Energy 259 (2022) 124852

Contents lists available at ScienceDirect

Energy

journal homepage: www.elsevier.com/locate/energy

Profit maximization for large-scale energy storage systems to enable fast EV charging infrastructure in distribution networks

Chun Sing Lai [a,b,*], Dashen Chen [a], Jinning Zhang [c], Xin Zhang [b], Xu Xu [d], Gareth A. Taylor [b], Loi Lei Lai [a,**]

[a] Department of Electrical Engineering, School of Automation, Guangdong University of Technology, Guangzhou, 510006, China
[b] Brunel Interdisciplinary Power Systems Research Centre, Department of Electronic and Electrical Engineering, Brunel University London, Kingston Lane, London, UB8 3PH, UK
[c] Center for Propulsion, School of Aerospace, Transport and Manufacturing, Cranfield University, Bedford, MK43 0AL, UK
[d] School of Electrical and Electronic Engineering, Nanyang Technological University, Jurong West, Singapore

ARTICLE INFO

Keywords:
Distribution network optimization
Fast EV charging Demand
Deep reinforcement learning
Battery energy storage systems

ABSTRACT

Large-scale integration of battery energy storage systems (BESS) in distribution networks has the potential to enhance the utilization of photovoltaic (PV) power generation and mitigate the negative effects caused by electric vehicles (EV) fast charging behavior. This paper presents a novel deep reinforcement learning-based power scheduling strategy for BESS which is installed in an active distribution network. The network includes fast EV charging demand, PV power generation, and electricity arbitrage from main grid. The aim is to maximize the profit of BESS operator whilst maintaining voltage limits. The novel strategy adopts a Twin Delayed Deep Deterministic Policy Gradient (TD3) algorithm and requires forecasted PV power generation and EV smart charging demand. The proposed strategy is compared with Deep Deterministic Policy Gradient (DDPG), Particle Swarm Optimization and Simulated Annealing algorithms to verify its effectiveness. Case studies are conducted with smart EV charging dataset from Project Shift (UK Power Networks Innovation) and the UK photovoltaic dataset. The Internal Rate of Return results with TD3 and DDPG algorithms are 9.46% and 8.69%, respectively, which show that the proposed strategy can enhance power scheduling and outperforms the mainstream methods in terms of reduced levelized cost of storage and increased net present value.

1. Introduction

Decarbonization in the transport sector largely accelerates the global uptake of electric vehicles (EVs). By 2030, EV market is estimated to reach 36 million in the UK [1]. The UK government has introduced a series of policies to promote EV deployment [2]. Consumers can receive a government subsidy of up to £2500 for EV purchased in the UK with a range of more than 70 miles on pure electricity and a price of less than £35,000. However, there are a limited number of grants available: 250 grants until March 31, 2022; 1000 grants between April 1, 2022 and March 31, 2023 [3]. The rapid transition and high penetration of EV significantly changes the travelling behaviors, energy usage pattern and consumer engagement and interaction with energy systems; which will have a considerable impact on the planning and operation of electricity networks at both local and national levels. This emerges an urgent need to identify and utilize the advanced energy storage technologies to mitigate the potential of wide-scale blackout caused by power supply and demand imbalance.

The evolution of UK electricity network is essential to integrate the large-scale influx of fast EV charging demand. Electrified transportation sector and electricity network are closely coupled with the development of vehicle-to-grid technology and Internet of Things platforms, which enables intelligent asset management platforms to promote low carbon energy systems. However, with these changes in demand side and supply side, it will become increasingly difficult for a congested electricity network to balance the widespread of fast charging EV requirements. One of the most promising solutions is to use large-scale battery energy

* Corresponding author. Brunel Interdisciplinary Power Systems Research Centre, Department of Electronic and Electrical Engineering, Brunel University London, Kingston Lane, London, UB8 3PH, UK.
** Corresponding author. Department of Electrical Engineering, School of Automation, Guangdong University of Technology, Guangzhou, 510006, China.
E-mail addresses: chunsing.lai@brunel.ac.uk (C.S. Lai), 2112004023@mail2.gdut.edu.cn (D. Chen), jinning.zhang@cranfield.ac.uk (J. Zhang), xin.zhang@brunel.ac.uk (X. Zhang), xu.xu@ntu.edu.sg (X. Xu), gareth.taylor@brunel.ac.uk (G.A. Taylor), l.l.lai@gdut.edu.cn (L.L. Lai).

https://doi.org/10.1016/j.energy.2022.124852
Received 15 March 2022; Received in revised form 22 June 2022; Accepted 14 July 2022
Available online 5 August 2022

Nomenclature

Abbreviations

BEV	Battery Electric Vehicles
BESS	Battery Energy Storage Systems
DDPG	Deep Deterministic Policy Gradient
ELM	Extreme Learning Machine
EV	Electric Vehicles
IRR	Internal Rate of Return
LCOS	Levelized Cost of Storage
MDP	Markov Decision Process
NPV	Net Present Value
NDC	Normalized Discharge Capacity
O&M	Fixed Operation and Maintenance
PSO	Particle Swarm Optimization
PV	Photovoltaic
PHEV	Plug-in Hybrid Electric Vehicles
RMSE	Root Mean Square Error
RL	Reinforcement Learning
SCD	Smart Charging Demand
SA	Simulated Annealing
TD3	Twin Delayed Deep Deterministic Policy Gradient
WACC	Weighted Average Cost of Capital

Symbols

α_t	Wholesale market price (£/MWh)
β_t	Retail electricity price (£/MWh)
τ	Soft update parameter
η_{ch}	Charging efficiency of BESS (%)
η_{dis}	Discharging efficiency of BESS (%)
ΔSOC_c	Deviation of SOC of the BESS in the c^{th} cycle (%)
γ_t	Loss factor (£/MW)
φ	The parameters of actor network
$\varphi^{'}$	The parameters of target actor network
$\theta_{1,2}$	The parameters of two critic network
$\theta^{'}_{1,2}$	The parameters of two target critic network
π_{φ}	Actor network
$\pi_{\varphi'}$	Target actor network
$Q_{\theta_{1,2}}$	Critic network
$Q_{\theta^{'}_{1,2}}$	Target critic network
a_t	Action of agent at time t
$C_{BESS_{Replacement}}$	BESS's degradation cost (£)
$C_{i,t}$	Cost of degeneration of battery at time t at i^{th} iteration (£)
C^B_t	Cost of degeneration of BESS at time t (£)
C_{BESS_y}	Total costs of the BESS at year y (£)
$C_{BESS_{Rated}}$	Capital cost of BESS (£/MWh)
$E_{Cap_{BESS}}$	Rated energy capacity of BESS (MWh)
E_{BESS_y}	BESS energy output at year y (MWh)
I_o	Initial investment cost (£)
$J(\varphi)$	Expected return function
k^{PV}	and k^{EV} PV and EV forecasted uncertainty factors
k^{min}	and k^{max} Flag variables of BESS denoting less than the minimum and above the maximum SOC
L_b	and U_b Lower and upper boundaries of the network voltage
N	Number of samples
N_c	Cycle number of SOC curve
n	BESS lifetime (year)
P^B_t	BESS output power for arbitrage at time t (MW)
P^{total}_t	Total output power of BESS (MW)
P^B_{Max}	BESS maximum output power for arbitrage (MW)
P^{back}_t	BESS output power for reserve power at time t (MW)
P^{PV}_t	PV system output power at time t (MW)
P^{EV}_t	Fast EV charging power at time t (MW)
P^{net}_t	Grid injected power (MW)
$P^{Net}_{t,base}$	Grid injected power without BESS (MW)
$P^{Net}_{t,TD3}$	Grid injected power with BESS scheduled by TD3 (MW)
$P^{Net}_{t,DDPG}$	Grid injected power with BESS scheduled by DDPG (MW)
$P^{Net}_{t,PSO}$	Grid injected power with BESS scheduled by PSO (MW)
$P^{Net}_{t,SA}$	Grid injected power with BESS scheduled by SA (MW)
P_{RMSE}	Forecasted root mean square error of PV generation divided by the generation capacity of PV
$P^{PV}_{t,on-line}$	and $P^{PV}_{t,off-line}$ PV power in on-line implementation and off-line training (MW)
$P^{EV}_{t,on-line}$	and $P^{EV}_{t,off-line}$ EV charging power in on-line implementation and off-line training (MW)
$P^{PV}_{t,noise}$	and $P^{EV}_{t,noise}$ Noise power for PV and EV (MW)
R_t	Reward of agent at time t
$Ratedcycle_c$	Equivalent full cycle number in the c^{th} cycle
R_{noise}	Range of noise
$R^{PV}_{t,stochastic}$	and $R^{EV}_{t,stochastic}$ Random value between −1 and 1 for PV and EV
SOC_t	State of charge at time t (%)
SOC_{min}	Minimum state of charge (%)
SOC_{max}	Maximum state of charge (%)
$SOC_{mean,\ c}$	Average value of SOC in the c^{th} cycle (%)
S^{back}	Amount of reserve energy used (MWh)
T	Maximum simulation time
V^{Min}_t	Minimum value of all node voltages at time t (p.u.)
V^{Max}_t	Maximum value of all node voltages at time t (p.u.)
Z_{BESS}	Profit of BESS (£)

storage systems (BESS) to meet fast EV charging demand. The capital and operational costs of BESS have been significantly reduced in the last decade due to technology advancement and economies of scale. BESS enables active distribution networks management which can effectively maintain system constraints and assist energy and transport sectors decarbonization.

Unmanaged PV power generation and fast EV charging demand can cause serious overvoltage issues in the distribution networks. In Ref. [4], a study on Saudi Arabia's electricity system to charge EVs shows that the system is no longer capable of sustaining a 20% EV penetration rate, or even a 10% EV penetration rate in the worst-case scenario. Ref. [5] considered a micro-grid composed of the power distribution such as wind power and PV, EV charging stations and energy storage systems. The uncertainties of EVs' charging demand and distributed renewable energy output are studied. Ref. [6] discussed that electricity retailers can use energy storage systems or demand response programs to profit from the electricity market and to overcome the stochastic nature of renewable energy. This facilitates the consumption of renewable energy and increases the proportion of clean energy use. Ref. [7] introduced a reinforcement learning method for the power scheduling of BESS for electricity arbitrage in electricity markets. Ref. [8] proposed a model-free deep reinforcement learning algorithm Rainbow Deep Q-Network used to control a battery in a microgrid to perform energy arbitrage and more efficiently utilize solar and wind energy sources. Ref. [9] used reinforcement learning for power scheduling of BESS with PV power generation. But the reviewed research papers do not consider the impact of fast EV charging using BESS on distribution networks. Ref. [10] employed DDPG algorithm for the economic dispatch of

Table 1
Comparison of recent related works.

Ref.	Reinforcement learning for BESS scheduling	Employed TD3 algorithm	Consider EV smart charging	Employed LCOS and IRR to analyze BESS	Consider renewable energy
[5,19]	×	×	×	×	✓
[6]	×	×	✓	×	✓
[7]	✓	×	×	×	×
[8,9]	✓	×	×	×	✓
[10]	✓	×	×	×	×
[11]	✓	×	✓	×	×
[14,15]	×	×	✓	×	✓
This work	✓	✓	✓	✓	✓

electric bus battery swapping stations. By the utilization of BESS in the bus battery swapping stations, the charging and discharging strategy is determined according to the real-time electricity price; such approach can guarantee the sufficient battery replacement and perform arbitrage with the electricity market. However, this reference did not consider the impacts of using batteries to absorb renewable power generation and fast EV charging behaviors on the distribution network. Ref. [11] proposed a deep reinforcement learning based demand response program with real-time pricing signals for EV charging schedule. Authors in Ref. [12] proposed an intelligent energy management and charging scheduling system for EVs. The system provides convenient energy management services by using battery control units and communication infrastructure for charging stations. This system facilitates the drivers to take the best charging decision. Ref. [13] proposed a location planning model of electric bus fast-charging stations for the electric bus transit system. The model considers the bus operation integrated with the distribution networks. Ref. [14] proposed a transactive real-time building energy management strategy considering high penetration of renewable power generation and EV charging demand. Ref. [15] presented a scheduling method for large-scale EV charging in a distribution network, considering random renewable power generation and electricity prices. However, the method did not consider the scenarios where the EV drivers are uncooperative with the demand response planning. Ref. [16] proposed a charging scheduling strategy for different EVs with optimization for the convenience of drivers, performance of transport system and distribution network. Ref. [17] proposed a hybrid strategy to manage the energy in electric vehicle charging station and distribution system. Ref. [18] proposed an optimal EV centralized charging strategy and scheduling algorithm for battery swapping station with an improved hybrid PSO-GA to reduce the power loss and voltage deviation of power networks. Due to the increasing in utilization of intermittent renewable energy resources, it will be more complicated to balance power supply and demand. Authors in Ref. [19] presented a stochastic unit commitment approach for a power system with BESS to facilitate integration of renewable resources, optimize the allocation of system resources, and minimize the system cost. This is a viable solution to deal with the variability and uncertainty of renewable resources by integrating BESS into the distribution network. In Ref. [20], a BESS arbitrage profit maximization problem is formulated considering uncertainty in day-ahead and real-time market prices. The proposed method can effectively help BESS owners to make operational decisions with the aim of maximizing arbitrage profit via market bidding.

Overall, reinforcement learning is a promising approach for profit maximization problem for active distribution networks, but currently has not been extensively explored for BESS, high penetration of PV power, and EV charging demand. Table 1 shows a comparison of recent

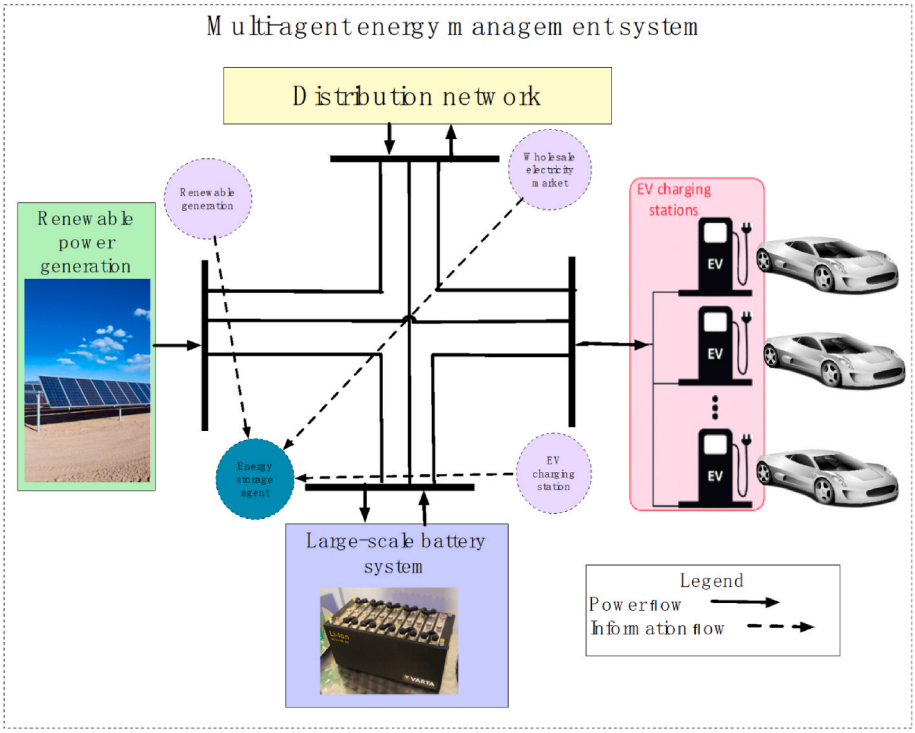

Fig. 1. Power flow and information flow of the distribution network connected with various assets.

related work and this study. It reflects that the present study is very much more comprehensive by including BESS as well.

