DOI:
10.1109/TII.2022.3163778

Publication date:
2022

Document Version
Peer reviewed version

Online Battery Protective Energy Management for Energy-Transportation Nexus

Shuangqi Li, Student Member, IEEE, Pengfei Zhao, *Member, IEEE*, Chenghong Gu, *Member, IEEE*, Jianwei Li, Shuang Cheng, Student Member, IEEE, and Minghao Xu, Member, IEEE

***Abstract*—Grid-connected electric vehicles (GEVs) and Energy-Transportation Nexus bring a bright prospect to improve the penetration of renewable energy and the economy of microgrids. However, it is challenging to determine optimal vehicle-to-grid (V2G) strategies due to the complex battery aging mechanism and volatile microgrid states. This paper develops a novel online battery anti-aging energy management method for Energy-Transportation Nexus by using a novel deep reinforcement learning framework. Based on battery aging characteristic analysis and rain-flow cycle counting technology, the quantification of aging cost in V2G strategies is realized by modelling the impact of number of cycles, depth of discharge, and charge and discharge rate. The established life loss model is used to evaluate battery anti-aging effectiveness of agent actions. The coordination of GEVs charging is modelled as multi-objective learning by using a deep reinforcement learning algorithm. The training objective is to maximize renewable penetration while reducing microgrid power fluctuations and vehicle battery aging costs. The developed Energy-Transportation Nexus energy management method is verified to be effective in optimal power balancing and battery anti-aging control on a microgrid in the UK. This research provides an efficient and economical tool for microgrid power balancing by optimally coordinating GEVs charging and renewable energy, thus helping promote a low-cost decarbonization transition.**

***Index Terms*—Transportation electrification, electric vehicle, microgrid, deep reinforcement learning, renewable energy, battery aging mitigation, vehicle to grid.**

ABBREVIATIONS

GEVs	Grid-connected EVs.
V2G	Vehicle-to-grid.
MG	Microgrids.
C-rate	Charging and discharging rate.
DoD	Depth of discharge.
SoC	State of charge.
NoC	Number of cycles.
DRL	Deep reinforcement learning.
CTF	Cycle-to-failure.
CTA	Cycle-to-aging.
DRLVM	Deep reinforcement learning V2G management.
PSOS	Peak-shaving-oriented scheduling.
REC	Renewable energy consumption.
VCC	Vehicle charging completion.

NOMENCLATURE

$\mathbf{VPS}$	Battery SoC and discharging power matrix.
SoC_i	Battery SoC state at i .
P_i	V2G power at i [W].
$\mathbf{RF}$	DoDs and C-rates matrix.
$Crate_i$	Battery C-rate in cycle i .
DoD_i	Battery DoD in cycle i .
U_i	Terminal voltage of the battery [V].
E	Rated capacity of the battery [Ah].
CTF	Battery equivalent cycle life calculation function.
CTA	Battery equivalent life loss calculation function
η	Battery life loss calculation function.
s	Reinforcement learning system state value.
a	Reinforcement learning system action value.
r	Reinforcement learning system reward value.
$Q^{\pi}(s,a)$	Q-value of the action.
γ^t	Reward discount factor.
π	Reinforcement learning strategy
$\mathbf{P}_t^{v2g}$	24-hour ahead V2G power sequence.
$\mathbf{S}_t^{v2g}$	24-hour ahead battery SoC states sequence.
$\mathbf{B}_t$	Power balancing requirement of the MG.
P_{load}	Power consumption state [W].
P_{solar}	Solar power generation state [W].
P_{wind}	Wind power generation state [W].
$P_{\mathrm{dis}}^{\max}$	Maximum discharging power of GEVs [W].
$P_{\mathrm{ch}}^{\max}$	Maximum charging power of GEVs [W].
D	Quantified battery life loss of GEVs.
G	Unbalanced power of the microgrid [W].
m	Number of GEVs.
$\mathbf{X}_{s_t}$	Training input of Q-network.
$\mathbf{Y}_{Q_t}$	Training output of Q-network.
C	Loss value of Q-network.
n	Size of the mini-batch for Q-network.

Manuscript received December 22, 2021; revised February 01, 2022 and March 09, 2022; received March 23, 2022. Date of publication March 23, 2022. Paper no. TII-21-4976 (Corresponding author: Chenghong Gu)

Shuangqi Li is with the Department of Electronic and Electrical Engineering, University of Bath, Bath, BA2 7AY, U.K., and also with National Engineering Laboratory for Electric Vehicles, Beijing Institute of Technology, Beijing 100081, China. (e-mail: sqli9966@gmail.com).

Pengfei Zhao is with the Institute of Automation, Chinese Academy of Sciences, Beijing, China and School of Artificial Intelligence, University of Chinese Academy of Sciences, Beijing, 100190, China. (e-mail: P.zhao0308@gmail.com).

Chenghong Gu, Jianwei Li, Shuang Cheng, and Minghao Xu are with the Department of Electronic and Electrical Engineering, University of Bath, Bath, BA2 7AY, U.K. (e-mail: c.gu@bath.ac.uk; jl3220@bath.ac.uk; sc2791@bath.ac.uk; mx266@bath.ac.uk).

$\mathbf{w}^l$ Weight matrix of neuron in network layer l.
$\mathbf{b}^l$ Bias matrix of neuron in network layer l.
α Learning rate of Q-network.

I. INTRODUCTION

THE adoption of renewable energy brings a bright prospect to resolve environmental concerns, but its high penetration also brings great challenges for power systems. Microgrids (MG) have a relatively low power transmission voltage and can enhance the active interactions between renewable energy and load demand [1]. Nevertheless, due to the intermittence of renewable energy and variability of demand, the power balance of MG is difficult to maintain, and the mismatch could cause abrupt short-term power fluctuation issues. Conventional MG power balancing solutions, including adding additional reserve power generation, energy storage devices, and advanced power management strategies, to some extent fail in the economy, stability, or reliability [2, 3].

The electrification of road vehicles and the concept of Energy-Transportation Nexus brings a bright prospect to solve the power balancing problems of MG [4]. GEV onboard batteries can offer the means to enhance MG flexibility, achieving uninterrupted operation by deferring their demand in time and space, acting as moving storage devices. In [5], [6], and [7], GEV batteries are used to provide peak-shaving, voltage regulation, and frequency regulation services through V2G regulation. Simulation and experiment results indicate that with effective charging management strategies, energy quality and system stability can be significantly improved by utilizing the moving energy storage capacity provided by GEVs. In [8] and [9], V2G services are further used to support the operation of wind and solar power generation. GEV batteries can be used to provide power balancing services and improve the economy and stability of the MG with renewable energy penetration. However, there is still a gap that constrains the use of vehicle batteries, i.e., the cost of GEVs battery life loss due to providing V2G service. The situation is worse when battery energy storage is used in MG with renewable energy because battery may undergo excessive short-term cycles [10].

