power multiplied by the unit time. where Q_t^{back} represents the amount of standby reactive power used at moment t. Equation (15) is the adjusted amount of active power, where P_t^{out} and P_t^B represent the active power output after the adjustment of Fig. 3 and the active power output by the reinforcement learning scheduling strategy, respectively. Equation (16) is the adjusted amount of reactive power, where Q_t^{out} and Q_t^B represent the reactive power output after the adjustment of Fig. 3 and the reactive power output by the reinforcement learning scheduling strategy, respectively.

$$LCOS = \frac{I_o + \sum_{y=1}^{Y} \frac{C_{BESS}^y}{\left(1+WACC\right)^y}}{\sum_{y=1}^{Y} \frac{E_{BESS}^y}{\left(1+WACC\right)^y}} \tag{17}$$

$$NPV = \sum_{y=0}^{Y} \frac{C_y}{\left(1+WACC\right)^y} \tag{18}$$

$$0 = \sum_{y=0}^{Y} \frac{C_y}{\left(1+IRR\right)^y} \tag{19}$$

In order to describe whether the battery energy storage system is profitable or not, this paper uses levelized cost of storage (LCOS), weighted average cost of capital (WACC), net present value (NPV) and internal rate of return (IRR) metrics to describe it. In Equation (17) I_o is the investment cost. where C_{BESS}^y and E_{BESS}^y are the O&M cost and total discharge of the battery storage system in year y, respectively. The IRR index can be obtained by calculating Equations (18-19), which can measure the risk resistance of the project.

The simulation model of the battery energy storage system, the response strategy for the violation of constraints, the sensitivity analysis index and the economic analysis index have been introduced in the previous section, and the next will be the example analysis of the method.

3 Results

3.1 Example of algorithm

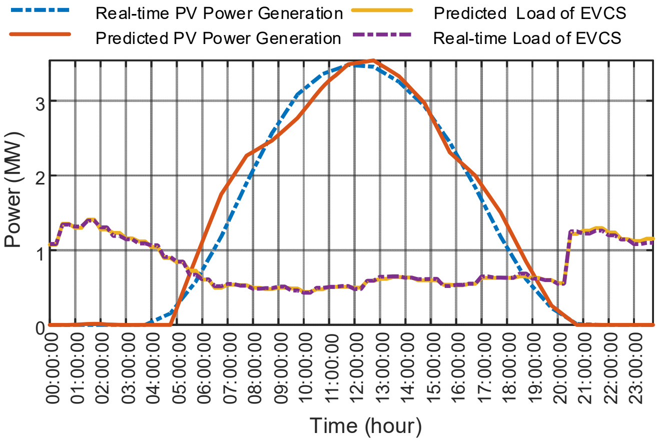


Fig. 4 PV and EVCS prediction curves and real-time curves

Fig. 4 illustrates the prediction and real-time curves of the photovoltaic (PV) power plant and electric vehicle charging station (EVCS). These curves are used for both the offline training and online execution phases of the reinforcement learning algorithm. The system parameters for the examples are presented in Table 1. Fig. 5 displays the real-time and retail electricity prices in the electricity market. The retail price of electricity in this figure includes both the price for demand response participation and the price for non-demand response participation. Furthermore, Fig. 6 shows the EVCS load corresponding to different retail electricity prices. A comparison between Figures 5 and 6 reveal that the load is low when the electricity price is high and high when the electricity price is low.

Table 1 System Parameters

Parameters	Value
O&M	2920 £/MW-yr
NDC	80%
PV installed capacity	4 MW
BESS installed capacity	6 MWh
BESS maximum discharge power	4 MVA
Δt	15 min
Battery unit price	150 £/kWh
Charging and discharging efficiency	95%