This paper presents a novel deep reinforcement learning-based power scheduling strategy for BESS which is installed in an active distribution network. The aim is to realize active management of distribution networks and maximize the profit of BESS. Technical constraints of the distribution network will be maintained including voltage limits. The problem of power scheduling is formulated as MDP with continuous action space. The TD3 algorithm, an actor-critic reinforcement learning agent, is utilized to search for an optimal policy that maximizes the cumulative long-term reward, by updating a neural network with gradient ascent method which provides the optimal power scheduling policy. The PV power is forecasted with ELM and the fast EV charging demand is generated by Monte Carlo method. The key contributions of this work are as follows:

- This work presents a novel methodology using TD3 reinforcement learning algorithm to maximize the BESS profit in a distribution network, consisting of fast EV charging demand, wholesale market electricity arbitrage, and PV power generation.
- The proposed methodology is compared with the state-of-the-art reinforcement learning algorithm (i.e., DDPG) and heuristic optimization algorithm (i.e., PSO and SA) using real-life datasets from the UK solar data and EV smart charging demand from Project Shift (UK Power Networks Innovation).
- The techno-economic and financial analysis has been conducted to evaluate the economic and financial viability of installing BESS in distribution networks.

The remaining sections of the paper are described as follows: Section 2 presents a novel deep reinforcement learning-based power scheduling strategy for BESS which is installed in an active distribution network. A techno-economic and financial analysis was conducted focusing on levelized cost of energy, internal rate of return, and net present value. Section 3 presents a case study of the charging and discharging strategy of BESS after it is connected to the distribution network and compares it with a state-of-the-art reinforcement learning algorithm and a heuristic optimization technique. Section 4 presents the conclusions and future work. Appendix presents and examines the datasets used for the model development.

2. Research context and methodology

In this section the research problems and methods will be introduced. The collection of needed datasets i.e., electricity price data, weather data for forecasting PV, and EV charging data are detailed in Appendix A. In this paper, ELM is employed to forecasted PV generation curve. The fast EV charging demand is generated by Monte Carlo method. For a detailed introduction, please refer to Appendix B.

2.1. Overall framework description

In this study, the distribution network operator is considered to have access to BESS owned by a third-party company. The owner of the BESS assumes the obligation to maintain the safe operation of the distribution network and, in return, the owner of the BESS has the right to provide charging services to EVs for profit and to arbitrage in the distribution network. Fig. 1 presents the main components of a distribution network with fast EV charging stations, including PV renewable power generation, and BESS.

The objective of the energy management system is to maximize the profit of the BESS whilst maintain distribution network voltage constraints. The distribution network can charge the BESS during off-peak hours, while at peak hours, the stored power can be recovered and provide electricity to fast EV charging stations to satisfy the fast EV charging demand. The local renewable power generation capacity is mainly used to meet the local EV electricity demand, the excessive part can be stored in battery energy storage or fed into the grid.

In this study, the optimal power scheduling of the active distribution network with BESS is obtained by using reinforcement learning (RL) algorithm. Different from conventional optimization methods, RL has excellent decision-making capability and considerable merits. Firstly, RL algorithm can achieve the optimal actions by interacting with the environment, so it does not require any prior knowledge of initial environment state, which is difficult to acquire in a real environment. Secondly, RL algorithm can be flexibly applied to application scenarios with uncertainty problems, through off-line training and on-line implementation. Thirdly, RL is easier to apply to real-time scenarios because it can use trained neural networks to determine the best results.

In reinforcement learning algorithms, an agent is trained to complete a task within an uncertain and complex environment. Agents are parameterized function approximators for training the policy, which can be classified as actor and critic. The aim of the actor neural networks is to find the sequence of actions that maximizes the cumulative long-term reward of the task, and the critic neural networks is to develop the expected value of the cumulative long-term reward. The agent receives observations and a reward from the environment, and then acts according to the environment. The reward is a measure to assess the performance of actions in a given environmental state, so that desired goals can be achieved by setting reasonable reward function.

For value-based methods, agents use only critic neural networks to determine their actions, which perform better in discrete action spaces but are computationally demanding in continuous action spaces. For policy-based method, agents use only actor neural networks to determine their actions, which are normally employed for continuous action spaces. For actor-critic methods, agents that use both actor and critic neural networks can handle both continuous and discrete action spaces. During the training process, the actor neural networks use the feedback from the critic neural networks to select the optimal actions. At the same time, the critic neural networks learn the value function from rewards to evaluate the actor properly.

The following action can be taken from actor neural networks to control the power flow for various assets in the distribution network:

- Electricity distribution networks transmit electricity from the medium voltage transmission grid to end users, which can charge EV and the BESS.
- Renewable electricity generated from PV systems can be 1) stored in BESS; 2) used for fast EV charging; 3) fed into the distribution network and sold to the wholesale market.
- The BESS can be:
  1. Charged by PV power generation system or distribution network.
  2. Discharged by fast EV charging or fed into the distribution network and sold to the wholesale market. A profit can be made via electricity arbitrage, i.e., storing electricity at off-peak hours when the wholesale market price is low and discharging electricity at peak hours when the wholesale market price is high.

This work considers the fast EV charging demand as a flexible load and smart charging is achieved via dynamic tariffs. To simplify the study, it is assumed that the electricity from EV will not be fed into the grid.

2.2. Mathematical formulation of the optimization problem

To maximize the profit of the BESS with high PV penetration and fast EV charging demand, the optimization problem is formulated with the objective function given in Equation (1) as follows:

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

$$R_t = \begin{cases} \alpha_t\left(P_t^B - P_t^{EV}\right)\Delta t + \beta_t P_t^{EV}\Delta t - C_t^B & if:\ P_t^B > 0, P_t^B > P_t^{EV} \\ \beta_t P_t^B \Delta t - C_t^B & if:\ P_t^B > 0, P_t^B \le P_t^{EV} \\ \alpha_t P_t^B \Delta t - C_t^B & if:\ P_t^B < 0 \end{cases} \quad (1)$$

To ensure that the distribution network operates within technical constraints, the constraints considered in the optimization problem are shown in Equations (2)–(4) as follows:

$$L_b \le V_{k,t} \le U_b\ , \forall k, \forall t \quad (2)$$

$$SOC_{min} \le SOC_t \le SOC_{max}\ , \forall t \quad (3)$$

$$\begin{cases} P_t^{net} = P_t^{total} + P_t^{PV} - P_t^{EV}\ , \forall t \\ P_t^{total} = P_t^{back} + P_t^B\ , \forall t \end{cases} \quad (4)$$

In Equation (1), Z_{BESS} is the profit of BESS. α_t and β_t are wholesale market price and retail electricity price, respectively. P_t^B, P_t^{PV}, and P_t^{EV} are BESS output power for arbitrage, PV system output power ($P_t^{PV} \ge 0$) and fast EV charging power ($P_t^{EV} \ge 0$) at time t, respectively. C_t^B is degradation cost of BESS at time t. Δt and R_t are the time slot and the reward at time t, respectively. In Equation (1), when the BESS discharging power P_t^B is less than the power required by fast EV charging P_t^{EV}, all the BESS discharging power P_t^B will be used for fast EV charging. When the BESS discharge power P_t^B is greater than the fast EV charging demand, the excessive part will be fed into the grid. In other words, the BESS prioritizes discharge to support fast EV charging and transmits power to the grid when the discharge power is greater than the fast EV charging demand. In the study, U_b and L_b are 1.05 and 0.95, respectively. SOC_{max} and SOC_{min} are 100% and 10% [21], respectively. Equations (2) and (3) ensure that the voltage of the node k and SOC of BESS are within the allowable range at time t. Equation (4) is the power balancing equation for the considered assets. P_t^{back} is the portion of the BESS allocated for reserve power, and the P_t^B is the portion of the BESS allocated for participation in the power market arbitrage. Equation (5) presents the SOC update process for the BESS. η_{ch} and η_{dis} are the charging and discharging efficiencies of BESS, respectively. And the conversion efficiency of BESS is taken as 95% [22]. $E_{Cap_{BESS}}$ is the rated energy capacity of BESS.

$$SOC_t = \begin{cases} SOC_{t-1} - \dfrac{\eta_{ch} P_t^B \Delta t}{E_{Cap_{BESS}}} & if:\ P_t^B < 0 \\ SOC_{t-1} - \dfrac{P_t^B \Delta t}{\eta_{dis} E_{Cap_{BESS}}} & if:\ P_t^B \ge 0 \end{cases} \quad (5)$$

The reinforcement learning model uses forecasted PV power generation and fast EV charging demand datasets in the training process. It is worth mentioning that the PV generation is susceptible to weather and fast EV charging load is forecasted by demand response. It is worth studying the volatility of PV generation and fast EV charging demand on the network. The employment of Equations (6) and (7) produces a forecast dataset with forecast errors. The PV forecasted dataset and the fast EV charging demand forecasted dataset in this study will be generated by this method.

$$\begin{cases} P_{t,on-line}^{PV} = P_{t,off-line}^{PV} + P_{t,noise}^{PV}\ , \forall t \\ P_{t,noise}^{PV} = k^{PV} * R_{t,stochastic}^{PV}\ , \forall t \\ R_{t,stochastic}^{PV} \in [-1, 1]\ , \forall t \end{cases} \quad (6)$$

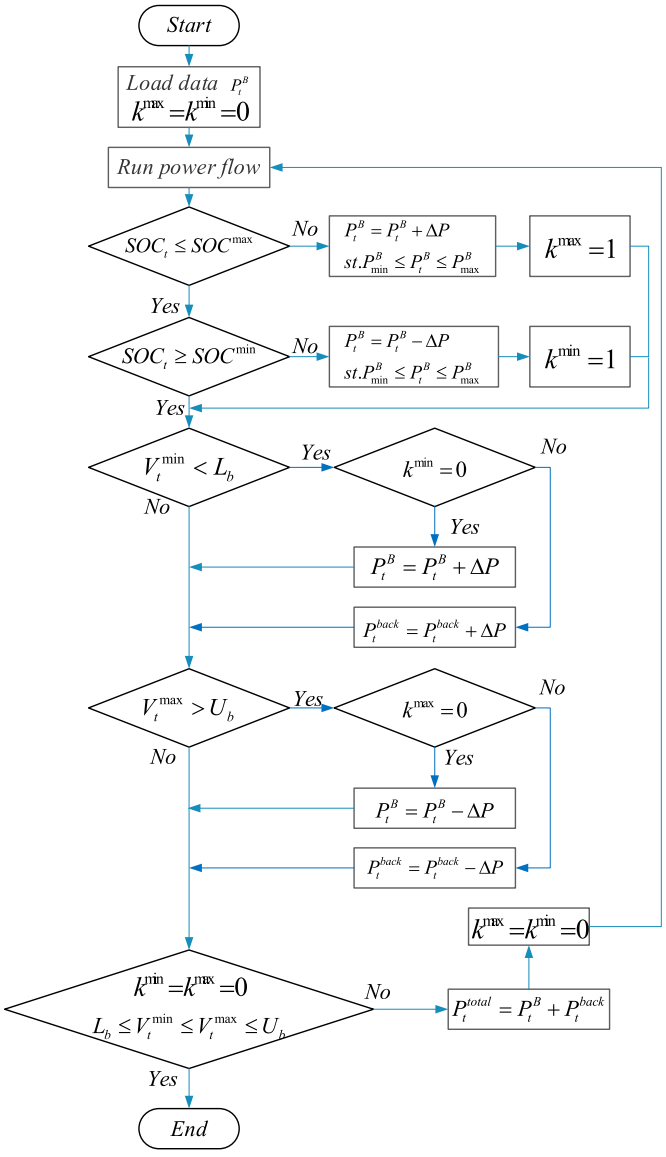

Fig. 2. The flowchart of real-time optimization method.

$$\begin{cases} P_{t,on-line}^{EV} = P_{t,off-line}^{EV} + P_{t,noise}^{EV}\ , \forall t \\ P_{t,noise}^{EV} = k^{EV} * R_{t,stochastic}^{EV}\ , \forall t \\ R_{t,stochastic}^{EV} \in [-1, 1]\ , \forall t \end{cases} \quad (7)$$

In Equation (6) $P_{t,on-line}^{PV}$ and $P_{t,off-line}^{PV}$ are the PV power in on-line implementation and off-line training, respectively. Here $P_{t,noise}^{PV}$ is assumed to be a noise factor, simulating the error in the forecast (to study generalization capability). $R_{t,stochastic}^{PV}$ is a random number between −1 and 1. Different forecasting accuracies are obtained by adjusting the size of k^{PV}. The larger the k^{PV} value is, the larger the PV forecasting error will be. k^{PV} is used to measure the PV forecasting error that RL can accept. That is, the larger the k^{PV} value that RL can accept, the better the generalization capability is. Equation (7) works similarly to Equation (6).

In real-time operation, there is a forecasting error that causes the day-ahead scheduling strategy to fail to meet the operational requirements. When the forecasting error is too large and exceeds the generalization capability of RL, measures need to be taken to ensure the reliable operation of the grid. For the stochastic optimization algorithm

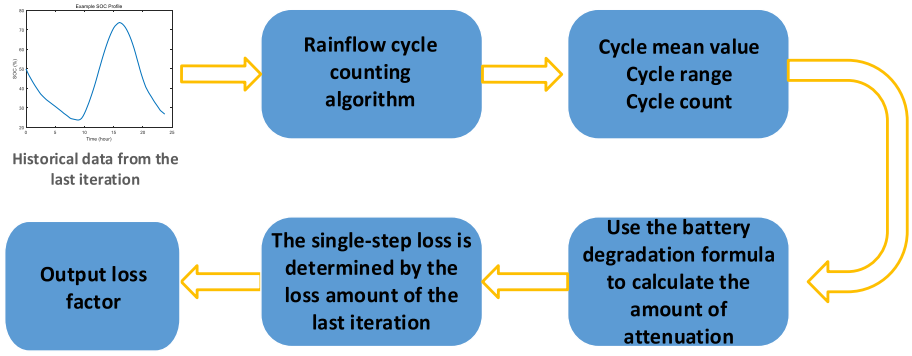

Fig. 3. Battery capacity loss calculation methodology.

that does not have the ability to output the scheduling policy in real-time, it is more important to take measures to ensure the reliable operation of the network. The method is shown in Fig. 2 and P_t^B represents the day-ahead discharging power of the BESS for arbitrage, calculated by the stochastic optimization algorithms i.e., PSO and SA. For RL i.e., TD3 and DDPG P_t^B represents the discharge power of the BESS for arbitrage calculated by RL in real-time. The ΔP is a very small positive real number. The P_t^{back} is the power provided by reserve power supply. P_t^{back} can be used as a quantitative criterion in the sensitivity analysis. Equation (8) defines the amount of reserve power used in a cycle T to quantify the sensitivity. The more sensitive the system is, the larger the value of S^{back} in Equation (8) will be. From Equation (8), the worst case is when $P_t^{back} \geq 0, \forall t$ or $P_t^{back} \leq 0, \forall t$. In this case, it is necessary to prepare the minimum reserve capacity of twice the S^{back} and the SOC of the reserve energy storage is 50%.

$$S^{back} = \sum_{t=1}^{T} \left| P_t^{back} \right| \Delta t \tag{8}$$

The real-time optimization method in Fig. 2 is divided into three stages as follows:

Stage 1) The charging strategy of BESS, the real-time PV generation power and EV charging power are loaded. Then run the power flow of the network at time t.
Stage 2) To ensure that the SOC of the BESS is within the set range, the discharge power P_t^B of the BESS is properly adjusted.
Stage 3) To ensure that the voltage of the grid is within the set range, adjust the discharge power P_t^B of the BESS and the power of the reserve power P_t^{back} without violating Stage 2.