The establishment of the life loss analysis model is of great significance to battery energy management in electric vehicles [11], hybrid powertrains [12], and grid energy storage systems [13]. The quantified battery aging cost can be used as a benchmark of degradation-oriented mode of operation for guiding battery energy management and providing a life-cycle cost analysis tool in Energy-Transportation Nexus. Many studies have been conducted to quantify battery aging costs in energy management. The single-factor bucket model [14] is one of the most commonly used methods in the existing literature to protect vehicle batteries. In [15], optimal GEV charging management is built as an event-based scheduling model. Battery anti-aging is realized by limiting the charging and discharging rate (C-rate) in V2G management. It was shown that the established event-based model could achieve quasi-instantaneous system responsiveness and protect vehicle battery from high Crate working condition. Nevertheless, battery life is still impaired by number of cycles (NoC) and depth of discharging (DoD). In [16], the DoD is selected as the aging observation variable to protect vehicle batteries when providing V2G services. The bucket model can be used in online V2G scheduling because of remarkable real-time performance and hardware applicability. However, batteries are complex electrochemical systems and multi factors influence their aging. The single-factor bucket model cannot systematically analyze the impact of these cycles on battery life, impairing the anti-aging and power balancing performance of V2G management.

Complex battery aging models have been proved to be effective and necessary for mitigating GEVs aging in V2G services [14]. A battery aging model that integrates the impact of temperature, C-rate, state of charge (SoC), and DoD is built in [17] to minimize the expected customer's charging cost when providing V2G services. Stochastic optimization is used to derive optimal strategies, and simulation results validate its satisfactory battery anti-aging performance. Nevertheless, the deployment of complex aging models and large-scale optimization algorithms makes the V2G model complex, which further weakens scheduling system's real-time performance [18]. In [19] and [20], the mitigation of aging costs in V2G services is realized by using comprehensive battery aging models and heuristic algorithms. Simulation results indicate that the battery degradation costs can be effectively reduced. However, the optimization-based V2G scheduling interval can hardly be shortened to 5 minutes even with the most advanced computing equipment, making it impossible to suppress transient MG demand and renewable energy fluctuations.

Deep reinforcement learning (DRL) has attracted much research attention in recent years for its high calculation efficiency and satisfactory real-time performance. The DRL algorithm-based energy management model and the derived battery anti-aging strategy can be used to define the optimal battery utilization strategy in Energy-Transportation Nexus, such as hybrid electric vehicle [21], rail transportation system [22], and GEVs charging scheduling [23]. In [24], DRL algorithm is used to realize online energy management for plug-in hybrid electric buses. The improvement of vehicle fuel economy and the mitigation of battery degradation costs are the learning target. Simulation results validate the effectiveness of the DRL method in reducing overall vehicle driving costs in real-time energy management. DRL provides a new solution to improve both real-time power balancing and battery anti-aging performance of V2G scheduling. However, to the authors' best knowledge, there is no published work reporting the realization of optimal battery anti-aging V2G scheduling based on a DRL framework.

This paper develops a novel online battery protective energy management method for Energy-Transportation Nexus under a DRL framework. Firstly, based on the battery degradation characteristic analysis and rain-flow cycle counting technology, the quantification of degradation cost in V2G scheduling is designed as a function of battery NoC, DoD, and C-rate. The established aging cost model is used to evaluate the battery anti-aging effectiveness of V2G strategies in DRL. Then, the coordination of GEVs charging is modeled as multi-objective learning under DRL framework. The training target of the DRL model is to maximize renewable penetration while reducing vehicle battery aging costs. Historical MG power balance and GEVs battery states are used to construct an experience pool, and the charging/discharging strategies are online scheduled based on the trained DRL model. GEVs energy storage capacity can be scheduled online with the developed method to absorb renewable energy while mitigating vehicle battery aging phenomenon in V2G service.

The major contributions of this paper are:

1) It establishes a novel battery life loss analysis model to quantify GEVs energy storage system aging costs during daily operation. Compared to the existing single-factor aging model, it can comprehensively reflect the influence of NoC, C-rate, and DoD on battery life and thus provide a more precise battery life loss evaluation result.
2) It designs a new DRL-based battery protective V2G behavior management framework, which provides a model-free solution for GEVs charging schedule. Compared to most existing methods, the developed DRL framework enables the online deployment of battery anti-aging V2G regulator by combining offline learning and the online strategy deployment process.
3) A multi-objective learning method is proposed to train the DRL model. Optimal strategies that comprehensively consider the MG power balancing requirement, GEVs charging requirement, and the mitigation of battery aging can be thus derived.

The rest of the paper is organized as follows. The vehicle battery life loss quantification model is established in Section II. A DRL-based V2G behavior management framework and multi-objective learning model are proposed in Sections III and IV, respectively. In Section V, simulation environment and numeric analysis are presented to verify the developed methods. Section VI concludes the whole paper.

II. Battery energy storage system life loss quantification model

Related literature identifies many factors influencing battery health, which can be broadly classified into calendar and cycle aging [25]. Calendar aging comprises all aging processes that lead to the degradation of a battery cell independent of charge-discharge cycling [26]. Literature [13] validates that calendar aging is unavoidable and shows a limited impact on battery life in energy management. Instead, cycle aging, which is caused by battery cycles, is the main factor that results in life loss of GEVs [27]. Therefore, only the cycle aging is considered in this study. According to battery degradation modes analysis results derived in [28], cycles with different C-rates and DoDs impact battery life from different aspects to different degrees. Based on the above discussion, this part develops a battery aging quantification model to quantify battery life loss in V2G strategies by analyzing battery NoC, DoD, and Crate information in SoC and discharging power profiles of GEVs.

To further extract battery aging features, the following time series are constructed as the input of the life loss model:

$$\mathbf{VPS} = \begin{bmatrix} SoC_1 & SoC_2 & \cdots & SoC_i & \cdots & SoC_m \\ P_1 & P_2 & \cdots & P_i & \cdots & P_m \end{bmatrix} \tag{1}$$

The rain-flow cycle-counting method, which has been proven effective in extracting industry material aging cycles [29], is used here to extract the battery aging features. Battery C-rate [30] is calculated by the following equation based on the discharging power of GEVs in (1):

$$Crate_i = \frac{P_i}{U_i \cdot E} \tag{2}$$

Where: U_i and E are the terminal voltage and rated capacity of the battery. The calculated battery DoDs and C-rates in $\mathbf{VPS}$ time series are further extracted and arranged in a feature matrix $\mathbf{RF}$ for quantifying aging cost:

$$\mathbf{RF} = \begin{bmatrix} Crate_1 & Crate_2 & \cdots & Crate_i & \cdots & Crate_n \\ DOD_1 & DOD_2 & \cdots & DOD_i & \cdots & DOD_n \end{bmatrix} \tag{3}$$

Where: $Crate_i$ and DOD_i are the extracted battery discharging rate and depth of discharge information in $cycle\ i$.