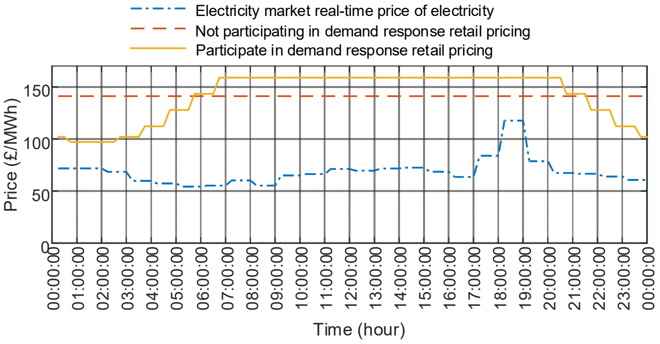


Fig. 5 Electricity market real-time price and retail electricity price

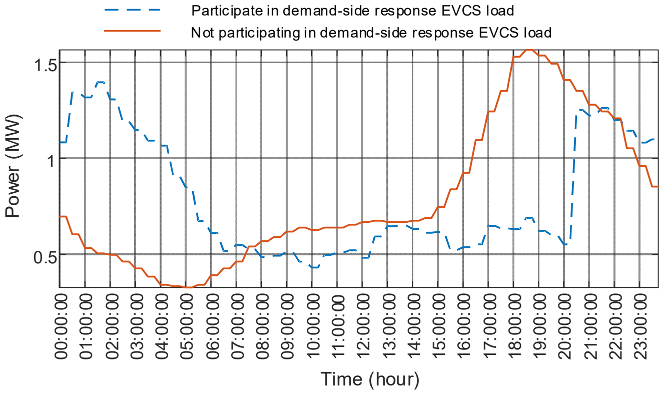


Fig. 6 EVCS load corresponding to different retail electricity prices

3.2 Simulation and analysis of examples

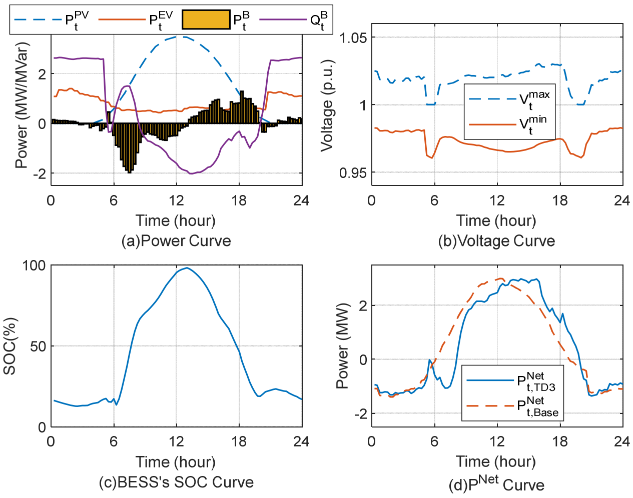


Fig. 7 EVCS participates in demand response TD3 algorithm scheduling results

Fig. 7 shows the results of TD3 dispatch when EVCS participates in demand response, and from Fig. 7(b-c), it can be found that both the grid voltage and the SOC of BESS can be limited to a reasonable range. Fig. 7(d) shows the output power of EVCS, BESS and PV together. From the figure, it can be found that the power curve shifts after the BESS is installed. Observing Fig. 7(a), it is found that the active power output from BESS is charged when the real-time electricity price in the electricity market is low and discharged when the price is high. Fig. 8 shows the results for EVCS without demand response. The price of electricity without demand response is unchanged, so BESS discharges when EVCS load is high and charges when EVCS load is low.

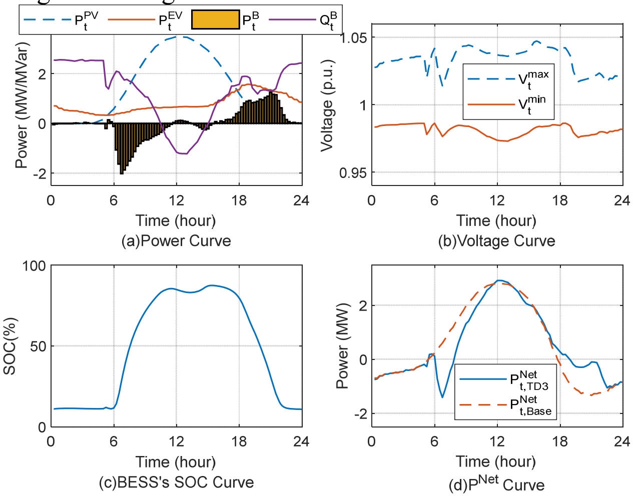
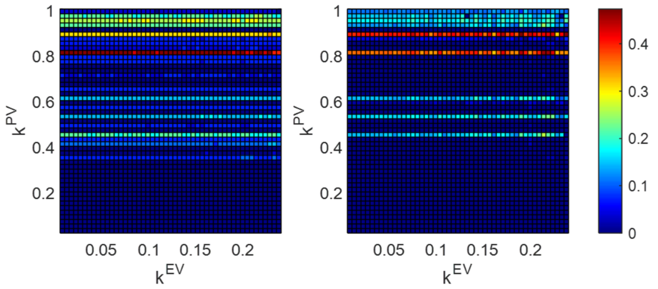


Fig. 8 EVCS does not participate in demand response TD3 algorithm scheduling results



(a) Q_c(MVarh) (b) P_c(MWh)

Fig. 9 Active and reactive power adjustment amount

The maximum EV load error in this paper is set at MAPE of 18.55%, and the square root of the PV prediction error divided by the maximum value of the installed capacity parameter of PV is 10.4%. The prediction errors in Fig. 9 (a-b) increase gradually from left to right and from bottom to top. Fig. 9 corresponds to Equation (15) and (16), respectively, where P_c and Q_c increase as the prediction error increases. Since no standby reactive power is used in the adjustment process, equation (14) is always zero. This indicates that the strategy of reinforcement learning scheduling does not need to use standby reactive power after processing in Fig. 3. Since there is reactive power adjustment, the voltage can be easily stabilized within the set range. This makes the active scheduling strategy tend to be a more economical one, and the economics of BESS is presented next.