2.3. Battery degradation model

A battery degradation model is included to enhance the calculation accuracy of the BESS operating cost [23]. Fig. 3 shows the calculation methodology of battery capacity loss factor [7]. This model uses the rainflow cycle counting algorithm to obtain cycle count, cycle range and cycle mean value based on the battery SOC curve as shown in Fig. 3. These parameters are then used to calculate the battery capacity loss factor.

Equations 9 and 10 are used to determine the SOC parameters. $SOC_{mean,\,c}$ and ΔSOC_c are SOC mean value of BESS and the deviation of BESS's SOC in the c^{th} cycle, respectively. These parameters are calculated by rainflow cycle counting algorithm.

$$SOC_{mean,\,c} = Cycle\ mean\ value \tag{9}$$

$$\Delta SOC_c = Cycle\ range \tag{10}$$

Equations 11 and 12 are used to calculate $Ratedcycle_c$, the battery equivalent full cycle of the c^{th} discharge cycle [24]. Fig. 4 shows the equivalent full cycle and normalized discharge capacity (NDC) as calculated by Equation (11) [24]. Fig. 4 shows that when the equivalent full cycle times of the battery reach about 4000 times, the corresponding NDC is 80%. The battery will be replaced when the available capacity degrades below the NDC value i.e., 80% [23]. The battery degradation model of Equation (11) has been used in Refs. [23,25] to describe the BESS degradation. For example, Ref. [23] examined a 2 MW/5 MWh BESS and Ref. [25] examined a 0.15 MW/0.6 MWh BESS.

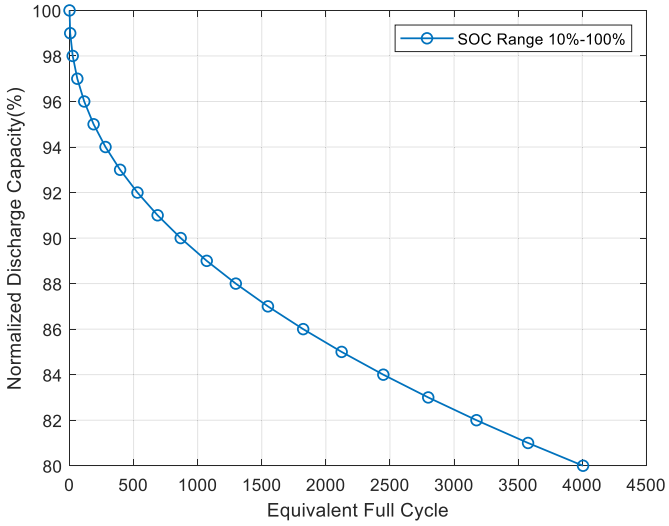

Fig. 4. Equivalent full cycle and normalized discharge capacity (NDC).

$$Ratedcycle_c = e^{\frac{ln\left(\frac{100^{0.453}*(100-NDC)}{aees_c}\right)}{0.453}} \tag{11}$$

$$a_{ees_c} = 3.25*SOC_{mean,\,c}*\left(1 + 3.25*\Delta SOC_c - 2.25*\Delta SOC_c^{\,2}\right) \tag{12}$$

$$C_{BESS_{Replacement}} = \sum_{c=1}^{N_c} \frac{E_{Cap_{BESS}} * C_{BESS_{Rate}} * \Delta SOC_c}{Ratedcycle_c} \tag{13}$$

In Equation (13), $C_{BESS_{Replacement}}$ is the BESS-degradation cost. N_c is the cycle number of SOC curve, which is calculated by rainflow cycle counting algorithm. The battery capacity loss factor is calculated with Equation (14) [7].

$$\gamma_{i+1} = \frac{C_{BESS_{Replacement},i}}{\sum_{t=1}^{T} \left| P_{i,t}^B \right|}, \forall i \tag{14}$$

$C_{BESS_{Replacement},i}$ is the battery degradation cost at i^{th} iteration. $P_{i,t}^B$ is the BESS output power in time t. T is equal to 96, which means that a day is divided into 96 time intervals of 15 min each.

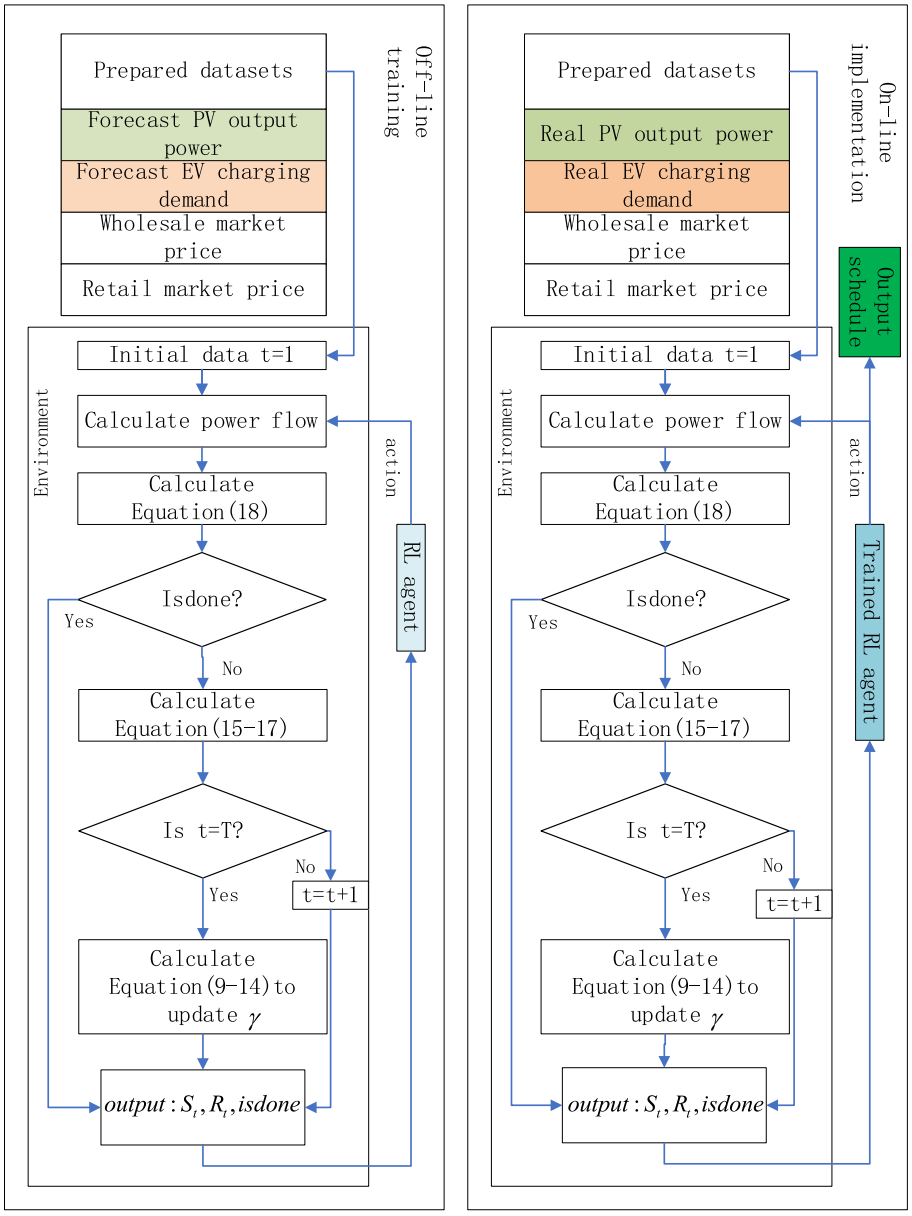

Fig. 5. Off-line training and on-line implementation flowchart.

$$C_{i,t} = \gamma_i * \left| P_{i,t}^{B} \right| \tag{15}$$

In Equation (15), $C_{i,t}$ (£) is the battery degradation cost at time t in the i^{th} iteration. In Equation (1), $C_t^B = C_{i,t}, \forall i$.

2.4. Reinforcement learning algorithm and optimization process

The agent for the BESS is trained by TD3 algorithm [26], which has three inputs i.e., observation, reward, and isdone; and one output which is the action. Observation is used to observe the state of the environment. The state variables in this environment include BESS's SOC at time t SOC_t, wholesale market price α_t, retail electricity price β_t, PV power generation P_t^{PV}, EV smart charging demand P_t^{EV}, the voltage value V_t^{Max} of the highest voltage node in the whole distribution network, and the voltage value V_t^{Min} of the lowest voltage node in the whole distribution network. The state space is given in Equation (16).

$$s_t = \left(SOC_t, \alpha_t, \beta_t, P_t^{PV}, P_t^{EV}, V_t^{Max}, V_t^{Min}\right) \tag{16}$$

The reward is specified according to the objective function in Equation (1). The reward function R_t is expressed in Equation (17). The purpose of the RL agent is to maximize the long-term cumulative reward.

$$R_t = \begin{cases} \alpha_t\left(P_t^B - P_t^{EV}\right)\Delta t + \beta_t P_t^{EV}\Delta t - C_t^B & if:\ P_t^B > 0, P_t^B > P_t^{EV} \\ \beta_t P_t^B \Delta t - C_t^B & if:\ P_t^B > 0, P_t^B \le P_t^{EV} \\ \alpha_t P_t^B \Delta t - C_t^B & if:\ P_t^B \le 0 \end{cases} \tag{17}$$

Isdone will be triggered when the constraints in Equations (2) and (3) are violated. The condition for isdone to be triggered is shown in Equation (18).

$$isdone = \begin{cases} Yes & if:\ V_t^{Max} > U_b, V_t^{Min} < L_b, SOC_t > 1, SOC_t < 0.1 \\ No & if:\ V_t^{Max} \le U_b, V_t^{Min} \ge L_b, SOC_t \le 1, SOC_t \ge 0.1 \end{cases} \tag{18}$$

Action a_t denotes the charging and discharging power of the BESS as shown in Equation (19). Here, P_{Max}^B is the maximum charging and discharging power of BESS for arbitrage.

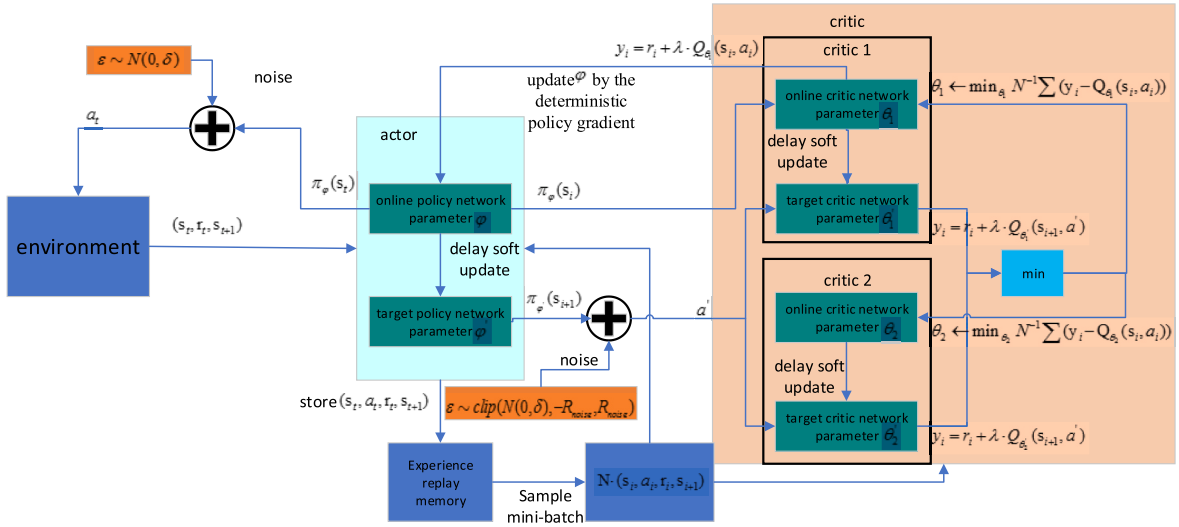

Fig. 6. Reinforcement learning agent of TD3 based on actor-critic approach.

$$a_t \in \left[-P_{Max}^{B}, P_{Max}^{B} \right], \forall t \tag{19}$$

Fig. 5 shows the off-line training and on-line implementation flowchart. The main process of off-line training is as follows: Firstly, the ELM and Monte Carlo methods are used to generate forecast datasets, including PV power generation and the fast EV charging demand. Secondly, the power flow is calculated using the power generated by the PV and fast EV charging demand as well as the charging and discharging power of the BESS (strategy for RL agent output). Thirdly, the Boolean variable isdone state is determined according to Equation (18). Isdone is triggered to terminate the training process due to constraints violation. Finally, the reward value and the state at the next moment are calculated and output to RL agent, and the battery capacity loss factor γ is updated at the last moment of the simulation cycle. The on-line implementation is similar to off-line training, except that the forecasted PV power generation and fast EV charging power data in off-line training are replaced with real-time data. The original RL agent is replaced with the trained RL agent. In this case, the BESS charging and discharging strategy can be output directly through the simulation without the need to train again, which greatly improves the speed of optimization.

A MDP model with discrete time step Δt is established for the BESS power scheduling problem. The model ensures the safe operation of distribution network with absorption of surplus renewable power generation, and providing electricity for fast EV charging stations. In this work, the discrete time step Δt is 15 min.

TD3 is a reinforcement learning algorithm based on actor-critic methods with the following procedures:

1) Initialize the network parameters $\varphi, \theta_1, \theta_2$ for actor network π_φ and two critic networks Q_{θ_1}, Q_{θ_2} respectively. The initialization parameters are then given to the target network $\theta_1 \rightarrow \theta_1'$, $\theta_2 \rightarrow \theta_2'$, $\varphi \rightarrow \varphi'$.
2) Select action $a \sim \pi_\varphi(s) + \varepsilon$, $\varepsilon \sim N(0,\delta)$, here s is the states, ε is noise. The action is applied to the environment to get the reward r and state s' of the next time step. The transition tuple (s, a, r, s') is saved in B.
3) When B has stored enough transition tuples (s, a, r, s'), mini-batch sampling of N transitions (s, a, r, s') from B is adopted to update critic network and actor network. The goal of updating the critic network is given in Equation (20). R_{noise} in Equation (20) is the range of noise and $R_{noise} > 0$. The parameters of the critic network are updated with Equation (21).

$$\begin{cases} y = r + \lambda Q_{\theta'}(s', \pi_{\varphi'}(s') + \varepsilon) \\ \varepsilon \sim clip(N(0, \widetilde{\delta}), -R_{noise}, R_{noise}) \end{cases} \tag{20}$$

$$\theta_i \leftarrow min_{\theta_i} N^{-1} \sum \left(y - Q_{\theta_i}(s,a) \right)^2 \tag{21}$$

4) Actor network parameters are updated every d step which can make the value estimation of critic network more accurate. Parameters φ of the actor network are updated by the deterministic policy gradient. $J(\varphi) = \mathrm{E}_{s_i \sim p_\pi, a_i \sim \pi}[R_0]$ in Equation (22) is the expected return function. The aim of updating φ is to maximize the expected return function $J(\varphi)$. Then parameters φ' of the target actor network and parameters θ_i' of the target critic networks are updated by soft updating Equation (23).

$$\nabla_\varphi J(\varphi) = N^{-1} \sum \nabla_a Q_{\theta_i}(s,a)|_{a = \pi_\varphi(s)} \nabla_\varphi \pi_\varphi(s) \tag{22}$$

$$\begin{cases} \theta_i' \leftarrow \tau\theta_i + (1-\tau)\theta_i' \\ \varphi' \leftarrow \tau\varphi + (1-\tau)\varphi' \end{cases} \tag{23}$$

The training of TD3 algorithm as an iterative process is depicted in Fig. 6. The TD3 algorithm uses neural networks to store the learned strategy for maximizing the cumulative reward in an uncertain and complex environment. The algorithm "TD3 with ENV" describes the iteration process of training an agent using TD3 algorithm.