The battery pack in vehicle energy storage systems contains hundreds of independent cells that own different aging states and characteristics. To simplify the life loss quantification process, the cycle-to-failure (CTF) characteristic profile provided by the manufacture [31], which describes battery pack aging characteristics by using cell aging experiment results, is employed in this study to quantify battery life loss in V2G services. Based on the extracted aging features data in (2), the following equation is used to describe battery nonlinear aging characteristics under different DoD states:

$$\lambda(DOD) = \alpha \cdot e^{-\frac{DOD-\sigma^2}{\varepsilon}} + \beta \cdot e^{-\frac{DOD-\varsigma^2}{\mu}} \tag{4}$$

Where: α , β , σ , ς , ε , and μ are curve fitting parameters, which can be calculated from battery CTF data derived in our previous work [32]. CTF profile reflects the influence of DoD and NoC on battery degradation. However, battery aging is also affected by the C-rate. In order to better quantify the impact of C-rate on life loss in V2G service, empirical models established in [33] is used here to correct the life loss map:

$$\hbar(\text{Crate}) = \frac{e^{0.05\times 25}}{e^{0.05} \times ((0.507 \cdot Crate^2 + 0.2906) \times 2 + 25)} \tag{5}$$

Based on the above analysis, battery equivalent cycle life under different DoD and C-rate working conditions can be calculated by multiplying λ and $\hbar$, which can be represented by the following equation:

$$CTF(DOD, Crate) = \lambda(DOD) \cdot \hbar(\text{Crate}) \tag{6}$$

Where: CTF is a function to calculate battery equivalent cycle life under the given working conditions (DoD and C-rate). The constructed CTF responding profile is shown in Fig. 1, where battery cycle life decreases with the increase of battery C-rate and DoD.

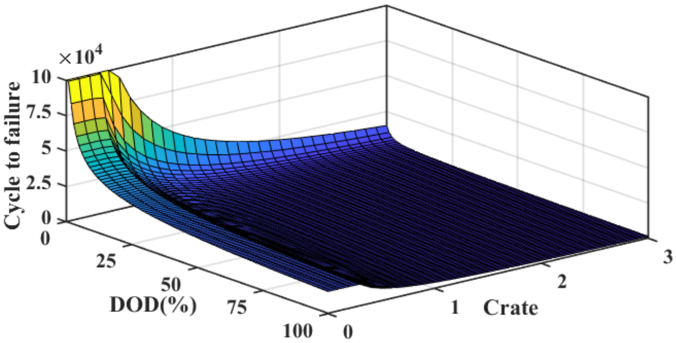


Fig. 1. The constructed battery CTF responding profile.

CTF profile describes battery-rated cycle life under different working conditions. In this study, the concept of cycle-to-aging (CTA) is further defined to quantify the impact of cycles with different DoDs and C-rates on battery life loss in V2G scheduling. The reciprocal of CTF value, which reflects percentage battery life loss, is defined as the CTA value of different cycles:

$$CTA(Cycle_i)=\frac{1}{CTF(Cycle_i)}=\frac{1}{\lambda(DOD_i)\cdot\hbar(\mathrm{Crate}_i)} \tag{7}$$

Battery percentage life loss in V2G strategies can be derived by aggregating the CTA value of different cycles, which can be depicted by the following function:

$$D=\eta(\mathbf{VPS})=\sum\nolimits_{i=1}^{n}CTA(Cycle_i)\times 100\% \tag{8}$$

Where: D is the quantified percentage vehicle battery system life loss in n cycles in V2G power and SoC time series $\mathbf{VPS}$.

III. Deep Reinforcement Learning-Based V2G Behavior Management Framework

In this study, dynamic V2G behavior management is solved with DRL algorithm to determine the optimal power exchange between the MG and GEVs. This section develops a deep reinforcement learning V2G management (DRLVM) framework to enable model-free GEVs charging scheduling by combining offline learning and online strategy deployment processes. As shown in Fig. 2, the DRLVM framework consists of 4 parts: agent, environment, experience pool, and Deep Q-network.

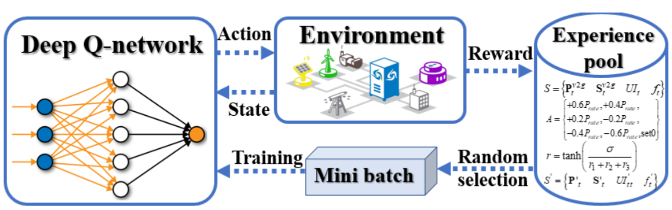


Fig. 2. The designed deep-reinforcement learning-based V2G behavior management framework.

The agent in the developed DRLVM framework is selected as an individual V2G participant, and the charging behavior of GEVs is the decision variable. The coordination of GEV charging at different times in the scheduling period is modelled as a Markov decision process, described by the following three essential elements: (a) state s, (b) action a, and reward r. The state variable in the DRLVM framework is designed as the historical V2G behavior of GEVs and power balance state of the microgrid, which can be represented by the following equations:

$$\mathbf{S}=\left\{\mathbf{P}_t^{v2g}\quad \mathbf{S}_t^{v2g}\quad \mathbf{B}_t\right\} \tag{9}$$

$$\mathbf{P}_t^{v2g}=\left\{P_1^{v2g}\quad\cdots\quad P_{t-1}^{v2g}\right\} \tag{10}$$

$$\mathbf{S}_t^{v2g}=\left\{SoC_1^{v2g}\quad\cdots\quad SoC_{t-1}^{v2g}\right\} \tag{11}$$

$$\mathbf{B}_t=\left\{P_{load,t}\quad P_{solar,t}\quad P_{wind,t}\right\} \tag{12}$$

Where: P^{v2g}, SoC^{v2g}, P_{load}, P_{solar} and P_{wind} are V2G power state, battery energy state, microgrid power consumption state, solar power generation state, and wind power generation states. Considering GEVs battery power dynamics, its power increment value is set as the action variable in the designed DRLVM:

$$A=\{+10,+5,+2,0,-2,-5,-10,\mathrm{set}0\} \tag{13}$$

Where: the units of all action values are kilowatt. The positive value indicates the improvement of V2G power, while the negative value represents the improvement of battery charging power. After the action is taken, the battery power output at each time step is constrained as:

$$-P_{\mathrm{dis}}^{\max}\le P_t^{v2g}\le P_{\mathrm{ch}}^{\max} \tag{14}$$

Where: $P_{\mathrm{dis}}^{\max}$ and $P_{\mathrm{ch}}^{\max}$ are the maximum discharging and charging power of GEV batteries.

In each iteration, the agent takes actions to schedule the charging power of GEVs based on the learned strategies and the state of the environment, including power generation and consumption states of the MG and SoC states of the vehicle battery. Meanwhile, according to the response of the environment, including the MG power balancing performance and the calculated battery aging cost, the quality of each action is evaluated by a reward function, and the state of the agent is automatically updated to s' based on the selected action and MG model.

In the training phase, the agent freely explores the action space as much as possible, and system state transfer process (a_t, s_t, r_t, s_{t+1}) is recorded in an experience pool. The Q-value of the action, which reflects the quality of the strategy, is calculated by the following equation:

$$Q^{\pi}(s,a)=\sum\nolimits_{t=0}^{\infty}\left\{\gamma^t R(s_t,a_t)\,||\,s_t=s,a_t=a\right\} \tag{15}$$

Where: γ^t is a constant variable within the range of 0 to 1, reflecting the discounted impact of future reward value on the current iteration step. In this study, the Epsilon-Greedy method [34] is employed to perform action selection based on the calculated Q-values during the DRLVM model training process.