Table 2 BESS profit and cost for TD3 algorithm scheduling

Demand Response/£	Charging cost	Degradation costs	O&M	Profits
Participate	368.22	190.27	42	147.83
Non-participation	306.48	144.32	42	150.04

Table 2 depicts the cost and profit of the BESS scheduled by the TD3 algorithm. The table shows that the profit of participating in demand response is slightly lower than the profit of not participating in demand response, because EVCS participating in demand response charge when the electricity price is low thus reducing the dependence on BESS. After Equations (17-19), the values of IRR index were obtained as 11.78% and 12.29% for EVCS participating in demand response and not participating in demand response, respectively.

4 Conclusion

This paper focuses on the utilization of reinforcement learning algorithms to schedule battery energy storage systems (BESS) in order to address the uncertainty and stochastic nature of photovoltaic (PV) generation and electric vehicle charging station (EVCS) loads in the grid. The BESS is carefully modeled, and an economic and sensitivity analysis is conducted. The findings suggest that by employing the TD3 algorithm scheduling and the proposed coping strategies, the BESS can achieve profitability and robustness in the grid. Considering active and reactive power scheduling is helpful to stabilize voltage and improve robustness, especially the effect of reactive power on voltage. The internal rate of return (IRR) index for the BESS, scheduled by the TD3 algorithm, is determined to be 11.78% and 12.29% for EVCS participating and non-participating demand response, respectively.

5 Acknowledgements

The authors acknowledge the financial support from EPSRC Supergen Energy Storage Network+ Early Career Researcher Committee Fund.

6 References

[1] G. Pulazza, N. Zhang, C. Kang, and C. A. Nucci, ‘Transmission Planning with Battery-Based Energy Storage Transportation for Power Systems with High Penetration of Renewable Energy’, *IEEE Transactions on Power Systems*, vol. 36, no. 6, pp. 4928–4940, Nov. 2021, doi: 10.1109/TPWRS.2021.3069649.

[2] C. S. Lai, D. Chen, X. Xu, G. A. Taylor, I. Pisica, and L. L. Lai, ‘Operational Challenges to Accommodate High Penetration of Electric Vehicles: A Comparison between UK and China’, in *2021 IEEE 15th International Conference on Compatibility, Power Electronics and Power Engineering, CPE-POWERENG 2021*, Institute of Electrical and Electronics Engineers Inc., 2021. doi: 10.1109/CPE-POWERENG50821.2021.9501197.

[3] A. Zahedmanesh, K. M. Muttaqi, and D. Sutanto, ‘A Cooperative Energy Management in a Virtual Energy Hub of an Electric Transportation System Powered by PV Generation and Energy Storage’, *IEEE Transactions on Transportation Electrification*, vol. 7, no. 3, pp. 1123–1133, Sep. 2021, doi: 10.1109/TTE.2021.3055218.

[4] B. C. Lai, W. Y. Chiu, and Y. P. Tsai, ‘Multiagent Reinforcement Learning for Community Energy Management to Mitigate Peak Rebounds under Renewable Energy Uncertainty’, *IEEE Trans Emerg Top Comput Intell*, vol. 6, no. 3, pp. 568–579, Jun. 2022, doi: 10.1109/TETCI.2022.3157026.

[5] Y. Wang, J. Chen, J. Liang, Z. Liu, H. Liu, and Z. Shen, ‘Benefits analysis of energy storage system configured on the renewable energy gathering stations’, *Energy Reports*, vol. 9, pp. 802–809, Sep. 2023, doi: 10.1016/j.egyr.2023.04.338.

[6] Q. Huang, L. Yang, C. Zhou, L. Luo, and P. Wang, ‘Pricing and energy management of EV charging station with distributed renewable energy and storage’, *Energy Reports*, vol. 9, pp. 289–295, Mar. 2023, doi: 10.1016/j.egyr.2022.10.418.

[7] Y. Zheng *et al.*, ‘DDPG based LADRC trajectory tracking control for underactuated unmanned ship under environmental disturbances’, *Ocean Engineering*, vol. 271, Mar. 2023, doi: 10.1016/j.oceaneng.2023.113667.