Algorithm.

Algorithm TD3 with ENV

1: Initialize critic networks Q_{θ_1}, Q_{θ_2}, and actor network π_φ with random parameters θ_1, θ_2, φ
2: Initialize target networks $\theta_1 \rightarrow \theta_1'$ $\theta_2 \rightarrow \theta_2'$, $\varphi \rightarrow \varphi'$
3: Initialize replay buffer B
4: Initialize *isdone* and *SOC*
5: **for** $t = 1$ **to** T **do**
6: Select action with exploration noise $a \sim \pi(s) + \varepsilon$, $\varepsilon \sim N(0,\delta)$
7: The environment receives the action a of the agent and calculates the power flow of the distribution network
8: Update Equations 15–18, get *isdone*, r and new state s'
9: Store transition tuple (s, a, r, s') in B
10: Sample mini-batch of N transitions (s, a, r, s') from B
11: $\widetilde{a} \leftarrow \pi_{\varphi'}(s) + \varepsilon$, $\varepsilon \sim$ clip $(N(0, \widetilde{\delta}), -c, c)$
12: $y \leftarrow r + \lambda min_{i=1,2} Q_{\theta_i'}(s', \widetilde{a})$
13: Update critics $\theta_i \leftarrow min_{\theta_i} N^{-1} \sum (y - Q_{\theta_i}(s,a))^2$
14: **if** *isdone* is *yes* **then**
15: **break,** end the current episode and move on to the next episode
16: **end if**
17: **if** t is T **then**
18: Solve Equations 9–14, update the battery loss factor γ_{i+1}

(*continued on next page*)

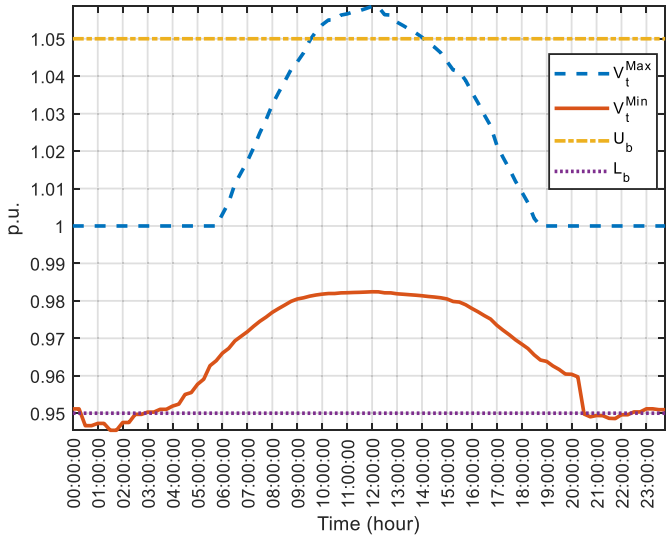

Fig. 7. The V_t^{Max} and V_t^{Min} curves of the network without the BESS.

(*continued*)

Algorithm TD3 with ENV

end if
if t mod d **then**
  Update φ by the deterministic policy gradient:
  $\nabla_{\varphi} J(\varphi) = N^{-1} \sum \nabla_a Q_{\theta_1}(s,a)|_{a=\pi_\varphi(s)} \nabla_\varphi \pi_\varphi(s)$
  Update target networks:
  $\theta_i' \leftarrow \tau\theta_i + (1-\tau)\theta_i'$
  $\varphi' \leftarrow \tau\varphi + (1-\tau)\varphi'$
end if
end for

2.5. Economic and financial analysis

The economic analysis mainly considers the levelized cost of storage (LCOS) and net present value (NPV) of BESS [27]. Considered by investors and decision makers, the BESS costs can be evaluated via LCOS which is commonly examined as a reference for energy storage economic. The battery lifetime, cost, and electricity price will all affect the LCOS. In this analysis, the battery degradation cost is also considered, and the LCOS of the battery is calculated according to the service life, service cost and the income. The financial analysis of the battery can be evaluated via NPV and IRR, which are common financial indicators for investors. The investment of BESS requires a large capital expenditure.

2.5.1. LCOS

LCOS can realize the economic comparisons between different types of energy storage technologies as well as operating scenarios as [22,28]:

$$LCOS = \frac{I_o + \sum_{y=1}^{n} \frac{C_{BESS_y}}{(1+WACC)^y}}{\sum_{y=1}^{n} \frac{E_{BESS_y}}{(1+WACC)^y}} \tag{24}$$

In Equation (24), I_o is the initial investment cost. $WACC$, C_{BESS_y} and E_{BESS_y} are the weighted average cost of capital (WACC), the operation and maintenance cost of BESS in year y and energy output, respectively.

2.5.2. NPV and IRR

NPV can be used to evaluate the financial viability of the project. A positive NPV denotes that the project is worthy for investment and profit can be made. The NPV is the present value of current and future benefit minus the present value of current and future costs. The NPV and IRR are calculated by Equations (25) and (26), respectively as follows: Where n is the total number of years. C_y is the cash flow for the BESS at the year y.

$$NPV = \sum_{y=0}^{n} \frac{C_y}{(1+WACC)^y} \tag{25}$$

$$0 = \sum_{y=0}^{n} \frac{C_y}{(1+IRR)^y} \tag{26}$$

3. Case study: EV smart charging in the UK distribution networks

3.1. Distribution network and node voltage evaluation instructions

Fig. B2 in Appendix B shows the forecasted and real curves for PV and EV smart charging demand. Ideally, any forecasting model can be used for training the RL. Of course, a higher accuracy forecasting model will provide better actions. In Fig. 7, V_t^{Max} denotes the voltage value of the node with the highest voltage among all nodes in the distribution network at time t; V_t^{Min} denotes the voltage value of the node with the lowest voltage among all nodes in the distribution network at time t. The advantage of this definition is that if V_t^{Max} and V_t^{Min} do not exceed the voltage limit, then the voltage of all nodes in the distribution network will not exceed the voltage limit. U_b and L_b denote the upper voltage boundary and lower voltage boundary, respectively. In Fig. 7, without BESS in the distribution network, as PV power generation is generally active in midday from 10:00 to 14:00, the PV power generation can be greater than the fast EV charging demand, so V_t^{Max} will exceed the voltage upper boundary U_b during this period. From 20:30 to 02:30 in the next day, V_t^{Min} is lower than the voltage lower boundary L_b. This will damage the safe and reliable operation of the power grid.

Fig. 8 presents the single-line diagram of a 10-bus distribution network for the case studies. Node 1 is the reference node in power flow analysis [29]. The power flow of the network was calculated by MATPOWER 5.0 with Newton-Raphson method. The BESS, PV, and SCD are connected to node 9. The BESS is installed to reduce the negative impacts on distribution network caused by fast EV charging demand and high penetration of renewable energy generation i.e., PV.

As discussed in Ref. [30], Li-ion batteries have achieved a capital cost

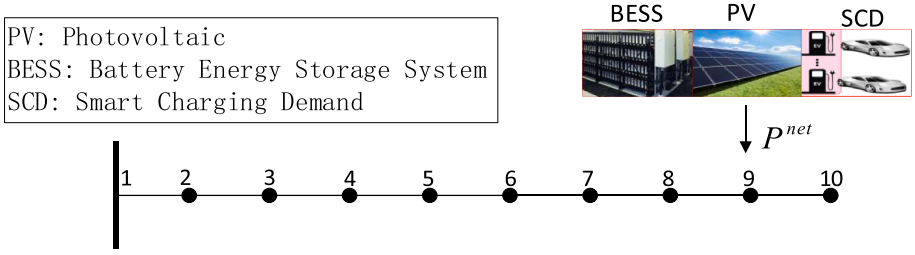

Fig. 8. A 10-bus distribution network.

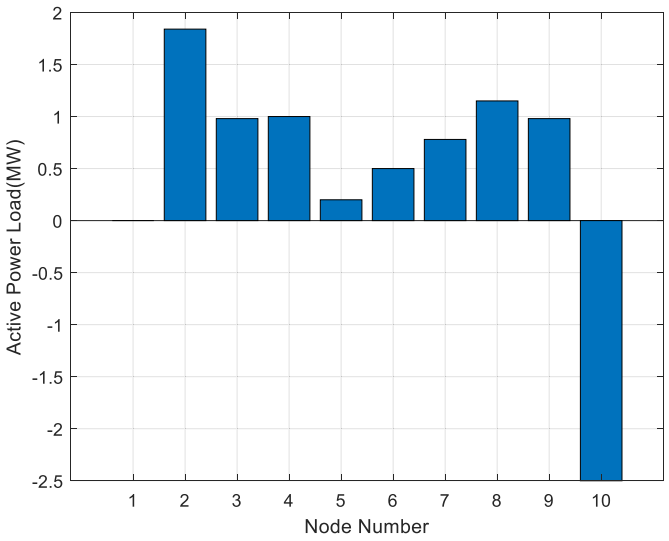

Fig. 9. The fixed load demand at different nodes of the distribution network.

Table 2
System parameters assumption.

Parameters	Value
Fixed operation and maintenance (O&M)	2920 [£/MW-yr] [34]
NDC	80% [23]
PV system rated capacity	4 MW
BESS rated energy capacity	6 MWh [31,32]
BESS rated power capacity	4 MW [31,32]
Δt	15 min
Capital cost of BESS $C_{BESS_{Rated}}$	150 £/kWh [30]
Charging efficiency η_{ch} /Discharging efficiency η_{dis}	97%/97% [22]

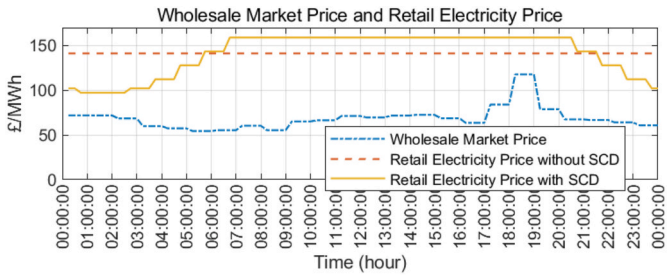

Fig. 10. Wholesale market price and retail electricity price with and without SCD.

of 150 £/kWh. The energy capacity of BESS is set to 6 MWh and the rated power of the BESS is set to 4 MW [31,32]. It is assumed that the PV power capacity is at 4 MW. There are already 10 MW PV power plants in UK [33], so it is possible to set up 4 MW of PV power generation. Considering that the maximum load of the distribution network is 12.37 MW [29], the maximum power of the charging station at a node is set to 1.5 MW [29]. Therefore, the environment to test the RL algorithm can be described as in Fig. 8 and Equations 16–18. The assumed values of the fixed load of the distribution network are shown in Fig. 9. The purpose of assuming these parameters of distribution network is to better show the impact of PV generation and EV loads on the grid voltage which is a minor modification from Ref. [29]. The system assumption parameters are shown in Table 2.

3.2. Experiment with TD3 algorithm

The agent is trained with forecasted PV power generation and fast EV charging demand datasets. Fig. 10 presents the wholesale electricity market price and retail electricity price with and without SCD. Subsequently, the trained agent can achieve optimal power scheduling in real-time. Fig. 11(a) shows that the BESS is charged between 06:00 and 16:00. In this period, the voltage of the distribution network in Fig. 11 (b) increases because of abundant PV power generation. However, the BESS discharges although the PV power generation exists at 16:00. This is because BESS's SOC at 16:00 as shown in Fig. 11(c) is relatively high and the distribution network voltage does not exceed 1.05 per unit. Then at 19:00, the wholesale electricity market price decreases and the BESS discharging behaviors almost stops. As the EV smart charging demand significantly increases at 20:30, the BESS begins to discharge again to ensure the voltage stability of the distribution network. The BESS will increasingly discharge when the voltage of distribution network is relatively low (i.e., 20:00 to 05:00). Comparing Figs. 7 and 11(b), the V_t^{Max} and V_t^{Min} curves can be effectively constrained within the upper and lower boundaries due to the integration of BESS. Fig. 11(c) shows that the BESS's SOC variation and the BESS's SOC can effectively operate within the constraints of 10% and 100%. Fig. 11(d) presents the power curve of P^{net} in Fig. 8. Fig. 11(d) shows that the difference between the maximum value and the minimum value of P^{net} decreases after BESS is added to the network. In Fig. 11(d), the $P_{t,base}^{Net}$ presents the power curve of P^{net} in Fig. 8 without BESS. In Fig. 11(d), the $P_{t,TD3}^{Net}$ presents the power curve of P^{net} in Fig. 8 with BESS and the BESS is scheduled by TD3.

During the training process, the agent determines the charging and discharging policy of the BESS based on the state s_t of the environment as determined by the forecasted datasets, i.e., PV generation and EV charging demand. The reward is an evaluation of the goodness of the policy taken by the agent. The agent will eventually learn to take the optimal policy to maximize the cumulative long-term reward for the task. The best charging and discharging policy in the training process, i.e., training action is shown in Fig. 12. Subsequently, the agent is then applied to a new environment with real-time datasets, i.e., the real PV generation curve and real fast EV charging demand. The agent will take a new policy in real-time according to the observation from the new environment. Fig. 12 shows the charging and discharging policy in the new environment as real-time action. The similar actions between the training and real-time scenarios proved that the trained agent can set up an effective policy for real-time applications according to the observations.

Fig. 13 shows the convergence curve of TD3 algorithm. The episode reward is the total reward for each episode. The average reward is the average reward for the last 100 episodes.

3.3. Comparison with state-of-the-art techniques

To verify the effectiveness of the proposed method, this section presents a comparison of the proposed method with a reinforcement learning technique based on the continuous action space and actor-critic approach i.e., DDPG and heuristic algorithms i.e., PSO and SA.

3.3.1. DDPG

DDPG is a type of reinforcement learning algorithm which is similar to TD3. Their main difference is that DDPG only has one pair of critic network, but TD3 has two critic networks, i.e., selecting the minimum value of two critic approximators for training. Thus, TD3 can reduce the overestimation and increase the training stability and speed. The results of BESS being scheduled by DDPG are as follows: Fig. 14(a) shows that the BESS charges from 9:00 to 15:00 due to surplus PV power generation. After 20:30, BESS starts to discharge because the EV smart charging demand significantly increases. The BESS can manage power flows to ensure that the distribution network operates within the voltage constraints. The BESS's SOC in Fig. 14(c) and the voltage of the distribution network in Fig. 14(b) are within the set range. Comparing Figs. 14(d) and Fig. 11(d), the fluctuation of $P_{t,DDPG}^{Net}$ is greater than that of $P_{t,TD3}^{Net}$.