Compared to conventional decision-making, the environment state variables in V2G scheduling are all continuous variables. Meanwhile, the historical V2G power and battery SoC should also be considered to enable battery anti-aging scheduling, which further complicates the computation burden during solving Q value. Therefore, a state continuous V2G scheduling algorithm is employed in this study based on deep neural network technology. As shown in Fig. 2, a deep network is used to estimate Q-values under continuous system state change in the decision system. The historical system state transfer processes (a_t, s_t, r_t, s_{t+1}) and corresponding Q-values are randomly selected from the experience pool to train the network. The estimated Q-value can be represented by the following equation:

$$Q^*(s,a) \approx \psi(s,a;w,b) \tag{16}$$

Where: ψ represents the transfer function of the trained deep neural network, w and b are the weights and biases in it.

The trained deep Q network is used to coordinate the charging of GEVs in the real world. The corresponding V2G scheduling strategies are derived by performing the action that has the maximum Q-value, expressed as:

$$\pi = \arg\max_a Q\left(s, a \mid w^Q, b^Q\right) \tag{17}$$

IV. Multi-objective learning model in V2G behavior management

This section provides the mathematical principle and establishes the learning model for DRLVM framework. Firstly, the V2G behavior learning is realized by establishing a multi-objective reward function that can comprehensively reflect the MG power balancing and battery anti-aging requirements. Then, the structure of the built deep-Q network and model training method are detailed.

A. Design of multi-objective reward function

The reward function is used to guide agents to make appropriate decisions, so its definition should be consistent with the objective of V2G scheduling. This part establishes a multi-objective reward function to minimize MG load fluctuation and battery life losses in DRLVM. The mitigation of battery degradation is the first target. GEVs charging power and SoC trajectory are extracted from the historical V2G strategy base and rearranged in a time series, as described in equations (10) and (11). Based on the established life loss quantification model in Section II, battery life loss in V2G strategies can be calculated as:

$$D = \eta(\mathbf{P}_t^{v2g}, \mathbf{S}_t^{v2g}) \tag{18}$$

The mitigation of load fluctuation and absorption of renewable power generation are also designed as the training targets of DRLVM to improve the economy and stability of the MG. The unbalanced power of the MG with GEVs penetration is selected as the second reward function:

$$G = P_{load} + m \cdot P_{v2g} - P_{\text{solar}} - P_{\text{wind}} \tag{19}$$

Where: m is the number of GEVs, which is used to reflect the aggregation effect in V2G service. It should be noted that all GEVs are assumed to contribute the same V2G power to the MG when calculating the reward function G. The reason is that the control object in DRLVM is the individual participant, and it is not permitted to set multi-step reward functions in DRL algorithm.

To comprehensively reflect battery aging mitigation, renewable energy fluctuation, and charging requirement of GEVs, the following multi-objective reward function are built to evaluate action quality:

$$\begin{cases} r_1 = \omega_1 D \\ r_2 = \omega_2 G \\ r_3 = \omega_3 (1 - SoC) \\ r = \tanh\left(\dfrac{\sigma}{r_1 + r_2 + r_3}\right) \end{cases} \tag{20}$$

Where: ω_1, ω_2, and ω_3 weight factors between the three different rewards. r_3 is used to reflect the charging requirement of participants. The larger the reward r, the worst the power balancing and battery anti-aging performance of the derived V2G strategies. Thus, minimizing its value in DRL training can help DRLVM coordinate the charging behavior of GEVs reasonably. σ is a constant to adjust the range of tangent function.

B. Deep-Q network structure and training method

In the designed DRLVM, the Q-value of different actions should be estimated to direct the charging behavior of GEVs. The estimation of Q-value in DRLVM can be regarded as a multi-input to multi-output regression problem. The complex mapping relationship between the outputs and inputs makes it difficult to learn the regularity between the state of the decision system and the Q-value of actions.

The neural network is one of the most commonly used artificial intelligence algorithms, which simulates the working mode of human brain neurons with abstract mathematical models and many nodes. Neural network composes different layers, and neurons in different layers perform operations according to different functions, transfer values, and finally merge into a complex network for curve fitting purposes. As long as the reasonable network structure and network parameters are properly designed, the neural network can theoretically map any relationships. This study uses a multi-layers deep neural network to fit the calculated Q-value for improving the generalization ability of the learning process, better dealing with continuous grid and GEVs state variables, and improving optimization effect of the established V2G coordinator. The network is trained by following loss function:

$$C = \frac{1}{2n} \sum_{i=1}^{n} \sum_{x=1}^{8} (Y_{Q_t,i}(\mathbf{X}_{s_t}) - \hat{Y}_{Q_t,i}(\mathbf{X}_{s_t}))^2 \tag{21}$$

Where: $\mathbf{X}_{s_t}$ is the training input of the Q-network, which consists of system state variable at t. $\mathbf{Y}_{Q_t}$ is the Q value of different actions, which can be calculated based on equation (20). $\hat{Y}_{Q_t}$ is the output of the Q-network. n is the size of the selected mini-batch.

The experience replay method [35] is used to update parameters of the Q network to boost training efficiency and accuracy. Neural network parameters can be updated by the following equation:

$$\mathbf{w}^l \leftarrow \mathbf{w}^l - \frac{\alpha}{n} \nabla_{W^l} C(\mathbf{W}, \mathbf{b}) \tag{22}$$

$$\mathbf{b}^l \leftarrow \mathbf{b}^l - \frac{\alpha}{n} \nabla_{b^l} C(\mathbf{W}, \mathbf{b}) \tag{23}$$

Where: $\mathbf{w}^l$ and $\mathbf{b}^l$ are the weight and bias of neuron in layer l; α is learning rate of Q-network.

V. CASE STUDY

This section illustrates the performance of the developed DRLVM method. The topology and parameters of the studied MG system are firstly presented, followed by the power balancing and vehicle battery anti-aging performances are evaluated.

A. The test microgrid system

The configuration of the test MG system with household load demand, GEVs, and renewable energy penetrations is shown in Fig. 3. Real grid demand and solar power generation data are provided by Western Power Distribution, an electricity distribution company in the UK. The demand data comes from the Stentaway Primary substation near Plymouth, and the corresponding solar generation data is from a 5MW solar near the studied community (longitude, latitude = 50.33, -4.034), UK. Wind power generation data is calculated based on the local wind speed and the stochastic simulation model in [36]. All the data used in this paper has been open-access provided on [37]. The steady-state MG simulation model in [38] is used to model the power conversion between different sectors to verify the effectiveness of the developed V2G behaviour management method. To facilitate the hardware deployment of the established coordinator, all developed methods are programmed with the deep reinforcement learning toolbox in Simulink.

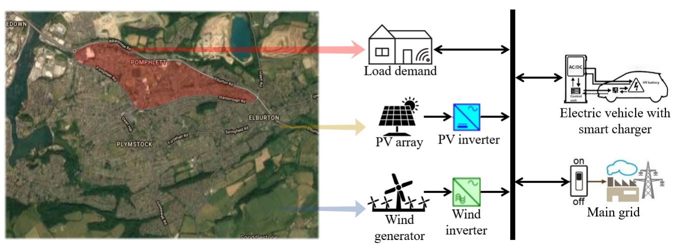


Fig. 3. The configuration of the studied MG system with renewable energy penetration.

In this study, the charging behaviors of 350 GEVs are simulated to provide power balancing services to the MG. The detailed battery characteristic parameter of the studied GEVs are illustrated in Table I. The rated capacity of the battery pack in each GEVs is 53 kWh, which consists of 10 modules connected with a 2p5s configuration. The battery module consists of 444 Lithium Ion cells with 3400mAh rated capacity and 3.8 V nominal voltage, and the rated discharging current reaches 500 A. The charging and discharging voltage cut-off of the battery cell is 4.2 V and 3.3 V, respectively.

TABLE I. BATTERY CHARACTERISTIC PARAMETERS OF THE SIMULATED GEVS FLEET

Parameters	Value
Battery cell type	Lithium-Ion 18650
Number of cells	444
Battery Module capacity	232Ah, 5.3 kWh
Voltage nominal	3.8V/Cell, 22.8V/Module
Charging voltage cut-off	4.2V/Cell, 25.2V/Module
Discharging voltage cut-off	3.3V/Cell, 19.8/Module
Rated discharging current	500 A
Battery pack configuration	2p5s
Battery pack capacity	53 kWh

B. Power balancing performance evaluation

The demand, solar power generation, and wind power generation profiles of the studied MG within one year are given in Fig. 4. The power consumption in winter and autumn is generally higher than that in spring and summer because of the use of heating installations, as shown in (a). Meanwhile, two peaks generally appear in grid demand profiles in the period of 08:00 to 10:00 and 17:00 to 20:00 because of the boom of commercial and household electricity consumption. Different from demand profiles, solar generations generally peak in the period of 10:00 to 16:00, while no PV output power can be provided after 19:00 until the morning. Compared to PV output profiles, the regularity of wind profile is not remarkable due to the uncertain wind speed, but the wind power generation in the evening is generally higher than daytime. The corresponding wind power generation states distribution is shown in (c), the average value is 1.95MW while the standard deviation (SD) reaches 1.874. Above demand and renewable generation profiles are used to train the established DRLVM model. The training targets are set to stabilize MG power balance state by using V2G services while mitigating vehicle battery aging costs.

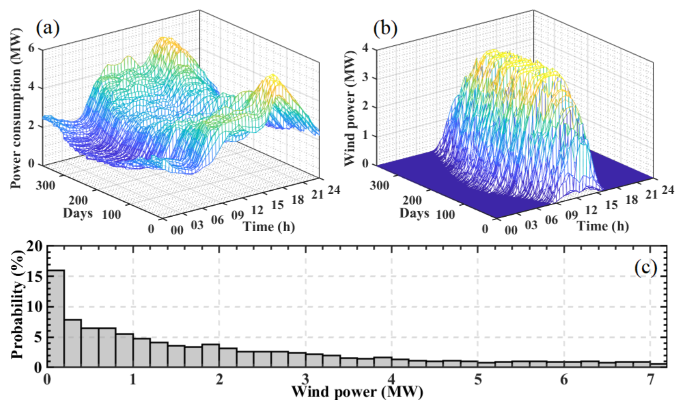


Fig. 4. DRL model training data. (a) MG demand profiles; (b) solar power generation profiles; (c) wind power generation states distribution.

Based on the above power system configuration, performances of four different V2G scheduling algorithms, including conventional fuzzy logic method [39] (Case 1), peak-shaving-oriented scheduling (PSOS) method [40] (Case 2), Q-learning method [41] (Case 3), and the DRLVM method (Case 4), are quantitatively compared in this section. The power balancing performance of different V2G scheduling methods within 250 working days is analyzed in Fig. 5. In terms of algorithm computation speed, the average simulation time of the PSOS method is as long as 265.4 s due to the complex optimization mechanism. Compared to the PSOS method, GEVs charging behavior can be directly scheduled based on the rules but free of optimization process in fuzzy logic method. As a result, the simulation time in Case 2 can be reduced to 0.13s. Owing to the offline training mechanism, the Q-learning and DRLVM methods achieve a similar calculation speed with the fuzzy logic method, and the simulation time can be limited to 0.25 s and 0.27 s. Therefore, online scheduling methods can better deal with variant renewable power generation and demand fluctuation compared to the optimization-based PSOS method. In this study, to guarantee system stability, the scheduling interval in Case 1 to 4 are set as 1 s, 300 s, 1 s, and 1 s, respectively.

The power balancing performance of four cases is compared in Fig. 5 (b) and (c). In V2G scheduling, GEV batteries are used to absorb renewable power generation as much as possible. The renewable energy consumption (REC) in different V2G schemes is shown in (b). V2G system with the fuzzy logic algorithm is not able to manage the charging behavior of GEVs synergistically, and thus the provided energy storage capacity is limited. As a result, only around 62.3% of renewable power generation can be consumed. Compared with the fuzzy logic method, the REC rate in PSOS method can be improved to 87.5% by better optimizing the charging behavior of GEVs. The variant renewable power generation and demand fluctuation can be better dealt with reinforcement learning method because of the shorter scheduling interval. Compared to the PSOS method, the REC rate can be further improved by 9.4% and 6.3% after the Q-learning and DRLVM methods are deployed.

The unbalanced power, which reflects the required power exchange between the MG and main grid, is further used here to evaluate the power balancing performance of V2G scheduling methods. As shown in Fig. 5 (c), in fuzzy logic, Q-learning, and DRLVM methods, the unbalanced power can be generally limited to 100 kW, which validates the gratifying power balancing performance of online methods. It should be figured out that the unbalanced power of the DRLVM method is around 25.2% higher than the Q-learning method. The reason is that the consideration of GEVs battery anti-aging requirement inevitably limits the potential utilization degree of GEV batteries in providing power balancing service. Furthermore, compared to conventional fuzzy logic and the Q-learning method, the developed DRLVM method can strictly satisfy the charging requirement of participants. As shown in (d), the vehicle charging completion (VCC) rate can be improved from 91.2% to 100% after deploying the DRLVM method.

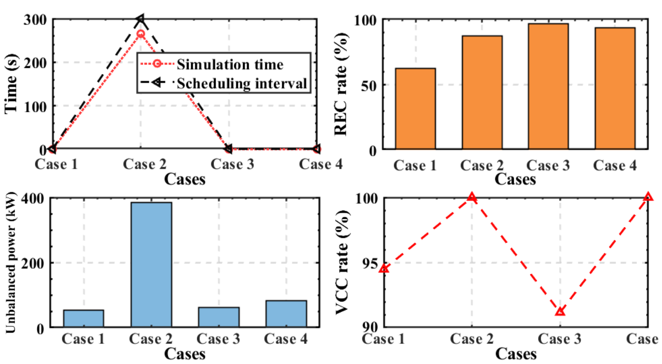


Fig. 5. Power balancing performance comparison of different cases. (a) simulation time; (b) V2G renewable energy consumption rate; (c) system unbalanced power; (d) vehicle charging completion rate;

The sensitivity analysis is carried out in this study to validate the power balancing performance of DRLVM method under the variation of vehicle battery capacity, fleet scale, and wind power generation fluctuation. Here, the fluctuation of wind energy is evaluated by the standard deviation (STD) of power generation data. The rate of change of wind power fluctuation in Day_i is calculated as:

$$RoC_{wind,i} = \frac{STD_i - \sum_{i=1}^{n} STD_i / n}{\sum_{i=1}^{n} STD_i / n} \times 100\% \qquad (24)$$

Where: STD_i is the standard deviation of wind power generation data in Day_i, n is the length of the simulation period. The corresponding sensitivity analysis result is shown in Fig. 6. System unbalanced grid power increases with the reduction of vehicle battery capacity and fleet scale in the DRLVM method. When battery capacity and fleet scale decrease by 30%, unbalanced grid power can still be limited within 121.4 kW and 153.8 kW, which indicates that the scheduling algorithm can keep stable operation under the variation of GEVs energy storage capacity. It should be figured out that the impact of fleet scale on system power balancing performance is higher than battery capacity. The reason is that the reduction of GEVs fleet can dramatically impair scheduling algorithm flexibility. Compared with the change of energy storage capacity, the renewable power generation fluctuation shows a limited influence on V2G scheduling. The unbalanced power can be limited to 107.6 kW even the rate of change of wind power fluctuation reaches 30%, which validates the robustness of the developed DRLVM method.