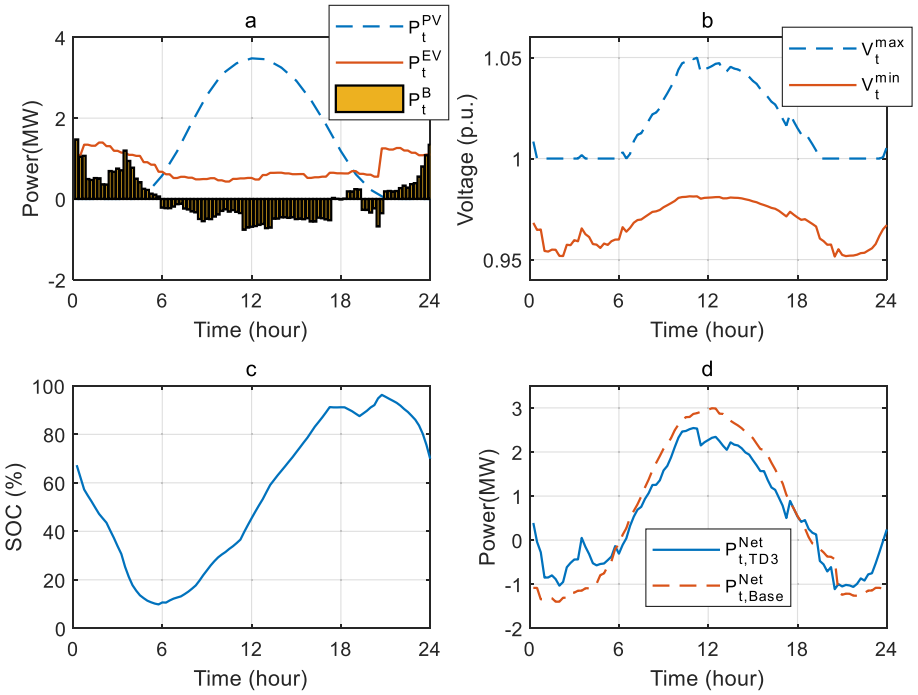

Fig. 11. Power flow, voltage level and SOC of distribution network and BESS (TD3).

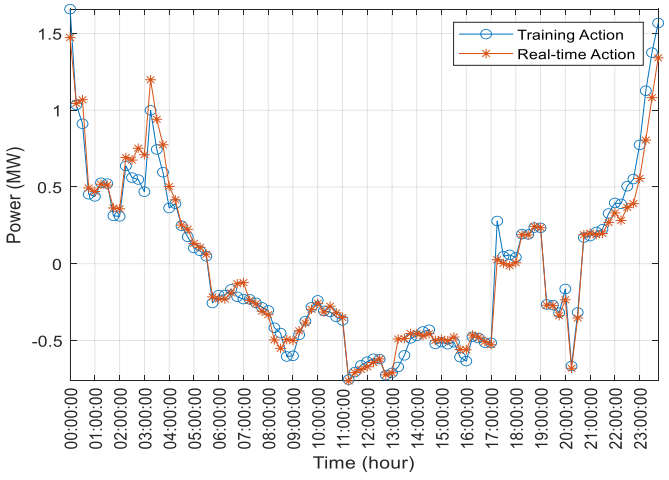

Fig. 12. Training action and real-time action in simulation.

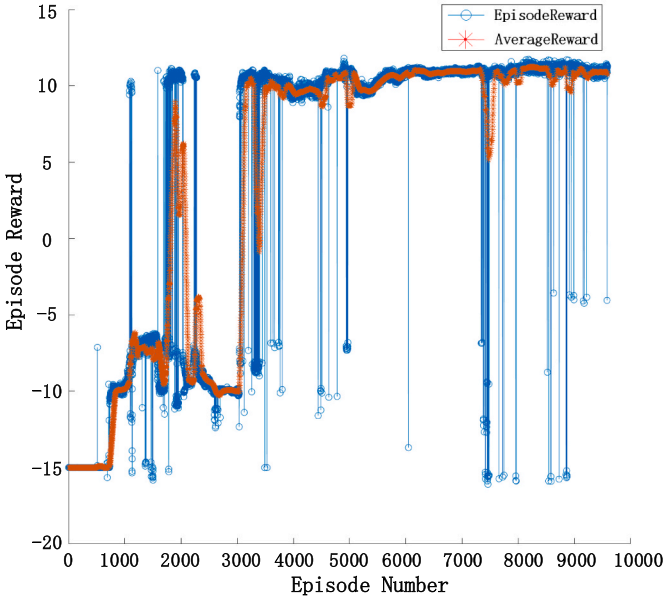

Fig. 13. Reinforcement learning convergence for TD3.

This should be due to the frequent switching of charging and discharging of the BESS. This also leads to voltage fluctuations, and the voltage fluctuations in Fig. 14(b) are significantly larger than those in Fig. 11(b).

3.3.2. PSO

To verify the effectiveness of reinforcement learning algorithms, PSO has been used to solve the same optimization problem. Using the original PSO algorithm, it is difficult to get a solution that satisfies the constraints. Some particles cannot be updated during iterations and stay at the boundaries of the search space. The simulations indicate that it is difficult for PSO to avoid constraints violation at each time step. Therefore, we have made some modifications to the PSO algorithm to improve the optimization performance. When the particles do not update after several iterations and no reasonable strategy can be found, we will save the particle with the highest score. Subsequently, all particles are then initialized and will move towards the particle with the best score. In the process of particle updates, particles with a higher score may appear to replace the previously saved particle. Fig. 15 shows that the BESS scheduling strategy determined by the modified PSO can effectively satisfy the constraint. Although the modified PSO algorithm can obtain a BESS scheduling strategy that satisfies the constraints, Fig. 15(a) shows that the modified PSO scheduling strategy still has the problem of frequent switching between charging and discharging. This will cause a problem as the voltage fluctuation and $P_{t,PSO}^{Net}$ are large as shown in Fig. 15(b) and (d), respectively. Fig. 15(c) shows the modified PSO which cannot fully utilize the BESS capacity for arbitrage. This is because the maximum value of the SOC in Fig. 15(c) only reaches about 50%.

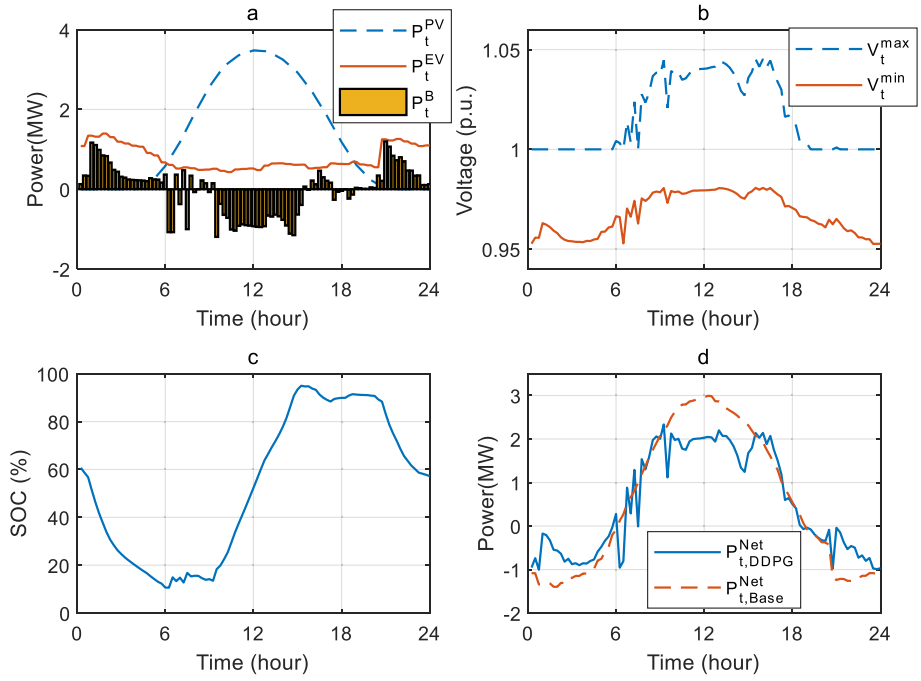

Fig. 14. Power flow, voltage level and SOC of distribution network and BESS (DDPG).

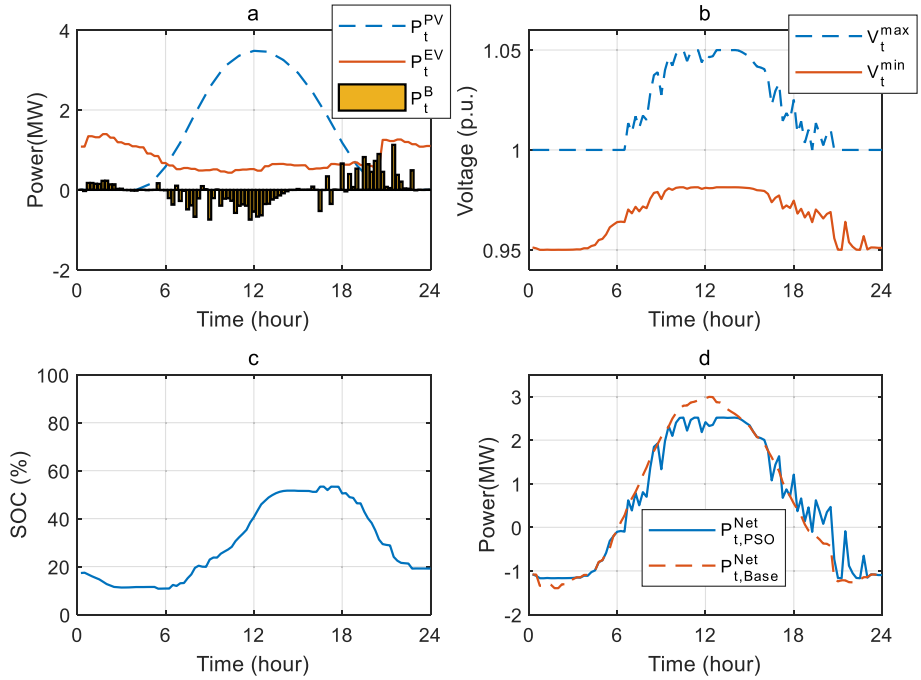

Fig. 15. Power flow, voltage level and SOC of distribution network and BESS (PSO).

3.3.3. SA

The simulated annealing (SA) method is a mature stochastic optimization method. It is a serial optimization algorithm that searches for the global optimum in random, there exists a probability to jump out of the local optimum and eventually converge to the global optimum. The magnitude of the SA random search decreases with the number of iterations, which is the process of simulated annealing. A point with a lower rating than the reference point will be randomly accepted as a new reference point based on the probability when it appears in the iteration. This is the reason why SA can go beyond the local optimum and converge to the global optimum.

The results obtained using SA are shown in Fig. 16. Fig. 16(b) shows that the network voltage can be stabilized within the set range. Fig. 16 (c) shows that the maximum value of the SOC is less than 60%, so the BESS still has a large unused capacity. In Fig. 16 (a), BESS charges at the peak of PV generation and discharges at the peak of EV charging. This is in accordance with the expected scheduling strategy of BESS.

Figs. 11(d), 14(d) and 15(d) and Fig. 16(d) show the power flow of the considered assets as determined from Equation (4). When the distribution network excludes BESS, the power flow P_t^{net} in Equation (4) is expressed as $P_{t,Base}^{Net}$. When the distribution network includes BESS, the power flow P_t^{net} in Equation (4) is expressed as $P_{t,TD3}^{Net}$, $P_{t,DDPG}^{Net}$, $P_{t,PSO}^{Net}$ and

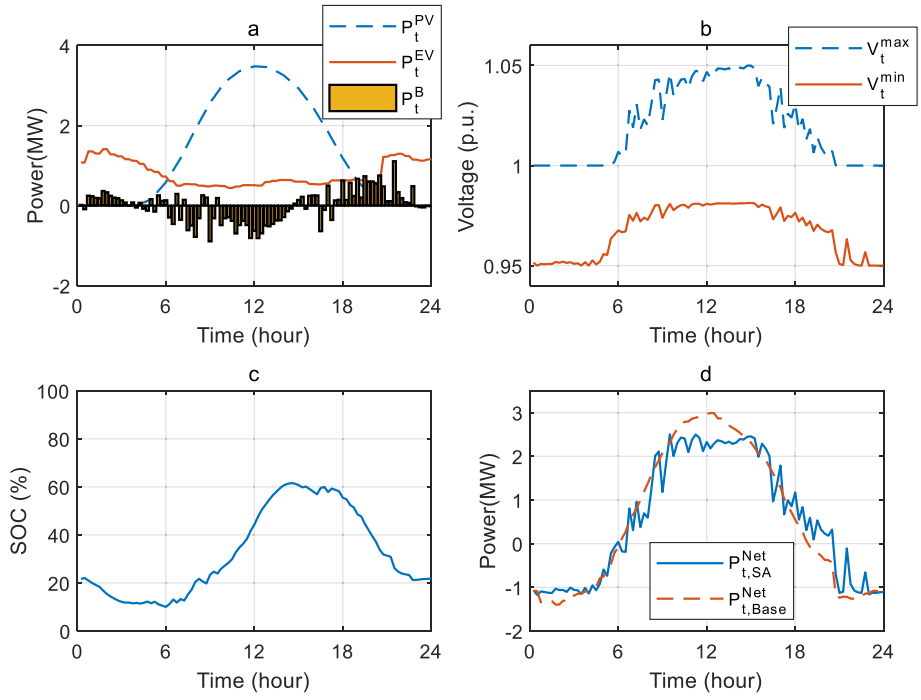

Fig. 16. Power flow, voltage level and SOC of distribution network and BESS (SA).

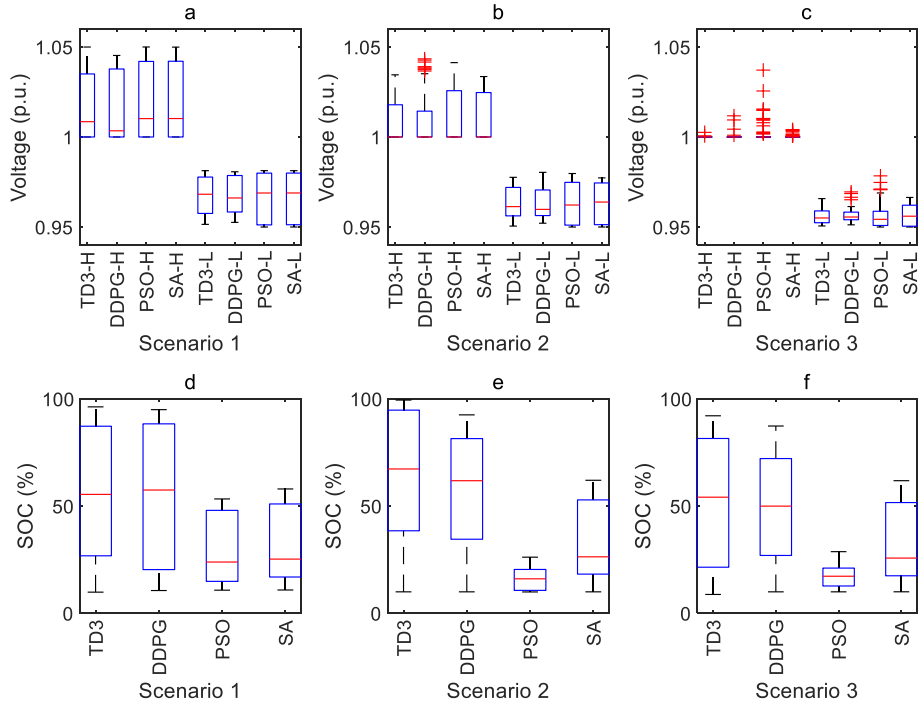

Fig. 17. Comparison of voltage and SOC results of different algorithms in different scenarios.

$P_{t,SA}^{Net}$ for the power flow in BESS charging and discharging strategies with TD3, DDPG, PSO and SA, respectively. The figures show that all the four algorithms can suppress power fluctuations. In contrast, the power flow of the BESS charging and discharging strategy with TD3 is more consistent while the power flow of PSO and SA has a greater fluctuation.