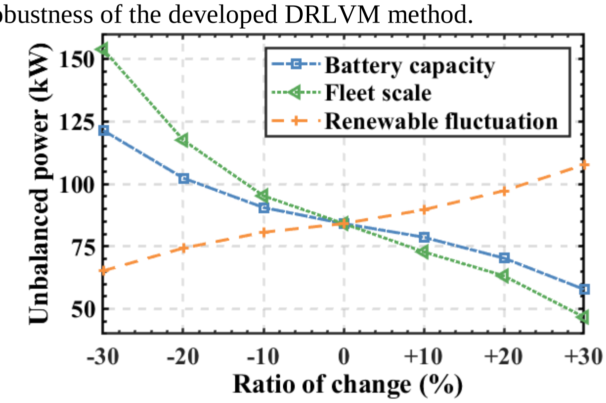


Fig. 6. DRLVM method power balancing performance sensitivity analysis.

Battery SoC profiles of a GEV in a regular working day in fuzzy logic and DRLVM methods are shown in Fig. 7. Compared to the fuzzy logic method, battery NoC in V2G scheme with DRLVM method is significantly reduced. GEVs are scheduled to absorb renewable energy as much as possible in fuzzy logic method. As a result, battery undergoes a great number of shallow cycles when dealing with variant wind power generation in the evening, as shown in Zone C. In the DRLVM method, instead of inversing battery charging state, V2G scheduling system can absorb renewable power generation by adjusting the battery working power. Therefore, battery cycles in DRLVM method can be significantly reduced, which validates the effectiveness of the established battery aging quantification model. The developed DRLVM method can also protect vehicle batteries from high C-rate working conditions that greatly impact their life. As shown in Zone A and B, the battery SoC rate of change during discharging and charging processes is significantly reduced in (b), which indicates that the developed DRLVM can reach a better trade-off between the power balancing and the battery lifetime protection.

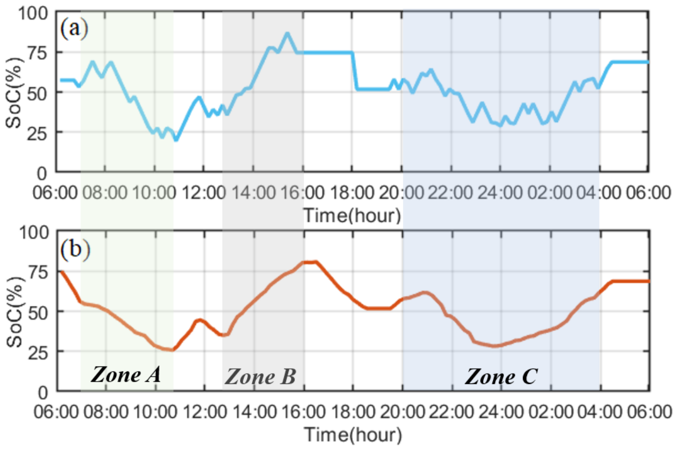


Fig. 7. Battery SoC profiles of a GEV in a regular working day in (a) fuzzy logic method and (b) the developed DRLVM method.

The battery anti-aging performance of different V2G scheduling methods in the whole simulation period is quantitatively analyzed in Table II. The charging behavior of GEVs can be better coordinated in the PSOS method because of the cooperative optimization mechanism. The battery number of cycles and C-rate in the simulation period can be reduced by 23.4% and 17.9% compared to the fuzzy logic method. The Q-learning method achieves a very similar performance compared to the fuzzy logic method, but the battery cycles and C-rate can be further reduced to 1875 and 1.24 after the developed aging model and multi-objective learning method is deployed. Based on the battery aging model in Section II, battery life loss within 250 working days under different V2G scheduling methods is quantified. Compared to fuzzy logic, PSOS, and Q-learning method, the developed DRLVM method can reduce battery life loss by 60.2%, 24.4%, and 51.2%, respectively. Battery life loss can be limited to 6.27% in the simulation period, which validates the effectiveness of the developed DRLVM method.

TABLE II. QUANTITATIVE PERFORMANCE EVALUATION OF DIFFERENT V2G SCHEDULING METHODS

Scenario	Case 1: Fuzzy logic method	Case 2: PSOS method	Case 3: Q-learning method	Case 4: DRLVM method
Number of cycles	2552	1954	2434	1875
Average C-rate	1.78	1.46	1.74	1.24
Battery life loss (%)	15.75	8.29	12.85	6.27

The sensitivity analysis is further carried out to analyze battery life loss under the variation of vehicle battery capacity, fleet scale, and wind power generation fluctuation. As shown in Fig. 8, both the battery capacity and fleet scale show a positive effect on the reduction of life loss in V2G services, while the renewable fluctuation shows a negative effect. The developed DRLVM method can stably work under GEVs energy storage capacity variation. When battery capacity and fleet scale decrease by 30%, battery life loss in V2G services can still be limited within 8.37% and 8.94%, respectively. Similarly, the impact of fleet scale is higher than battery capacity on V2G system battery protective performance. Renewable power generation fluctuation shows a very limited influence on V2G scheduling. Battery life loss can be limited to 6.62% even the rate of change of wind power fluctuation reaches 30%, which validates the robustness of the developed DRLVM method.

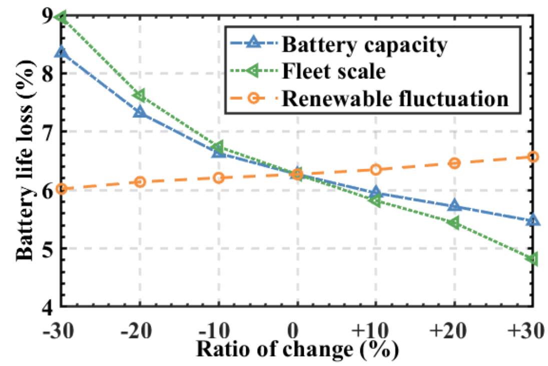


Fig. 8. DRLVM method battery protective performance sensitivity analysis.