3.4. Different PV power generation scenarios

The results of the four algorithms for Scenario 1 were analyzed in detail in Sections 3.2, 3.3.1, 3.3.2 and 3.3.3. The peak PV generation in Scenario 1 is greater than that in Scenarios 2 and 3, which considers the most extreme cases of grid operation where PV generation may cause overvoltage. Scenario 1 is an important experiment to verify that BESS can guarantee the safe operation of the system. Scenarios 1, 2 and 3 are established according to the amount of PV power generated which can more adequately reflect the working conditions of the system under different weather conditions. Fig. 17 shows the voltage variations and SOC usage of algorithms TD3, DDPG, PSO and SA under Scenarios 1 to 3.

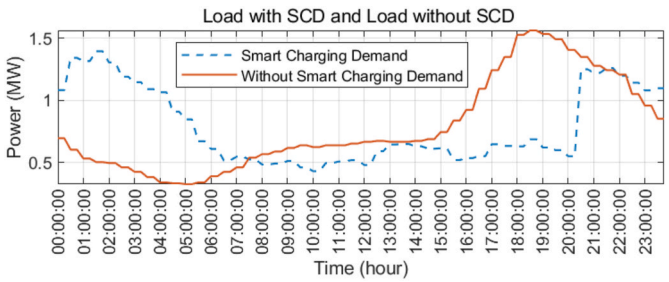

Fig. 18. Load with SCD and without SCD

The TD3-H, DDPG-H, PSO–H and SA-H in Fig. 17(a–c) represent the variation of the distribution network voltage V_t^{Max} when the BESS is scheduled by the TD3, DDPG, PSO and SA algorithms. The TD3-L, DDPG-L, PSO-L and SA-L in Fig. 17(a–c) represent the variation of the distribution network voltage V_t^{Min} when the BESS is scheduled by the TD3, DDPG, PSO and SA algorithms. Fig. 17(a–c) show that when the BESS is scheduled by the TD3 algorithm, the voltage variations of the distribution network is the smallest. Fig. 17(d–f) show that the TD3 algorithm can maximize the use of BESS.

3.5. *Effect on smart charging demand in BESS power scheduling*

This section examines the scheduling strategy of BESS related to the two different fast EV charging demands, i.e., with and without demand

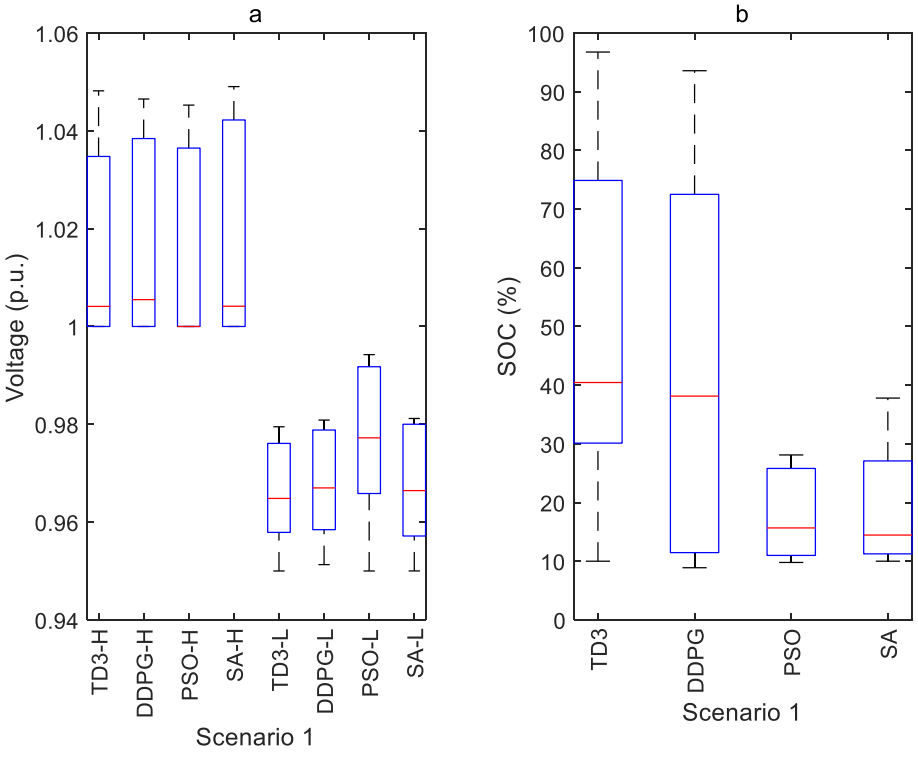

Fig. 19. Comparison of voltage and SOC results of different algorithms without SCD in Scenario 1.

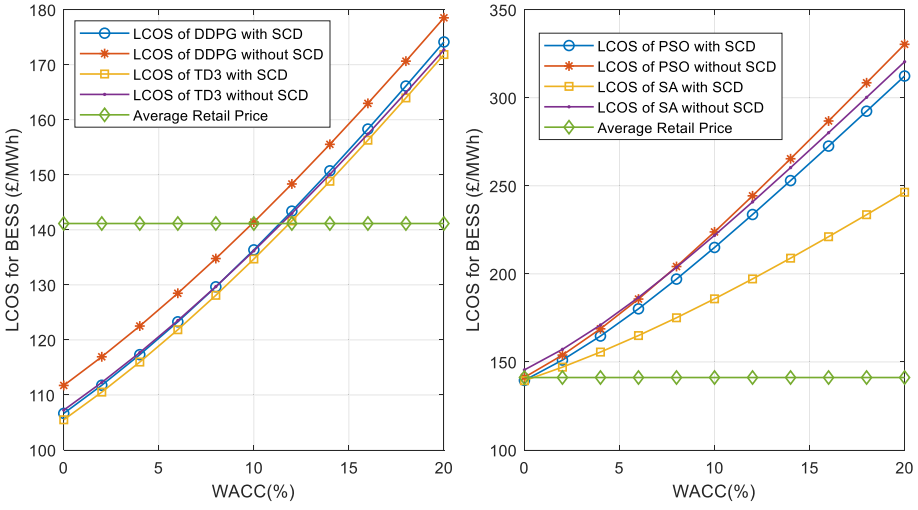

Fig. 20. LCOS for BESS in comparison with and without SCD for DDPG, TD3, PSO and SA.

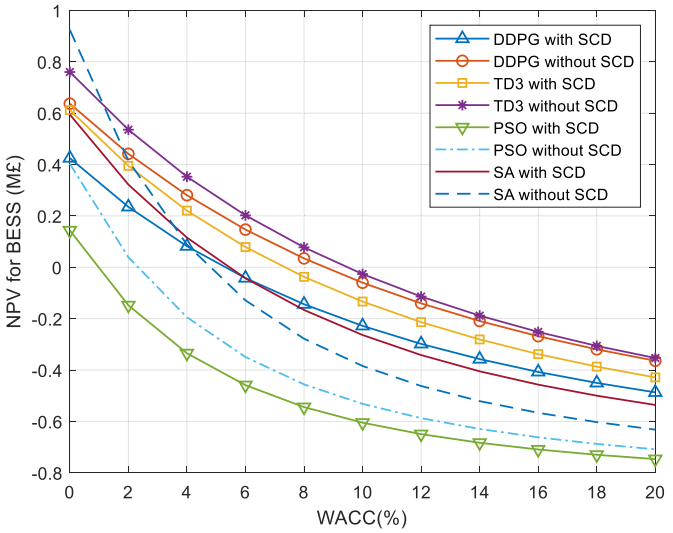

Fig. 21. NPV for BESS in comparison with and without SCD for DDPG, TD3, PSO and SA.

response. Fig. 10 shows the wholesale market price and retail electricity price with and without SCD. In Fig. 10 the retail electricity price with SCD is the electricity price where EV can participate in demand response. The fast EV charging demand considering price incentive i.e., SCD and without SCD are shown in Fig. 18. The total daily energy of fast EV charging with and without SCD are the same. This is to ensure that the BESS scheduling strategies under these two EV loads could be compared.

The results in Fig. 19(a) show that the BESS scheduled by the TD3 algorithm has the smallest voltage variation of the distribution network. Fig. 19(b) shows that the BESS scheduled by TD3 can maximize the utilization of BESS. Combining Sections 3.4 and 3.5, it can be concluded that the use of TD3 for BESS can minimize voltage variation of distribution networks. TD3 has a high utilization of BESS, so that BESS can be used to arbitrage in the power market as much as possible.

3.6. Economic and financial analysis

This section presents the economic and financial analysis for BESS in a distribution network considering with and without EV demand response i.e., smart charging. WACC is a key input for financial models. The WACC is considered as an overall combined effect of the cost of debt, cost of equity, share of CAPEX, and the corporate tax rate [27]. This work examines a range of WACC between 0% and 20%. Due to the importance and uncertainty in WACC, Fig. 20 and Fig. 21 show the LCOS and NPV for the BESS, respectively. Fig. 20 shows the LCOS and average retail electricity price.

In this study, the fixed operation and maintenance (O&M) cost of BESS is set to 2920 £/MW-yr [34] and the average retail electricity price is 141.15 £/MWh. Fig. 20 shows that the BESS scheduled by the TD3 algorithm is slightly better than the BESS scheduled by the DDPG algorithm in the evaluation of LCOS. Fig. 20 also shows that the BESS scheduled by TD3 is significantly better than the BESS scheduled by PSO or SA in the evaluation of LCOS. To make sense of numbers, the calculated LCOS values are compared with the work in Ref. [35]. Ref. [36] claimed that the LCOS of battery technologies can achieve 175 £/MWh in the future. Ref. [22] claimed that the LCOS can be reduced if the system operates in "High-PV" (scenario with high penetration of photovoltaic power) scenario. "High-PV" contributes to a reduced LCOS due to higher lifetime energy output. The WACC required for the LCOS to be greater than the retail electricity price is 10% (High-PV) [22]. In this study, for RL algorithms the LCOS was below 175 £/MWh except for "DDPG without SCD". For stochastic optimization algorithms, the LCOS was above 175 £/MWh. For RL algorithms, the range of WACC required for the LCOS to be greater than the retail electricity price is between 10% (LCOS of DDPG without SCD) to 12% (LCOS of TD3 with SCD). But for PSO and SA, the range of WACC required for the LCOS to be greater than the retail electricity price is less than 2%. This is an important indicator as the BESS can be economic if the cost of electricity per MWh (LCOS) is less than the revenue generated per MWh (average retail electricity price).

Ref. [22] studied the IRR distribution of BESS from 2.13% to 9.55%. The BESS in Ref. [22] is applied in systems with a large amount of PV i.e., solar PV energy output 1.53 GWh/year and anaerobic digestion biogas power generation. Fig. 21 shows the NPV for BESS and there is a quadratic relationship between the NPV and WACC. The NPV at the current WACC is greater than zero indicating that BESS is profitable. The discount rate which provides an NPV equals to zero is known as the IRR. The IRRs of the strategies using TD3, DDPG, PSO and SA with SCD are 7.31%, 5.27%, 0.86% and 5.39%, respectively. The IRRs of the strategies using TD3, DDPG, PSO and SA without SCD are 9.46%, 8.69%, 2.27% and 4.17%, respectively. The results show that BESS can be more profitable when EVs are not involved in demand response and BESS can get a better return with TD3.

The daily costs and profit of BESS with SCD under different scenarios and optimization algorithms (as examined in Sections 3.3 and 3.4) are given in Table 3. Table 4 presents the daily costs and profit of BESS without SCD and different optimization algorithms are examined in Section 3.5. Considering Tables 3 and 4 when comparing Scenario 1, the economic performance of BESS without SCD is better than that with SCD. This is due to SCD will reduce the peak charging demand in the evening, which could have been supported by BESS to make additional

Table 3
Daily costs and profit of BESS with SCD for the three scenarios.

Algorithm	Scenario	Energy cost (£/day)	Battery degradation cost (£/day)	O&M cost (£/day)	Profit (£/day)
TD3	Scenario 1	384.49	174.06	42.00	132.91
	Scenario 2	369.68	165.85	42.00	132.49
	Scenario 3	395.26	172.27	42.00	159.88
DDPG	Scenario 1	346.73	148.85	42.00	102.26
	Scenario 2	364.68	155.85	42.00	132.49
	Scenario 3	419.54	186.21	42.00	65.41
PSO	Scenario 1	237.99	100.89	42.00	57.74
	Scenario 2	189.36	85.38	42.00	52.52
	Scenario 3	74.99	32.41	42.00	6.08
SA	Scenario 1	201.85	105.64	42.00	72.64
	Scenario 2	255.09	115.55	42.00	103.42
	Scenario 3	251.08	114.86	42.00	114.53

Table 4
Daily costs and profit of BESS without SCD for Scenario 1.

Algorithms	Energy cost (£/day)	Battery degradation cost (£/day)	O&M cost (£/day)	Profit (£/day)
TD3	378.61	174.49	42.00	183.23
DDPG	382.60	176.46	42.00	167.83
PSO	131.75	59.11	42.00	81.41
SA	160.14	90.36	42.00	84.27

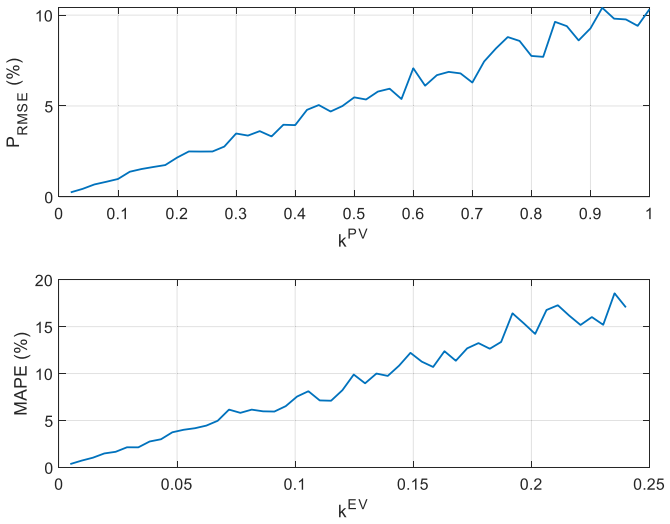

Fig. 22. Correspondence of k^{EV} and k^{PV} with forecast error parameter.

profit.

3.7. Case study of generalization capability and sensitivity analysis

Equations (6) and (7) are used to generate 50 SCDs and 50 PV generation curves, respectively. 50 k^{PV} values are generated which are evenly distributed between zero and one. Then using Equation 6 and (50) PV generation curves are generated and the forecast error parameter P_{RMSE} of PV generation curves is distributed between 0% and 10.4%. P_{RMSE} is defined as the forecasted root mean square error (RMSE) of PV generation divided by the generation capacity of PV. 50 k^{EV} values are generated which are evenly distributed between zero and 0.24, to produce a mean absolute percentage error (MAPE) of up to 20%. Subsequently, 50 SCDs are generated using Equation (7) and the MAPE of SCD can reach up to 18.55%. The correspondence of k^{EV} and k^{PV} with forecast error parameters are shown in Fig. 22. The combinations of 50 SCDs and 50 PV generation curves produce 2500 scenarios. For the RL algorithms, the charging and discharging strategy of BESS is conducted in real-time. The power P_t^{back} (i.e., the reserve power) is calculated by the flowchart of real-time optimization method in Fig. 2, and subsequently, S^{back} is calculated by Equation (8) as the reserve power usage for the scenario. For PSO and SA, the day-ahead charging and discharging strategy of the BESS is used to calculate P_t^{back} by Fig. 2, and then S^{back} is obtained by the steps of the RL algorithms described above.

Fig. 23 presents the reserve energy usage for different optimization methods. The experimental results show that TD3 has minimal use of reserve power. For the TD3 algorithm, no reserve power is required when the PV forecast error parameter P_{RMSE} is less than 9% and the SCD forecast MAPE is less than 5%. In addition, TD3 algorithm can be used in most scenarios without the use of reserve power. The DDPG algorithm does not require reserve power when the forecast error is small, while the demand for reserve energy increases when the forecast error becomes large. For SA and PSO, all scenarios require reserve energy. SA and PSO algorithms are more sensitive to the forecast error of SCD. The maximum, minimum and average values of the capacity of reserve energy required by different algorithms for 2500 scenarios are shown in Table 5. Compared to the other mentioned algorithms, the TD3 algorithm has a smaller maximum, minimum and average value of S^{back}. In general, the TD3 algorithm has a strong generalization capability and requires only a small amount of reserve energy to operate properly even when the forecast error is large. In Equation (8), the worst case is when $P_t^{back} \geq 0, \forall t$ or $P_t^{back} \leq 0, \forall t$. In the worst case, it is necessary to prepare the minimum reserve capacity of twice the S^{back} and the SOC of the reserve energy storage is 50%. In this case, TD3 requires a minimum reserve energy storage capacity of 0.492 MWh.

4. Conclusions

This paper presents a novel optimal power scheduling methodology

Table 5
Capacity of reserve energy required for different algorithms.

S^{back}	TD3	DDPG	PSO	SA
Average (MWh)	0.027	0.124	0.168	0.071
Maximum (MWh)	0.246	0.775	0.488	0.256
Minimum (MWh)	0.000	0.000	0.015	0.008

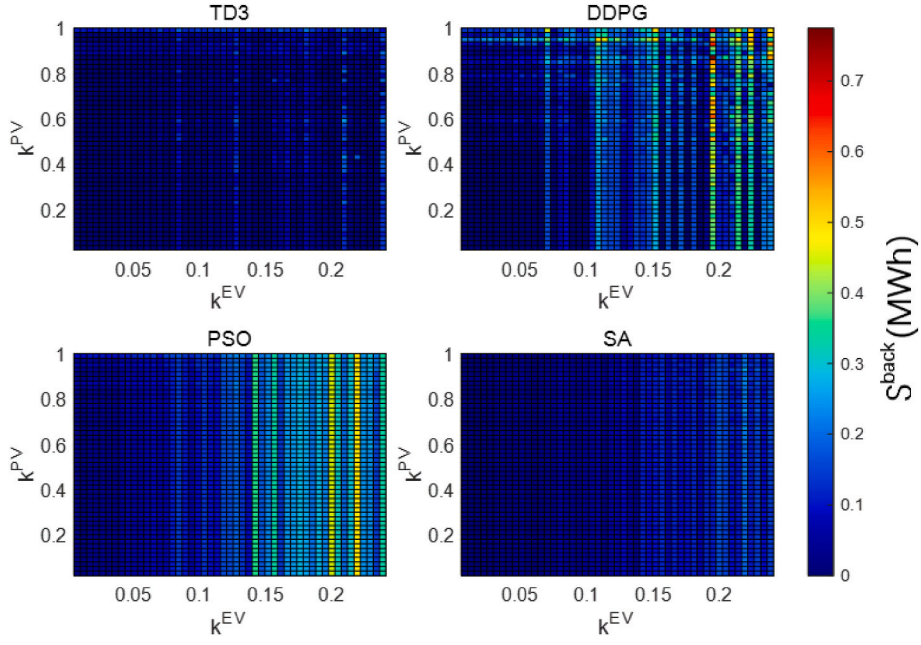

Fig. 23. Reserve energy usage for different optimization methods.

for the BESS in a distribution network with the fast EV charging demand and high penetration of PV power generation. BESS can support fast charging stations and better utilize renewable energy.

ELM and Monte Carlo algorithms were used to forecast PV power generation dataset and to sample the fast EV charging demand dataset for RL training. Case studies are conducted with Smart EV charging dataset from UK Power Networks, Project Shift. The cost of battery degradation is calculated by employing a rainflow cycle counting algorithm and a state-of-the-art battery degradation model. For such a complex model, it is a common practice to simplify and then linearize the model and use a commercial solver to solve it. Alternatively, the problem can be solved using a stochastic optimization algorithm i.e., PSO and SA, and a reserve power supply can be introduced in the real-time phase to ensure the normal operation of the system. The main contribution of this paper is the effective exploitation of the offline training and online implementation of reinforcement learning algorithms. Not only it is fast in online implementation, but it also ensures that the model is not simplified i.e., linearize, etc. The proposed model is solved using the state-of-the-art reinforcement learning algorithm TD3. BESS with additional energy reserve is considered to address the uncertainty of renewable energy and EV charging during real-time scheduling. The results show that the IRRs of the strategies using TD3 and DDPG algorithms with SCD are 7.31% and 5.27%, respectively. The IRRs of the strategies using TD3 and DDPG algorithms without SCD are 9.46% and 8.69%, respectively. The IRRs of the strategies using PSO and SA algorithms with SCD are 0.86% and 5.39%, respectively. The IRRs of the strategies with PSO and SA algorithms without SCD are 2.27% and 4.17%, respectively. A sensitivity analysis of the BESS real-time scheduling is conducted, and the results show that the TD3 algorithm uses the least amount of reserve energy in all scenarios.

Author statement

Chun Sing Lai: Conceptualization, Methodology, Investigation, Validation, Writing – original draft, Writing- Reviewing and editing, Resources, Supervision, Project administration, Funding acquisition; **Dashen Chen**: Methodology, Software, Investigation, Data curation, Formal analysis, Writing – original draft, Writing- Reviewing and editing; **Jinning Zhang**: Writing – original draft; Xin Zhang: Writing – original draft; **Xu Xu**: Conceptualization, Writing- Reviewing and editing; **Gareth A Taylor**: Conceptualization, Methodology, Writing – original draft, Supervision, Funding acquisition; **Loi Lei Lai**: Writing-Reviewing and editing, Resources, Project administration, Funding acquisition, Supervision

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

This study is a re-analysis of existing data which are openly available at locations cited in the References section of this paper.

Acknowledgements

This work is sponsored by Brunel University London BRIEF Funding; the Department of Finance and Education of Guangdong Province 2016 [202]: Key Discipline Construction Program, China; Education Department of Guangdong Province: New and Integrated Energy System Theory and Technology Research Group [Project Number 2016KCXTD022]. We are grateful for support from the DTE Network + funded by EPSRC grant reference EP/S032053/1.

Appendix A

In this section, data acquisition and processing of PV data, wholesale electricity market data, retail electricity price data, and EV smart charging demand data are discussed.

PV Data

A comparison of different irradiance data sources including National Aeronautics and Space Administration (NASA) [37] and historical solar insolation data in the UK from 2017 to 2019 [38] is depicted in Fig. A1. As the UK is situated in the northern hemisphere, the solar insolation is the highest in June and July, and it is the lowest in January and December.

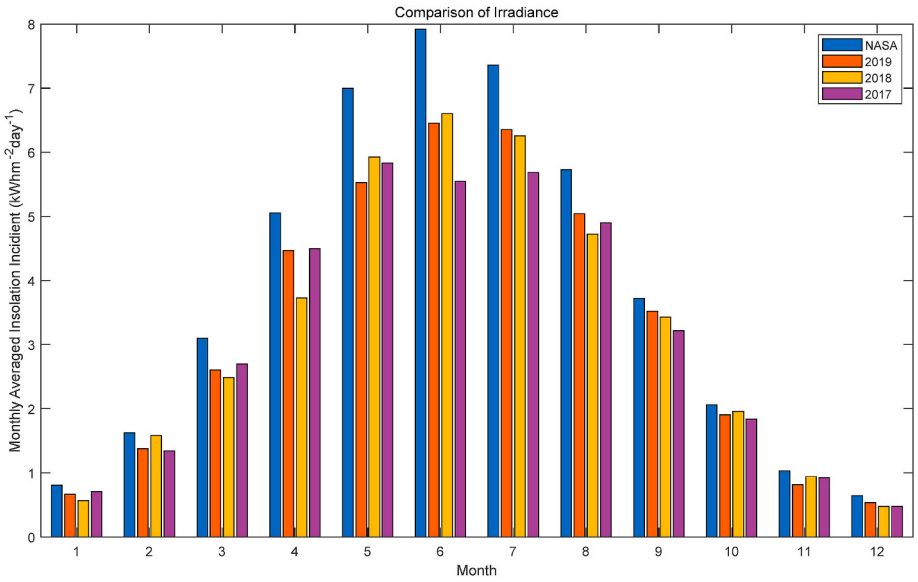

Fig. A.1. Comparison of solar insolation data.

Wholesale electricity market data

Electricity is a commodity traded in the wholesale market. The wholesale price increases with the increased electricity demand. Nord Pool AS is a European power exchange which is responsible for delivering power traded across Europe [35,39]. The wholesale electricity can be used for charging BESS, as well as BESS energy can be sold to the wholesale market. The wholesale electricity can be used to fulfil electricity demand from charging stations directly.

Retail electricity price

The retail electricity price refers to the price that consumers pay for the fast EV charging as described by the Project Shift project contributed by UK Power Networks. The price can be different for each timeslot which incentivizes consumers to charge their EV at low prices. The project has subsequently investigated a tariff, i.e., Octopus Go Faster [40], which provides consumers with inexpensive charging windows at different timeslots of the day. Specifically, the consumers can select both a duration of charging timeslots (e.g., 3-h, 4-h and 5-h period) with reduced electricity rate and a start time. The start time of 20:30, 21:30, 22:30, 23:30, 00:30, 01:30, 02:30 and 03:30 with a tariff choice of 3-h, 4-h and 5-h period is at a reduced electricity rate of 45 £/MWh, 50 £/MWh and 55 £/MWh respectively. All possible options for charging schemes are shown in Table A1 [40]. In the other period, the retail electricity price is at the level of 158.89 £/MWh. Considering that the electric vehicle load curve includes the charging loads of multiple electric vehicles, the charging schemes they choose are also varied. Assuming that the probability of each charging scheme being selected is the same, the charging price of all electric vehicles can be better described by the average retail electricity price which is shown in Fig. A2. The average retail electricity price is calculated by taking the average of all the 3-h, 4-h and 5-h period retail electricity price curves in Fig. A2.

Table A.1
All possible charging schemes options [40].

Time	Price	Time	Price	Time	Price
20:30–23:30	45 £/MWh	21:30–00:30	45 £/MWh	22:30–01:30	45 £/MWh
23:30–02:30	45 £/MWh	00:30–03:30	45 £/MWh	01:30–04:30	45 £/MWh
02:30–05:30	45 £/MWh	03:30–06:30	45 £/MWh	20:30–00:30	50 £/MWh
21:30–01:30	50 £/MWh	22:30–02:30	50 £/MWh	23:30–03:30	50 £/MWh
00:30–04:30	50 £/MWh	01:30–05:30	50 £/MWh	02:30–06:30	50 £/MWh
20:30–01:30	55 £/MWh	21:30–02:30	55 £/MWh	22:30–03:30	55 £/MWh
23:30–04:30	55 £/MWh	00:30–05:30	55 £/MWh	01:30–06:30	55 £/MWh

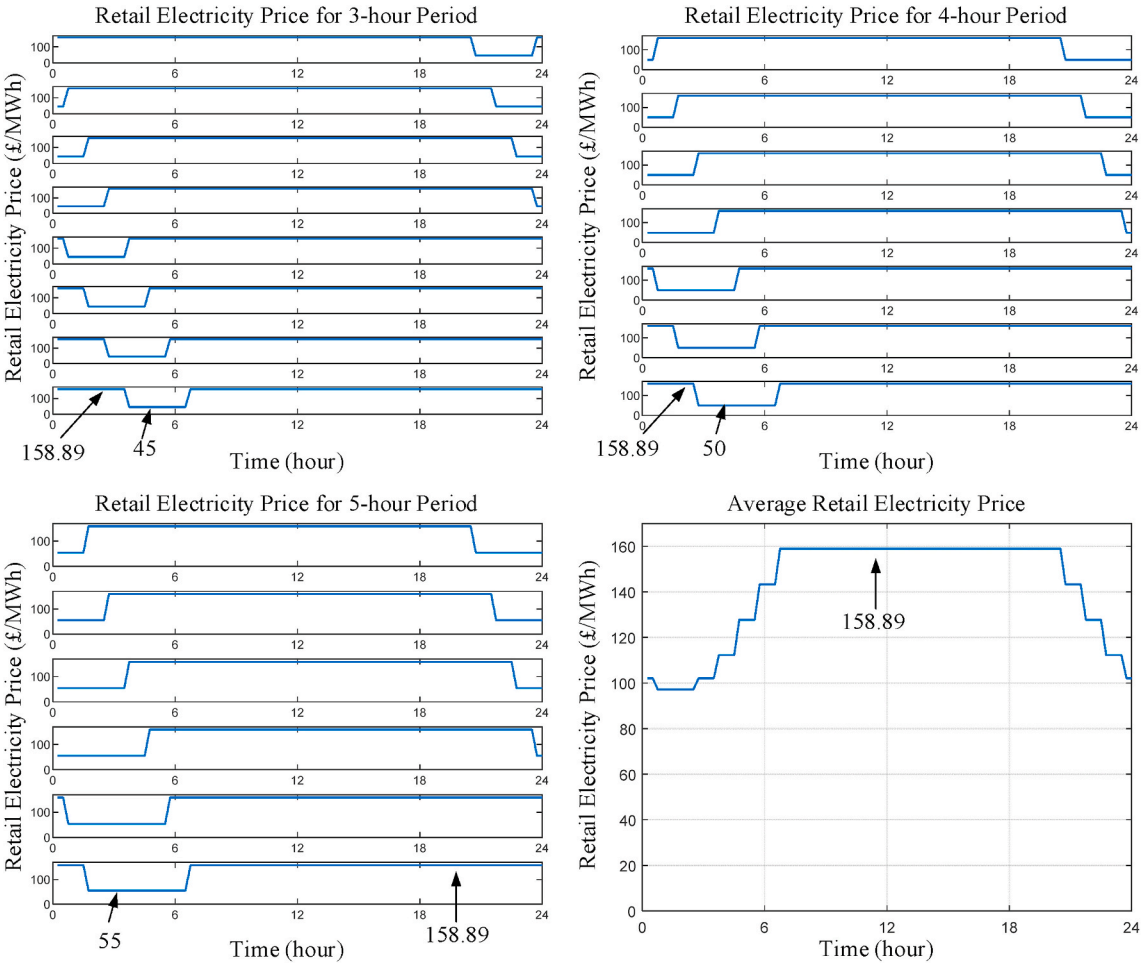

Fig. A.2. Retail electricity price for smart charging demand.

Smart charging demand data

The smart charging demand data is obtained from the Project Shift project contributed by UK Power Networks. Monte Carlo simulations were performed to generate synthetic smart charging demand data as presented in Fig. A3 [1].