VI. CONCLUSION

A novel battery anti-aging V2G scheduling method that can provide power balancing services for the MG by utilizing GEVs energy storage capacity is developed in this paper. GEVs aging cost in V2G scheduling is quantified by a battery degradation model. The optimal GEVs charging coordination is modelled as a multi-objective learning problem under DRL framework. Through extensive simulations on an MG system built with real power generation and consumption data in the UK, the key findings are as follows: (1) Compared to the bucket model, the established aging cost analysis model can model battery aging characteristics more comprehensively. Vehicle battery life loss in V2G service can be significantly reduced after the developed battery aging quantification model is deployed. (2) Benefiting from offline training, the reinforcement learning-based V2G scheduling can real-time schedule the charging behavior of GEVs to mitigate the volatility of renewable energy. As a result, MG unbalanced power and REC rate can be significantly reduced and improved.

Furthermore, the applicability of the methodology developed in this paper can be summarized as follows: (1) The established battery life loss analysis model can be used as a benchmark of degradation-oriented mode of operation for guiding battery energy management and providing an effective life-cycle cost analysis tool. (2) The established DRL-based V2G scheduling model and the simulation results in this study define the optimal vehicle battery utilization strategy in smart energy systems considering degradation, which can further improve Energy-Transportation Nexus efficiency.

REFERENCES

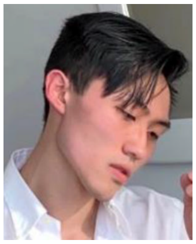
[1] B. Fan, S. Guo, J. Peng, Q. Yang, W. Liu, and L. Liu, "A Consensus-Based Algorithm for Power Sharing and Voltage Regulation in DC Microgrids," *IEEE Transactions on Industrial Informatics,* vol. 16, no. 6, pp. 3987-3996, 2020.

[2] B. Chen, J. Wang, X. Lu, C. Chen, and S. J. I. T. o. S. G. Zhao, "Networked microgrids for grid resilience, robustness, and efficiency: a review," vol. 12, no. 1, pp. 18-32, 2020.

[3] B. Wang, M. Sechilariu, and F. J. I. t. o. s. g. Locment, "Intelligent DC microgrid with smart grid communications: Control strategy consideration and design," vol. 3, no. 4, pp. 2148-2156, 2012.
[4] S. Li, C. Gu, X. Zeng, P. Zhao, X. Pei, and S. Cheng, "Vehicle-to-grid management for multi-time scale grid power balancing," *Energy,* vol. 234, p. 121201, 2021/11/01/ 2021.
[5] X. Li *et al.*, "A cost-benefit analysis of V2G electric vehicles supporting peak shaving in Shanghai," vol. 179, p. 106058, 2020.
[6] Y. Rao, J. Yang, J. Xiao, B. Xu, W. Liu, and Y. J. E. Li, "A frequency control strategy for multimicrogrids with V2G based on the improved robust model predictive control," vol. 222, p. 119963, 2021.
[7] M. Mazumder and S. J. I. S. J. Debbarma, "EV Charging Stations With a Provision of V2G and Voltage Support in a Distribution Network," vol. 15, no. 1, pp. 662-671, 2020.
[8] S. Vachirasricirikul and I. J. I. T. o. S. G. Ngamroo, "Robust LFC in a smart grid with wind power penetration by coordinated V2G control and frequency controller," vol. 5, no. 1, pp. 371-380, 2014.
[9] M. J. E. Alam, K. M. Muttaqi, and D. J. I. T. o. S. G. Sutanto, "Effective utilization of available PEV battery capacity for mitigation of solar PV impact and grid support with integrated V2G functionality," vol. 7, no. 3, pp. 1562-1571, 2015.
[10] C. Zhou, K. Qian, M. Allan, and W. J. I. T. o. E. C. Zhou, "Modeling of the cost of EV battery wear due to V2G application in power systems," vol. 26, no. 4, pp. 1041-1050, 2011.
[11] M. Dubarry and D. J. J. o. P. S. Beck, "Big data training data for artificial intelligence-based Li-ion diagnosis and prognosis," vol. 479, p. 228806, 2020.
[12] L. De Pascali, F. Biral, and S. J. I. T. o. V. T. Onori, "Aging-aware optimal energy management control for a parallel hybrid vehicle based on electrochemical-degradation dynamics," vol. 69, no. 10, pp. 10868-10878, 2020.
[13] A. Maheshwari, N. G. Paterakis, M. Santarelli, and M. Gibescu, "Optimizing the operation of energy storage using a non-linear lithium-ion battery degradation model," *Applied Energy,* vol. 261, p. 114360, 2020/03/01/ 2020.
[14] J. M. Reniers, G. Mulder, S. Ober-Blöbaum, and D. A. Howey, "Improving optimal control of grid-connected lithium-ion batteries through more accurate battery and degradation modelling," *Journal of Power Sources,* vol. 379, pp. 91-102, 2018/03/01/ 2018.
[15] H. Krueger and A. Cruden, "Multi-Layer Event-Based Vehicle-to-Grid (V2G) Scheduling With Short Term Predictive Capability Within a Modular Aggregator Control Structure," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 5, pp. 4727-4739, 2020.
[16] M. Ihara, S. Tianmeng, and H. Nishi, "A simulation study of electric power leveling using V2G infrastructure," in *IECON 2011 - 37th Annual Conference of the IEEE Industrial Electronics Society*, 2011, pp. 3224-3229.
[17] M. Ebrahimi, M. Rastegar, M. Mohammadi, A. Palomino, and M. Parvania, "Stochastic Charging Optimization of V2G-Capable PEVs: A Comprehensive Model for Battery Aging and Customer Service Quality," *IEEE Transactions on Transportation Electrification,* vol. 6, no. 3, pp. 1026-1034, 2020.
[18] C. Chen, J. Chen, Y. Wang, S. Duan, T. Cai, and S. Jia, "A Price Optimization Method for Microgrid Economic Operation Considering Across-Time-and-Space Energy Transmission of Electric Vehicles," *IEEE Transactions on Industrial Informatics,* vol. 16, no. 3, pp. 1873-1884, 2020.
[19] X. Zeng, M. S. Nazir, M. Khaksar, K. Nishihara, and H. Tao, "A day-ahead economic scheduling of microgrids equipped with plug-in hybrid electric vehicles using modified shuffled frog leaping algorithm," *Journal of Energy Storage,* vol. 33, p. 102021, 2021/01/01/ 2021.
[20] S. Tabatabaee, S. S. Mortazavi, and T. Niknam, "Stochastic scheduling of local distribution systems considering high penetration of plug-in electric vehicles and renewable energy sources," *Energy,* vol. 121, pp. 480-490, 2017/02/15/ 2017.
[21] T. Liu, X. Hu, W. Hu, and Y. Zou, "A Heuristic Planning Reinforcement Learning-Based Energy Management for Power-Split Plug-in Hybrid Electric Vehicles," *IEEE Transactions on Industrial Informatics,* vol. 15, no. 12, pp. 6436-6445, 2019.
[22] K. Deng *et al.*, "Deep reinforcement learning based energy management strategy of fuel cell hybrid railway vehicles considering fuel cell aging," *Energy Conversion and Management,* vol. 251, p. 115030, 2022/01/01/ 2022.
[23] M. Dubarry, A. Devie, K. Stein, M. Tun, M. Matsuura, and R. J. J. o. P. S. Rocheleau, "Battery Energy Storage System battery durability and reliability under electric utility grid operations: Analysis of 3 years of real usage," vol. 338, pp. 65-73, 2017.
[24] J. Wu, Z. Wei, W. Li, Y. Wang, Y. Li, and D. U. J. I. T. o. I. I. Sauer, "Battery thermal-and health-constrained energy management for hybrid electric bus based on soft actor-critic DRL algorithm," vol. 17, no. 6, pp. 3751-3761, 2020.
[25] D.-I. Stroe, M. Swierczynski, A.-I. Stroe, R. Laerke, P. C. Kjaer, and R. J. I. T. o. I. A. Teodorescu, "Degradation behavior of lithium-ion batteries based on lifetime models and field measured frequency regulation mission profile," vol. 52, no. 6, pp. 5009-5018, 2016.
[26] M. Dubarry, A. Devie, and K. J. J. o. P. S. McKenzie, "Durability and reliability of electric vehicle batteries under electric utility grid operations: Bidirectional charging impact analysis," vol. 358, pp. 39-49, 2017.
[27] K. Uddin, M. Dubarry, and M. B. J. E. P. Glick, "The viability of vehicle-to-grid operations from a battery technology and policy perspective," vol. 113, pp. 342-347, 2018.
[28] M. Dubarry, C. Truchot, and B. Y. Liaw, "Synthesize battery degradation modes via a diagnostic and prognostic model," *Journal of Power Sources,* vol. 219, pp. 204-216, 2012/12/01/ 2012.
[29] R. C. Ugras, O. K. Alkan, S. Orhan, M. Kutlu, A. J. M. S. Mugan, and S. Processing, "Real time high cycle fatigue estimation algorithm and load history monitoring for vehicles by the use of frequency domain methods," vol. 118, pp. 290-304, 2019.
[30] F. Todeschini, S. Onori, and G. Rizzoni, "An experimentally validated capacity degradation model for Li-ion batteries in PHEVs applications," *IFAC Proceedings Volumes,* vol. 45, no. 20, pp. 456-461, 2012/01/01/ 2012.
[31] Q. Badey, G. Cherouvrier, Y. Reynier, J. Duffault, and S. J. C. T. E. Franger, "Ageing forecast of lithium-ion batteries for electric and hybrid vehicles," vol. 16, pp. 65-79, 2011.
[32] S. Li, H. He, C. Su, and P. Zhao, "Data driven battery modeling and management method with aging phenomenon considered," *Applied Energy,* vol. 275, p. 115340, 2020/10/01/ 2020.
[33] Y. Wang, X. Jiao, Z. Sun, and P. Li, "Energy Management Strategy in Consideration of Battery Health for PHEV via Stochastic Control and Particle Swarm Optimization Algorithm," *Energies,* vol. 10, no. 11, 2017.
[34] J. Wu, H. He, J. Peng, Y. Li, and Z. Li, "Continuous reinforcement learning of energy management with deep Q network for a power split hybrid electric bus," *Applied Energy,* vol. 222, pp. 799-811, 2018/07/15/ 2018.
[35] S. Kapturowski, G. Ostrovski, J. Quan, R. Munos, and W. Dabney, "Recurrent experience replay in distributed reinforcement learning," in *International conference on learning representations*, 2018.
[36] C. Nayar, T. Markson, and W. Suponthana, "Wind/PV/diesel micro grid system implemented in remote islands in the Republic of Maldives," in *2008 IEEE International Conference on Sustainable Energy Technologies*, 2008, pp. 1076-1080.
[37] "https://www.westernpower.co.uk/innovation/pod/dataset/data-licences," ed.
[38] P. R. Mendes, L. V. Isorna, C. Bordons, and J. E. J. J. o. P. S. Normey-Rico, "Energy management of an experimental microgrid coupled to a V2G system," vol. 327, pp. 702-713, 2016.
[39] B. Sah, P. Kumar, and S. K. J. I. S. J. Bose, "A fuzzy logic and artificial neural network-based intelligent controller for a vehicle-to-grid system," 2020.
[40] N. I. Nimalsiri, E. L. Ratnam, D. B. Smith, C. P. Mediwaththe, and S. K. J. I. T. o. I. T. S. Halgamuge, "Coordinated Charge and Discharge Scheduling of Electric Vehicles for Load Curve Shaping," 2021.
[41] H. Ko, S. Pack, and V. C. M. Leung, "Mobility-Aware Vehicle-to-Grid Control Algorithm in Microgrids," *IEEE Transactions on Intelligent Transportation Systems,* vol. 19, no. 7, pp. 2165-2174, 2018.