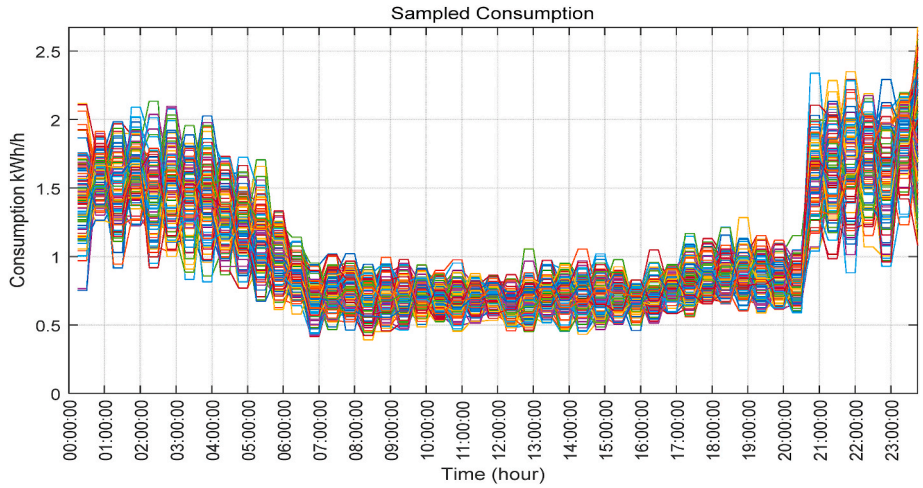

Fig. A.3. Charging profiles of maximum electricity consumption for 20 customers from the Octopus Energy trial (100 samples) [1].

Appendix B

Solar PV power forecasting model

Extreme Learning Machine (ELM) has been widely used for solar energy forecasting [41]. Since it is not the focus of this work, the principle of ELM is briefly described here. Fig. B1 shows that the ELM is a feed-forward neural network whose input weights a_i and biases b_i are fixed, and the error function is reduced by updating the output weights β_i. Because a_i and b_i are fixed, the output of L hidden neurons is also fixed and can be represented by the H matrix. The output weights β_i are obtained by calculating $\min_{\beta \epsilon R_{L\times 1}} H\beta - T^2$, where T is the expected output of the training set. ELM models can achieve extremely fast learning speed and good generalization capability.

ELM is used to forecast photovoltaic power generation [41]. In the PV power forecasting model, the solar radiation intensity and cloud cover conditions are the inputs of ELM, and the output is the PV power generation value. The forecast of PV power generation will be used in the training process of reinforcement learning based optimal power scheduling strategy.

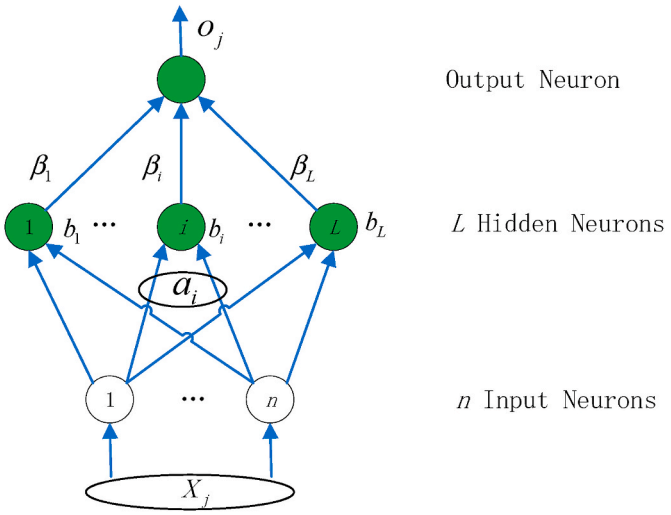

Fig. B.1. ELM network model.

Forecasting data for agent training

The datasets required to train a RL agent are described in this section. Fig. B2 shows the PV power generation forecast using ELM. Monte Carlo method is used to generate the EV smart charging demand data. The mean absolute percentage error (MAPE) of EV smart charging demand forecast and PV power generation forecast are 4.30% and 12.69%, respectively [1]. The MAPE of EV smart charging demand forecasted by the method proposed in Ref. [42] is 2.62%, 2.90%, 3.25% and 3.78% in spring, summer, autumn, and winter, respectively. In Ref. [43], the MAPE of PV power generation forecast is 7.65% and can be as high as 10.00% at noon. Compared with other studies, the MAPE in this work is relatively high. Even with a lower forecast accuracy data, the trained agent can achieve a better power scheduling strategy, which also shows that the generalization capability of the algorithm can meet the application requirements in a real environment.

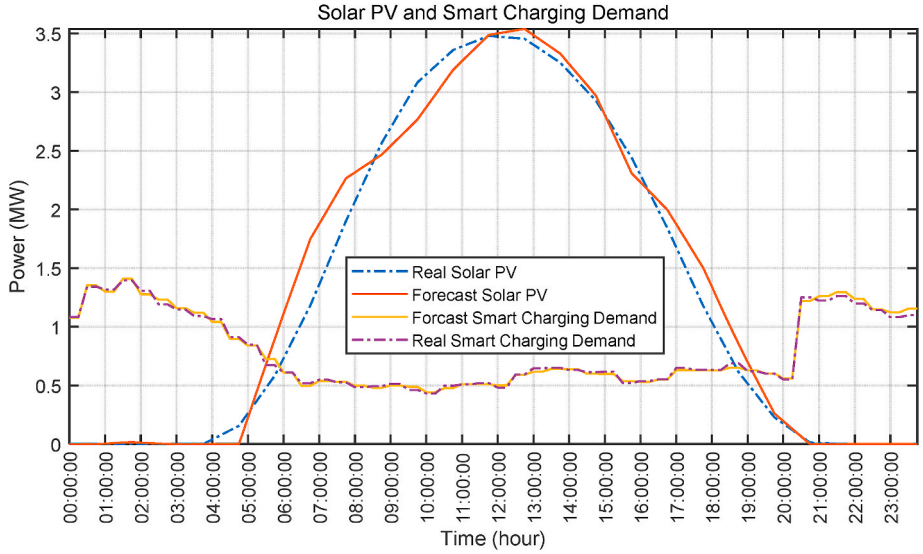

Fig. B.2. The forecasted and real PV power generation and EV smart charging demand.

References

[1] Trialling market based incentives for domestic smart electric vehicle charging. 2021. https://innovation.ukpowernetworks.co.uk/wp-content/uploads/2021/02/UKPN_Shift_Interim_Report_v05.pdf.

[2] Promoting vehicle efficiency and electrification through stimulus packages. 2020. https://www.iea.org/articles/promoting-vehicle-efficiency-and-electrification-through-stimulus-packages.

[3] Low-emission vehicles eligible for a plug-in grant. 2020. https://www.gov.uk/plug-in-car-van-grants.

[4] Almutairi A. Plug-in electric vehicles and their impact on power generation availability: a real survey-based analysis in Saudi Arabia. Sustain Cities Soc 2021;75(Dec). https://doi.org/10.1016/j.scs.2021.103389.

[5] Li C, Zhang L, Ou Z, Wang Q, Zhou D, Ma J. Robust model of electric vehicle charging station location considering renewable energy and storage equipment. Energy 2022;238. https://doi.org/10.1016/j.energy.2021.121713. Jan.

[6] Zeynali S, Rostami N, Ahmadian A, Elkamel A. Stochastic energy management of an electricity retailer with a novel plug-in electric vehicle-based demand response program and energy storage system: a linearized battery degradation cost model. Sustain Cities Soc 2021;74. https://doi.org/10.1016/j.scs.2021.103154. Nov.

[7] Cao J, Harrold D, Fan Z, Morstyn T, Healey D, Li K. Deep reinforcement learning-based energy storage arbitrage with accurate lithium-ion battery degradation model. IEEE Trans Smart Grid 2020;11(5):4513–4521, Sep. https://doi.org/10.1109/TSG.2020.2986333.

[8] Harrold DJB, Cao J, Fan Z. Data-driven battery operation for energy arbitrage using rainbow deep reinforcement learning. Energy Jan. 2022;238. https://doi.org/10.1016/j.energy.2021.121958.

[9] Huang B, Wang J. Deep-reinforcement-learning-based capacity scheduling for PV-battery storage system. IEEE Trans Smart Grid 2021;12(3):2272–83. https://doi.org/10.1109/TSG.2020.3047890. May.

[10] Gao Y, Yang J, Yang M, Li Z. Deep reinforcement learning based optimal schedule for a battery swapping station considering uncertainties. IEEE Trans Ind Appl 2020;56(5):5775–84. https://doi.org/10.1109/TIA.2020.2986412. Sep.

[11] Wan Z, He HLH, Prokhorov D. Model-free real-time EV charging scheduling based on deep reinforcement learning. IEEE Trans Smart Grid 2018. https://doi.org/10.1109/TSG.2018.2879572.

[12] Qureshi KN, Alhudhaif A, Jeon G. Electric-vehicle energy management and charging scheduling system in sustainable cities and society. Sustain Cities Soc 2021;71. https://doi.org/10.1016/j.scs.2021.102990. Aug.

[13] Wu X, Feng Q, Bai C, Lai CS, Jia Y, Lai LL. A novel fast-charging stations locational planning model for electric bus transit system. Energy 2021;224. https://doi.org/10.1016/j.energy.2021.120106. Jun.

[14] Liu Z, Wu Q, Shahidehpour M, Li C, Huang S, Wei W. Transactive real-time electric vehicle charging management for commercial buildings with PV on-site generation. IEEE Trans Smart Grid 2019;10(5):4939–50. https://doi.org/10.1109/TSG.2018.2871171.

[15] Jin J, Xu Y. Optimal policy characterization enhanced actor-critic approach for electric vehicle charging scheduling in a power distribution network. IEEE Trans Smart Grid 2021;12(2):1416–28. https://doi.org/10.1109/TSG.2020.3028470. Mar.

[16] Luo Y, Feng G, Wan S, Zhang S, Li V, Kong W. Charging scheduling strategy for different electric vehicles with optimization for convenience of drivers, performance of transport system and distribution network. Energy 2020;194. https://doi.org/10.1016/j.energy.2019.116807. Mar.

[17] Rajani DB, Kommula BN. An optimal energy management among the electric vehicle charging stations and electricity distribution system using GPC-RERNN approach. Energy 2022;245. https://doi.org/10.1016/j.energy.2022.123180. Apr.

[18] Kang Q, Wang J, Zhou M, Ammari AC. Centralized charging strategy and scheduling algorithm for electric vehicles under a battery swapping scenario. IEEE Trans Intell Transport Syst 2016;17(3):659–69. https://doi.org/10.1109/TITS.2015.2487323. Mar.

[19] Li N, Uckun C, Constantinescu EM, Birge JR, Hedman KW, Botterud A. Flexible operation of batteries in power system scheduling with renewable energy. IEEE Trans Sustain Energy 2016;7(2):685–96. https://doi.org/10.1109/TSTE.2015.2497470. Apr.

[20] Krishnamurthy D, Uckun C, Zhou Z, Thimmapuram PR, Botterud A. Energy storage arbitrage under day-ahead and real-time price uncertainty. IEEE Trans Power Syst 2017;33(1):84–93. https://doi.org/10.1109/tpwrs.2017.2685347. Apr.

[21] Saxena S, Hendricks C, Pecht M. Cycle life testing and modeling of graphite/LiCoO2 cells under different state of charge ranges. J Power Sources 2016;327:394–400. https://doi.org/10.1016/j.jpowsour.2016.07.057. Sep.

[22] Lai CS, Locatelli G, Pimm A, Tao Y, Li X, Lai LL. A financial model for lithium-ion storage in a photovoltaic and biogas energy system. Appl Energy 2019;251. https://doi.org/10.1016/j.apenergy.2019.04.175. Oct.

[23] Lai CS, et al. Levelized cost of electricity for photovoltaic/biogas power plant hybrid system with electrical energy storage degradation costs. Energy Convers Manag 2017;153:34–47. https://doi.org/10.1016/j.enconman.2017.09.076. Dec.

[24] Saxena S, Hendricks C, Pecht M. Cycle life testing and modeling of graphite/LiCoO2 cells under different state of charge ranges. J Power Sources 2016;327:394–400. https://doi.org/10.1016/j.jpowsour.2016.07.057. Sep.

[25] Liu W, Xu Y, Feng X, Wang Y. Optimal fuzzy logic control of energy storage systems for V/f support in distribution networks considering battery degradation. Int J Electr Power Energy Syst 2022;139. https://doi.org/10.1016/j.ijepes.2021.107867. Jul.

[26] Fujimoto S, van Hoof H, Meger D. Addressing function approximation error in actor-critic methods. 2018 [Online]. Available: https://github.com/.

[27] Lai CS, Locatelli G. Economic and financial appraisal of novel large-scale energy storage technologies. Energy 2021;214. https://doi.org/10.1016/j.energy.2020.118954. Jan.

[28] Lai CS, McCulloch MD. Levelized cost of electricity for solar photovoltaic and electrical energy storage. Appl Energy 2017;190:191–203. https://doi.org/10.1016/j.apenergy.2016.12.153.

[29] Baghzouz Y, Ertem S. Shunt capacitor sizing for radial distribution feeders with distorted substation voltages. IEEE Trans Power Deliv 1990;5(2):650–7. https://doi.org/10.1109/61.53067.

[30] Nykvist B, Nilsson M. Rapidly falling costs of battery packs for electric vehicles. Nat Clim Change 2015;5(4):329–32. https://doi.org/10.1038/nclimate2564. Mar.

[31] Smarter network storage (SNS). 2019. https://innovation.ukpowernetworks.co.uk/projects/smarter-network-storage-sns/.

[32] Smarter network storage low carbon network Fund. 2019. https://innovation.ukpowernetworks.co.uk/wp-content/uploads/2019/06/SNS1.12-EnergyStorageAsAnAsset-Learning-Report-v1.0.pdf.

[33] Newsroom Edie. Gridserve buys UK's first subsidy-free solar farm to power EV forecourts. Aug. 10. EDIE; 2020. Aug. 10, 2020), https://www.edie.net/gridserve-buys-uks-first-subsidy-free-solar-farm-to-power-ev-forecourts/.

[34] Lazard's levelized cost of storage analysis – version 4.0. Lazard; 2018. November, https://www.lazard.com/media/450774/lazards-levelized-cost-of-storage-version-40-vfinal.pdf.

[35] Lai CS, Locatelli G. Are energy policies for supporting low-carbon power generation killing energy storage? J Clean Prod 2021;280. https://doi.org/10.1016/j.jclepro.2020.124626. Jan.

[36] Jülch V. Comparison of electricity storage options using levelized cost of storage (LCOS) method. Appl Energy 2016;183:1594–606. https://doi.org/10.1016/j.apenergy.2016.08.165. Dec.

[37] National Aeronautics and space administration Goddard Institute for space studies. 2019. https://data.giss.nasa.gov/modelE/ar5plots/srlocat.html.

[38] Solar PV. 2019. www.renewables.ninja.

[39] N2EX Day ahead Auction prices. 2021. https://www.nordpoolgroup.com/Market-data1/GB/Auction-prices/UK/Hourly/?view=table.

[40] Go and go faster. https://www.guylipman.com/octopus/formulas.html.

[41] Fu W, Wang K, Li C, Tan J. Multi-step short-term wind speed forecasting approach based on multi-scale dominant ingredient chaotic analysis, improved hybrid GWO-SCA optimization and ELM. Energy Convers Manag 2019;187:356–77. https://doi.org/10.1016/j.enconman.2019.02.086. May.

[42] Zhang X, Chan KW, Li H, Wang H, Qiu J, Wang G. Deep-learning-based probabilistic forecasting of electric vehicle charging load with a novel queuing model. IEEE Trans Cybern 2021;51(6):3157–70. https://doi.org/10.1109/TCYB.2020.2975134. Jun.

[43] Liu J, Fang W, Zhang X, Yang C. An improved photovoltaic power forecasting model with the assistance of aerosol index data. IEEE Trans Sustain Energy 2015;6(2):434–42. https://doi.org/10.1109/TSTE.2014.2381224. Apr.