Shuangqi Li (S'18) was born in Beijing, China. He received the B.Eng. degree in vehicle engineering from the Beijing Institute of Technology, Beijing, China, in 2018. He is currently pursuing the Ph.D. degree in electrical engineering with the Department of Electronic and Electrical Engineering, University of Bath, Bath, U.K. He was a Research Assistant with the National Engineering Laboratory for Electric Vehicles, Beijing Institute of Technology, Beijing, from 2018 to 2019. His major research interests include the big data analysis, deep-learning algorithm, deep reinforcement

learning algorithm, operation and planning of smart grid systems, hybrid energy storage system energy management, and V2G service.

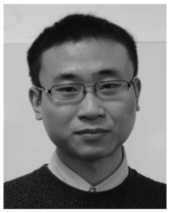


Pengfei Zhao was born in Beijing, China. He received the double B.Eng. degree from the University of Bath, Bath, U.K. and North China Electric Power University, Baoding, China, in 2017, and the Ph.D. degree from the University of Bath, Bath, U.K., in 2021 He was a Visiting Ph.D. Student with the Smart Grid Operations and Optimization Laboratory, Tsinghua University, Beijing, China, in 2019. He is currently an Assistant Professor with the State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing. His major research interests include the multienergy systems, public health emergency management, and smart city management.

Chenghong Gu (M'14) was born in Anhui province, China. He received the Bachelor's degree from the Shanghai University of Electric Power, Shanghai, China, in 2003, and the Master's degree from the Shanghai Jiao Tong University, Shanghai, China, in 2007, both in electrical engineering. He received the Ph.D. degree from the University of Bath, U.K. He is currently a Lecturer and EPSRC Fellow with the Department of Electronic and Electrical Engineering, University of Bath. His major research interest is in multi-vector energy system, smart grid, and power economics.

Jianwei Li was born in Gansu, China. She received the Bachelor's degree in smart grid information engineering from the Xi'an University of Technology, Xi'an, China, in 2018, and the Master's degree in electrical power systems from the University of Bath, Bath, U.K., in 2019. She is currently pursuing the Ph.D. degree in electronic and electrical engineering with the University of Bath, Bath, U.K. Her main research interests include power system planning, analysis, and power system economics.

Shuang Cheng (S'19) received the B.Eng. and MPhil degree in electrical and electronic engineering from North China Electric Power University, Beijing, China in 2017 and 2020, respectively. She received the MSc. degree from the University of Bath, Bath, UK in 2019. She was a visiting scholar in Wroclaw University of Science and Technology. She is currently pursuing the Ph.D. degree with the University of Bath. Her main research interests include energy system planning, power system economics, and market design.



Minghao Xu received the B.Eng. degrees in electrical and electronic engineering from the University of Bath, U.K., and electrical power engineering from North China Electric Power University, Baoding, China, in 2014. He received the Ph.D. degree from the University of Bath, U.K. His research interests include big data, machine learning, and deep learning applications in power systems.